
Measurement of the $t\bar{t}\gamma$ production cross section with the
ATLAS experiment at $\sqrt{s} = 13$ TeV using the full Run 2
data set

Dissertation

zur Erlangung des mathematisch-naturwissenschaftlichen Doktorgrades
„Doctor rerum naturalium“
der Georg-August-Universität Göttingen

im Promotionsprogramm ProPhys
der Georg-August University School of Science (GAUSS)

vorgelegt von

Andreas Achim Kirchhoff

aus Wipperfürth

Göttingen, 2024



Betreuungsausschuss

Prof. Dr. A. Quadt
Prof. Dr. S. Lai
Prof. Dr. S. Schumann

Mitglieder der Prüfungskommission:

Referent: Prof. Dr. A. Quadt
II. Physikalisches Institut, Georg-August-Universität Göttingen
Koreferentin: PD Dr. Carmen Diez Pardos
Fakultät IV, Department Physik, Experimentelle Teilchenphysik, Universität Siegen

Weitere Mitglieder der Prüfungskommission:

Prof. Dr. Stan Lai
II. Physikalisches Institut, Georg-August-Universität Göttingen
Prof. Dr. Steffen Schumann
Institut für Theoretische Physik, Georg-August-Universität Göttingen
Prof. Dr. Ansgar Reiners
Institut für Astrophysik und Geophysik, Georg-August-Universität Göttingen
PD Dr. Ralph Bernhard
II. Physikalisches Institut, Georg-August-Universität Göttingen

Tag der mündlichen Prüfung: 14. März 2024
Referenz: II.Physik-UniGö-Diss-2024/01

*"Der Vorwurf, meine Doktorarbeit sei ein Plagiat, ist abstrus.
Die Anfertigung dieser Arbeit war meine eigene Leistung."*
Baron von Münchhausen (frei nach "Der falsche Kalender" von Marc-Uwe-Kling)

*Für meine verstorbene Mutter Hildegard und meinen verstorbenen Bruder Lars.
Für meinen Vater Hermann, meinen Sohn Camilo und meine Liebe Aless.*

Measurement of the $t\bar{t}\gamma$ production cross section with the ATLAS experiment at $\sqrt{s} = 13$ TeV using the full Run 2 data set

Abstract

With Run 1 of the LHC, an unequalled top-quark factory started its operation. While during Run 1 the focus was on $t\bar{t}$ measurements, the unprecedented size of the ATLAS Run 2 data set of 140 fb^{-1} allows for precise measurements even for rare processes involving top quarks. One of these is the $t\bar{t}\gamma$ process, where a pair of top quarks is produced in association with a photon. This process is sensitive to the electromagnetic coupling of top quarks and hence to the charge of the top quark. Besides the general background of $t\bar{t}$ measurements that come with an additional photon, the largest irreducible background comes from those events where the emission of the photon happens in the decay of the top quarks. This reduces the sensitivity to the top-photon coupling. Hence, it is desirable to differentiate between photons emitted by top quarks (also called the production part) and photons emitted during the decay. Strictly speaking, there is an interference between these two processes but within the narrow-width approximation, which is valid for the $t\bar{t}\gamma$ process, it vanishes. Consequently, this analysis for the first time performs a cross-section measurement only on the $t\bar{t}\gamma$ from production part, where initial state radiation from $q\bar{q}$ initiated events is not decoupled. The inclusive cross-sections are measured in the dilepton and lepton+jets decay channels of the $t\bar{t}$ pair. The separation of $t\bar{t}\gamma$ from production and $t\bar{t}\gamma$ from decay is done with feed-forward neural networks, where a multiclass classification process was implemented for the lepton+jets channel, while in the dilepton channel binary classification was used. The signal strength and consequently the cross-sections are extracted with a profile likelihood fit. The result for the combined phase space of dilepton and lepton+jets channel at particle level is $322^{+16}_{-15} \text{ fb} = 322 \pm 15(\text{syst.}) \pm 5(\text{stat.}) \text{ fb}$ and is in agreement with the standard model prediction. Furthermore, a summary of differential measurements, an EFT interpretation and a combination with the ATLAS $t\bar{t}Z$ measurement are given.

Messung des $t\bar{t}\gamma$ Produktionswirkungsquerschnitts mit dem ATLAS Experiment bei $\sqrt{s} = 13$ TeV unter Nutzung des vollen Datensatzes aus Run 2

Zusammenfassung

Mit Run 1 des LHC begann der Betrieb einer zuvor nie da gewesene Top-Quark-Fabrik. Während der Fokus in Run 1 noch auf $t\bar{t}$ Messungen lag, erlaubt der beispiellose ATLAS-Datensatz von Run 2 (140 fb^{-1}) präzise Messungen auch von seltenen Prozessen, in die Top-Quarks involviert sind. Einer davon ist der $t\bar{t}\gamma$ Prozess, bei dem ein Top-Quark-Paar zusammen mit einem Photon erzeugt wird. Dieser Prozess ist sensitiv auf die elektromagnetische Kopplung von Top-Quarks und damit auch auf die Ladung des Top-Quarks. Neben dem generellen Untergrund von $t\bar{t}$ Messungen mit zusätzlichem Photon stammt der größte nicht reduzierbare Untergrund von solchen Ereignissen, bei denen die Abstrahlung des Photons im Zerfall des Top-Quarks passiert. Dies reduziert die Sensitivität zur Top-Photon-Kopplung. Insofern ist es wünschenswert, zwischen Photonen, die von Top-Quarks abgestrahlt werden (auch Produktionsteil genannt) und solchen aus dem Zerfall, zu unterscheiden. Streng genommen gibt es zwischen diesen Prozessen Interferenzen, aber im Rahmen der Narrow-Width-Approximation, welche für den $t\bar{t}\gamma$ Prozess gültig ist, verschwinden diese. Konsequenterweise führt diese Analyse zum ersten Mal eine Wirkungsquerschnittsmessung aus, die nur den Anteil von $t\bar{t}\gamma$ aus der Produktion betrachtet, wobei Anfangszustandsstrahlung von $q\bar{q}$ initiierten Ereignissen nicht entkoppelt ist. Die inklusiven Wirkungsquerschnitte werden im dileptonischen und Lepton+Jets Zerfallskanal des Top-Quark-Paares gemessen. Die Trennung von $t\bar{t}\gamma$ aus der Produktion und $t\bar{t}\gamma$ aus dem Zerfall erfolgt durch vorwärtsgerichtete neuronale Netze, wobei eine Mehrklassenklassifikation im Lepton+Jets Kanal implementiert wird, während eine binäre Klassifikation im dileptonischen Kanal genutzt wird. Die Signalstärke und konsequenterweise die Wirkungsquerschnitte werden mit Profile-Likelihood-Fits extrahiert. Das Resultat für den kombinierten Phasenraum des dileptonischen und des Lepton+Jets Kanals auf Teilchen-Level ist $322^{+16}_{-15} \text{ fb} = 322 \pm 15(\text{syst.}) \pm 5(\text{stat.}) \text{ fb}$ und steht damit in Übereinstimmung mit der Vorhersage des Standardmodells. Des Weiteren wird eine kurze Zusammenfassung der differentiellen Messungen, der EFT-Interpretation und der Kombination mit der ATLAS $t\bar{t}Z$ Messung gegeben.

Author's contribution

The operation, maintenance and data analysis for experiments like ATLAS is a huge effort that needs contributions from many people. Without all the experts running the detector, assuring data quality and working on the event reconstruction (only to mention a few things), this analysis and dissertation would not have been possible. But even the analysis (as most analyses within ATLAS) itself is not done by a single person but by a whole team. For this analysis, people from the University of Göttingen, the University of Siegen and the University of Dortmund worked together. As a consequence, this means that not all parts of the analysis presented in this thesis have been done exclusively by the author. To avoid referring to internal or private communication, it will be listed here what the author's work in this analysis is, where the work is based on the work of others within the analysis and where the work of others is based on the author's work. The Monte-Carlo (MC) simulation samples used for this analysis were produced centrally by and for the whole ATLAS collaboration. Object definitions were taken from general ATLAS-wide recommendations and in communication with the analysis of the $t\bar{t}Z$ process. Both analyses use coherent samples for their EFT-combination (EFT=effective field theory). From this information, the N-tuples (but the EFT-N-tuples) for this analysis were produced by the author.

From previous ATLAS analyses of the $t\bar{t}\gamma$ -process, the general cuts to define the fiducial phase space were mostly adopted by the analysis [1, 2]. Slight modifications of the b-tagging working point and the ΔR distance between the photon and the nearest lepton were tested by the author, and based on those studies, the analysis team defined the fiducial phase space.

For the background estimations, the author looked at the control regions populated with $W\gamma$ and $Z\gamma$ events (see Section 5.2.1). The derivations of the hadron-fake scale factors, the electron-fake scale factors and the lepton fake scale factors were done by Carmen Diez Pardos and Buddhadeb Mondal from the University of Siegen (see Sections 5.2.2, 5.2.3 and 5.2.4). The author's main work can be found in Chapter 5.3 and Sections 6.2 and 6.3 on the neural network and the inclusive cross-section fits. The results from the neural network training were not only used for the inclusive cross-section fits but also for the differential fits and EFT fits, which were made by Buddhadeb Mondal and Jahn Hahn from Siegen (see Sections 6.4.1 and 6.4.2). The combination with the $t\bar{t}Z$ analysis was made by Tomas Dado (TU Dortmund) and Baptiste Ravina (Georg-August-University Göttingen) (see Section 6.4.3). The conclusions drawn at the end are mainly from the author but they profited a lot from the internal discussions involving the whole analysis team (see Chapter 7).

The work presented in this thesis was also written down in the ATLAS-internal documentation of this analysis. Most of the technical information about simulations, event reconstruction and the analysis strategy, as well as the work not done by the author, is based on the information given in the internal note¹. An official ATLAS paper was also published [3].

¹The internal note can be found here: <https://cds.cern.ch/record/2712912>. It is not publicly accessible.

Contents

1. Introduction	1
2. The standard model of particle physics	3
2.1. The evolution of the standard model of particle physics	3
2.2. Summary of the standard model	6
2.2.1. The particle content of the standard model	6
2.2.2. The standard model Lagrangian	8
2.3. Limits of the standard model	15
2.4. The proton structure	17
2.5. The top quark within the standard model	18
2.5.1. Top quark production and decay at the LHC	19
2.5.2. Top quark properties	22
2.6. Top quark pair production in association with a photon	27
2.6.1. General introduction to top quark cross-section measurements with associated boson production	27
2.6.2. The top-photon coupling	28
2.6.3. Probing the top-photon coupling	30
3. Experimental setup	33
3.1. CERN and the LHC	33
3.1.1. CERN	33
3.1.2. LHC	33
3.2. The ATLAS detector	36
3.2.1. The Inner Detector	37
3.2.2. The electromagnetic calorimeter	40
3.2.3. The hadronic calorimeter	41
3.2.4. The muon spectrometer	42
3.2.5. The ATLAS trigger system	43
3.2.6. Luminosity measurements	44

4. From data and MC simulations to the final event selection	45
4.1. Data	45
4.2. Monte Carlo simulations	45
4.2.1. MC samples	46
4.2.2. Sample overlap removal	49
4.3. Object definitions	50
4.3.1. Photons	51
4.3.2. Electrons	51
4.3.3. Muons	52
4.3.4. Jets	52
4.3.5. E_T^{miss}	53
4.3.6. Object overlap removal	53
4.4. Event reconstruction and selection	54
4.4.1. Photon categorisation	55
4.4.2. Event categorisation	55
4.4.3. Final event selection	56
4.5. Fiducial phase space at particle level	59
5. Analysis Strategy	61
5.1. General outline	61
5.2. Background modelling	62
5.2.1. Prompt photon background	62
5.2.2. $e \rightarrow \gamma$ fake background	69
5.2.3. $h \rightarrow \gamma$ fake background	72
5.2.4. Lepton fake background	75
5.3. Neural network	78
5.3.1. Introduction to neural networks	79
5.3.2. Event classification in the lepton+jets channel	81
5.3.3. Event classification in the dilepton channel	97
5.4. Systematic uncertainties	108
5.4.1. Theoretical uncertainties	109
5.4.2. Experimental uncertainties	111
5.4.3. Background normalisation uncertainties	113
6. Results	115
6.1. Profile likelihood fits and inclusive cross-section measurements	115
6.2. Tests of the fit setup	117
6.2.1. Asimov tests	118
6.2.2. Tests for fits to data	128
6.3. Final fit and inclusive cross-section extraction	139
6.4. Further results	146
6.4.1. Differential cross-sections	146
6.4.2. EFT interpretation	148
6.4.3. $t\bar{t}Z$ - $t\bar{t}\gamma$ combination	150

7. Conclusions	153
Bibliography	159
Appendices	179
A. Monte Carlo Samples	181
B. Additional information on the top quark reconstruction	185
B.1. Quadratic equation for neutrino p_z	185
B.2. Plots on top-quark reconstruction	186
C. Additional plots on prompt photon background control regions	189
C.1. $W\gamma$ control region	189
C.2. $Z\gamma$ control region	193
D. Additional plots on the neural network training	197
D.1. L+jets channel	197
D.1.1. Input distributions	197
D.1.2. Separation plots	205
D.1.3. Negative weight comparisons	209
D.1.4. Comparison of training and testing output distributions	214
D.1.5. Permutation importance	217
D.1.6. Output correlation plots	221
D.2. Dilepton channel	225
D.2.1. Input distributions	225
D.2.2. Separation plots	228
D.2.3. Negative weight comparisons	230
D.2.4. Comparison of training and testing output distributions	233
E. Additional material on inclusive cross-section fits	235
E.1. Additional material on Asimov fits	235
E.2. Additional material on unblinded fits	240
E.3. Post-fit plots	257
E.3.1. $t\bar{t}\gamma$ production: Post-fit plots for the fit in the lepton+jets channel with two NF	257
E.3.2. Inclusive $t\bar{t}\gamma$: Post-fit plots for the fit in the lepton+jets channel with one NF	261
E.3.3. $t\bar{t}\gamma$ production: Post-fit plots for the fit in the dilepton channel with two NF	266
E.3.4. Inclusive $t\bar{t}\gamma$: Post-fit plots for the fit in the dilepton channel with one NF	268

F. Results on differential cross-sections	271
F.1. Absolute cross-sections	271
F.1.1. Lepton+jets channel	271
F.1.2. Dilepton channel	272
F.2. Normalized cross-sections	273
F.2.1. Lepton+jets channel	273
F.2.2. Dilepton channel	274
F.3. χ^2 and p -values	275
G. EFT results and combination with $t\bar{t}Z$	277
G.1. EFT	277
G.2. Combination with $t\bar{t}Z$	278
Danksagung	279

CHAPTER 1

Introduction

The standard model of particle physics (SM) [4, 5] is among the most successful theories in physics, making precise predictions about all known elementary particles and their interactions (see e.g. [6]). It emerged out of several decades of intensive research. The theory was completed at the beginning of the 1970s with the proof of its renormalisability [7–9], the development of QCD to describe strong interactions and coloured particles (quarks and gluons) [10–12] and the prediction of the third generation to allow charge-parity (CP) violation within the theory [13]. Experimentally, the last piece that was missing was the Higgs boson. It was discovered in 2012 at the Large Hadron Collider (LHC) by the experiments ATLAS and CMS [14, 15].

So why is research still needed in particle physics? Even before the discovery of the Higgs boson, it was known that there are experimentally well-established effects that cannot be explained by the SM, among them the questions on the nature of dark matter and dark energy as well as the matter-antimatter asymmetry in the universe (see Section 2.3). While all these observations were made in astrophysics [16–20], the answer is believed to be found in particle physics. That is why one of the main motivations to build the LHC, besides the search for the Higgs boson, was the search for physics beyond the standard model (BSM)(see e.g. [21–25]) to hopefully be able to explain the mentioned observations. But no signal of new physics was found so far [5]. A few results show mild tensions but there is no smoking gun signature telling us where to look for new physics. As a consequence of this, there are two possibilities: either the physics processes produced in pp collisions at the LHC are not sensitive to BSM physics, or the analyses need to get more precise.

Assuming that the latter is true, it is necessary to reduce systematic uncertainties and redo analyses that were already done at lower centre-of-mass energies or with lower statistics. One of these processes one can look at is the production of a pair of top quarks (one top quark and one anti-top quark) in association with a photon ($t\bar{t}\gamma$). This

1. Introduction

process can be used to study the coupling between the top quark and the photon, as it is sensitive to the charge of the top quark, the coupling strength and the topological structure of the coupling (see Section 2.6). However, only photons emitted during the production stage of the process are sensitive to the top-photon coupling. While photons emitted by decay products of the top quarks are not sensitive to the coupling, those events form an irreducible background. Previous analyses [1, 2, 26–30] of this process did not take this into account and measured both contributions (photons from production and decay) together. This new analysis [3] of the $t\bar{t}\gamma$ -process will differentiate for the first time these two origins, using a machine learning technique where deep neural networks (DNN) are used to classify signal and background (see Chapter 5.3). The implementation of it was done by the author of this thesis. It will be shown that this approach is valid, and fiducial (inclusive and differential) cross-sections can be measured with similar or even improved sensitivity compared to previous analyses (see Chapters 6 and 7).

The thesis is structured as follows: the next chapter gives a theoretical introduction into the topic. Chapter 3 describes CERN, the LHC and the ATLAS experiment, while Chapter 4 explains how events are selected based on the reconstructed events. Afterwards the analysis strategy (Section 5.1), the background modelling (Section 5.2), the usage of neural networks (Section 5.3) and the treatment of systematic uncertainties (Section 5.4) are explained. The results are shown in Chapter 6 and Chapter 7 gives a conclusion and an outlook.

The standard model of particle physics

The first section of this chapter is devoted to giving a historical introduction to how the standard model of particle physics was developed. Up to a certain level, the selected experimental and theoretical breakthroughs given here and the mentioned people who were part of this development are selected based on the subjective opinion of the author. Because this should be just a short overview and not a historical work, it is also impossible to name all the people who made contributions to the actual development and how the different contributions are linked in detail, as well as how some first misunderstandings of results were overcome.

The second section is then a concise, mathematical description of the standard model. Section 2.3 points out the limits of the current knowledge, while Section 2.4 describes the theoretical basis of how the LHC allows us to do particle physics. Sections 2.5 and 2.6 are used to present the theory details behind the actual thesis.

2.1. The evolution of the standard model of particle physics

In 1897 the first elementary particle, the electron, was found (although it was unknown at the time that the electron was an elementary particle) by J.J. Thomson [31], performing experiments with so-called cathode rays. At the time, there was no quantitative microscopic theory of matter and its interactions. The first step towards such a theory was made by Planck around the year 1900 in solving the black body radiation problem, assuming that energy only comes in quantised portions [32]. Generally, this is seen as the first move towards quantum mechanics. Although the electron was known, the structure of matter was still unknown, as well as how to interpret the quantised energy portions. The next important steps were made by Rutherford and his collaborators Geiger and Marsden on the structure of matter, and by Einstein on the energy quantisation. From the experimental results obtained by Geiger and Marsden [33], Rutherford could develop

2. The standard model of particle physics

an atomic model where the atomic mass is concentrated in a small, positively charged nucleus while the electrons orbit it [34]. An explanation of how these orbits can be stable was given with the development of quantum mechanics in the 1920s (see below). On the other hand, Einstein could give an explanation of the photoelectric effect by using Planck's quantisation hypothesis, and describing light as made of particles [35]. These particles were later called photons. At the time, this was a strong hypothesis as light, and electromagnetic radiation in general, was described very well by Maxwell's theory as a continuous field.

The further development in understanding the structure of matter and the strange behaviour of light (depending on the experiment it behaves like particles or waves) led to the development of quantum mechanics. While Bohr developed a semi-classical model that could explain the line spectrum of hydrogen by imposing stable electron orbits, to overcome the theoretical problems of a purely classical theory based on Rutherford's model [36–38], the underlying reason for these stable orbits was given by Schrödinger and Heisenberg, Born and Jordan in their respective formulations of a fully quantised non-relativistic theory [39–42].

In 1932 Anderson detected a particle that had the same properties as the electron but was positively charged [43]. It turned out that this particle was the anti-particle of the electron, later called the positron, predicted by Dirac [44], who developed a relativistic version of quantum mechanics and predicted the existence of anti-matter. Later, in 1937, Anderson and Neddermeyer also discovered the muon in cosmic rays [45]. In the 1920s and 1930s, it had also been realised that the atomic nucleus was composed of two sorts of particles. The proton had been discovered in 1919 [46] and the neutron in 1932 [47]. In the 1960s scattering experiments at the Stanford-Linear-Accelerator-Center (SLAC) showed that the proton and neutron were not point-like particles but possessed an inner structure [48–54]. Before that, quite a large number of particles were known. The different types of baryons and mesons were found in cosmic rays and particle accelerators. At first sight, it was unclear how all these particles were related to each other. This question was solved by Gell-Mann, Zweig, and Ne'eman proposing the existence of particles with fractional charges [55–58], which were then called quarks, a term introduced by Gell-Mann.

On the theoretical side, the experimental development and all the discoveries were accompanied by the development of quantum field theory (QFT) and quantum electrodynamics (QED) as the first full QFT to explain the interactions between the different particles, in this case between fermions and photons. These ideas, mainly developed by Feynman, Schwinger and Tomonaga, were later used to describe the weak and strong force with QFTs too¹.

While particles were identified to be fermions having a spin² of 1/2, the vector bosons, possessing spin 1, were identified to exchange the interactions between the particles, e.g.

¹This thesis has not the scope to give a full account of the development of QED and QFT. For this, it is referred to these two books: [59, 60].

²All particles have a property called spin which is effectively an internal angular momentum. It was found by Stern and Gerlach [61].

2.1. The evolution of the standard model of particle physics

the photon as the mediator of the electromagnetic force. As long as these bosons are massless the needed gauge invariance of the underlying quantum field theory is maintained. Without gauge invariance, those theories would not be unitary and hence not able to make any predictions. For massive bosons, this becomes a problem. The solution is the so-called Brout-Englert-Higgs (BEH) mechanism [62–65]. At the time, no massive bosons were known but there had been attempts to explain the strong force via massive bosons [66]. Later, the idea could be applied to the weak interactions. The masses of fermions can also be explained with the BEH mechanism via a Yukawa coupling between the Higgs boson and the fermions.

The electron neutrino, predicted by Pauli to ensure energy conservation in the nuclear β -decay, was found in 1956 by Cowan and Reines [67, 68]. Its existence was used in the construction of a unified theory of the weak and electromagnetic force, proposed by Glashow, Weinberg (implementing the BEH mechanism into the theory), Salam, and Ward [69–71]. Their model predicted three massive bosons. From the experiments of Wu and Goldhaber [72, 73], it was already known that the weak force was violating parity. In 1964, Christenson, Cronin, Fitch, and Turley showed that it was also violating CP-invariance [74]. The idea to have massive bosons, as used in electro-weak unification, was proposed by Yukawa as early as 1935 [66]. In 1962 the μ -neutrino was also found [75]. That the muon and the μ -neutrino form the second lepton generation was understood later. The renormalizability of the electroweak theory was shown by t’Hooft and Veltman [7–9].

In 1963, Cabibbo introduced a matrix to explain the weak decays of hadrons where it was observed that hadrons containing a strange quark could decay into hadrons without strangeness [76]. The angle of this rotation matrix describes the mixing of weak eigenstates to form the observed mass eigenstates that can move through space and time. Based on this work, Glashow, Iliopoulos and Maiani, in 1970, proposed the existence of a fourth quark, the c -quark, to explain the suppression of flavour-changing neutral current decays of kaons³ [78]. This quark was found in 1974 by two independent groups at SLAC and BNL (Brookhaven National Laboratory) [79, 80].

Around the same time, the quark model from the 1960s was put into a full theory called quantum chromodynamics (QCD). With it, the main questions the quark model raised, namely the fact that only hadrons but no bare quarks could be observed and the breakdown of perturbation theory at low energies, could be answered [10–12].

So by the mid-1970s, one had a full electroweak theory⁴ with two generations of leptons and a theory for the strong force with two generations as well (with the gluon still being unobserved). In 1975, the first particle of the third generation, the τ -lepton, was discovered without being predicted [83]. The existence of the corresponding neutrino was inferred more or less immediately. The prediction of a third quark generation was already present since 1973, formulated by Kobayashi and Maskawa. In order to explain CP violation, one condition to explain the matter-antimatter asymmetry observed in the universe [20], they had to extend the matrix introduced by Cabibbo from a 2x2 matrix

³Although they were not the first to predict a fourth quark [77].

⁴The predicted neutral current interactions were seen in 1973 [81, 82].

2. The standard model of particle physics

to a 3x3 matrix [13]. The complex phase that was introduced thereby, does allow CP-violation in weak interactions. From the kaon system, CP violation was already known. It was later also found in hadrons containing b - or c -quarks [84–86]. The predicted b -quark was discovered in 1977 [87] at Fermilab by the E288 experiment, while for the top quark discovery by the CDF and D0 experiments at the Tevatron accelerator (again at Fermilab) one had to wait until 1995 [88, 89]. In 1979 the gluon could be discovered eventually at DESY at the PETRA storage ring with the TASSO experiment [90].

During the 1980s and 1990s, more measurements (e. g. at the Large Electron Positron Collider (LEP) at CERN and the HERA collider at DESY) supported the standard model, especially the direct discovery of the W and Z bosons [91–94] at the SPS/Sp̄pS accelerator with the UA1 and UA2 experiments. The last fermion, the τ -neutrino, was then discovered in 2001 at Fermilab with the DONUT experiment [95]. The last missing piece was the Higgs boson, which was eventually discovered in 2012 [14, 15] by the ATLAS and CMS experiments at the LHC at CERN.

Since then, there has been a self-consistent theory supported by experimental evidence. Unfortunately, we know of observations that cannot be explained by this theory. These will be presented in Section 2.3. The next section will give a full account of the SM.

2.2. Summary of the standard model

In the last section, a short description of the development of the SM was given. In this section, the theory will be described in detail.

2.2.1. The particle content of the standard model

The gauge groups

The SM is a renormalisable quantum field theory which is invariant under local gauge transformations. The corresponding gauge groups have the following structure:

$$U_Y(1) \times SU(2)_L \times SU_C(3) . \quad (2.1)$$

Y represents the hypercharge singlets on which the $U(1)$ transformation acts, while L stands for the chiral left-handed weak isospin doublets on which the $SU(2)$ transformation acts. Together, they form the gauge group of electroweak unification. C represents the colour triplets on which the $SU(3)$ transformation acts. It is the gauge group of the strong force. Each gauge transformation leads to a conserved quantum number, which is hypercharge⁵, the third component of the weak isospin and colour charge, respectively.

Gauge bosons

From the gauge groups, one can furthermore deduce the number of so-called bosonic fields, which are the fields that exchange the forces between the different particles. While

⁵Due to electroweak unification, the conserved quantity actually being measured is electric charge.

2.2. Summary of the standard model

$U(1)$ has one generator, this leads to one bosonic field. Respectively, $SU(2)$ has three bosonic fields and $SU(3)$ comes along with eight bosonic fields. Due to electroweak unification, the bosonic field that emerges from the $U(1)$ gauge group mixes with one of the bosonic fields created by the $SU(2)$ group. The mixing between these two fields gives rise to the photon and the Z boson fields. The other two fields from the $SU(2)$ -group also mix and form the bosonic fields that give rise to the two massive W bosons. The $SU(3)$ gauge group of the strong force leads to the existence of eight gluons. All these bosons, appearing as excitations of bosonic fields, have spin 1 and are called vector bosons, as they behave like vectors under parity transformation. Due to their relation to the gauge groups, they are also called gauge bosons. They act as the force carriers of the three fundamental forces described by the SM, which are the electromagnetic force, the weak force and the strong force. Gravitation, which is the fourth fundamental force known, is not included as there is no quantum theory of it yet.

Fermions

Besides the bosonic fields and their excitation appearing as particles, there are also fermionic fields. These fields describe ordinary matter that interacts with each other by the exchange of bosons. Fermions have half-integer spin, and the fundamental fermions within the SM all have spin $1/2$. There are 12 different types of fermions in the standard model, grouped into two categories called quarks and leptons. Within these groups, three generations are observed, where each generation builds a weak $SU(2)$ doublet. The particle masses increase with each generation. While leptons only interact via the electroweak force and either have an electric charge of ± 1 (electrons, muons, tau-leptons)⁶ or do not have electric charge (neutrinos)⁷, quarks do also interact via the strong force. They have fractional charges of $+2/3$ or $-1/3$. Independently of being a lepton or a quark, the particles with the third component of the weak isospin being $+1/2$ are put at the top of the weak isospin doublets (called up-type), while those with $-1/2$ appear at the bottom (called down type). The up-type quarks (up, charm and top) have an electric charge of $+2/3$, while the down-type quarks (down, strange, bottom) have an electric charge of $-1/3$. For the leptons, the neutrinos are the up-type particles, and the charged leptons are the down-type ones. In addition, to each fermion corresponds an antiparticle with opposite charge⁸.

Higgs boson and summary

The last particle is the scalar Higgs boson. It has none of the aforementioned charges. Its existence gives rise to the masses of the particles via the BEH mechanism and assures the gauge invariance and renormalisability of the standard model. In other words, without the Higgs boson, the world of elementary fields/particles could not be described with

⁶They can interact with all four gauge bosons of the electroweak force.

⁷They only interact via the W and Z bosons.

⁸For neutrinos, it is not clear whether they are their own antiparticles (Majorana fermions) or not (Dirac fermions).

2. The standard model of particle physics

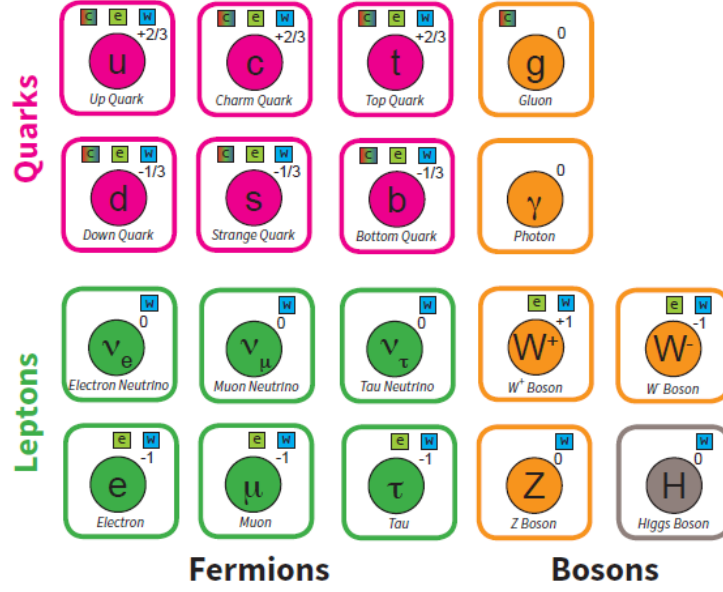


Figure 2.1.: The figure shows the particle content of the SM. The quarks are marked violet, the leptons green, the gauge bosons orange and the Higgs boson grey. For each particle it is marked which charges they carry (c=colour-charge, w=weak-charge and e=electric-charge). Furthermore, the value of the electric charge is shown. The figure was taken from [96] under the CC BY-NC-ND 3.0 DEED license⁹.

a quantum field theory. In Figure 2.1 the particle content of the standard model is summarised.

2.2.2. The standard model Lagrangian

So far, only the different fields/particles and their general properties were listed. Nothing was said about how they interact. The interactions are described, as usual in field theories, by a Lagrangian. To construct the Lagrangian of the SM, it makes sense to go through the different forces described by the SM and construct their Lagrangians first, before formulating the full theory.

⁹The license terms can be found here: <https://creativecommons.org/licenses/by-nc-nd/3.0/>. No changes were made.

QED

From the Dirac equation [44] the free Lagrangian of any fermion with the mass m is given by¹⁰

$$\mathcal{L}_{\text{free}} = \bar{\Psi}(\gamma^\mu \partial_\mu - m)\Psi, \quad (2.2)$$

where Ψ is the spinor describing the fermion, $\bar{\Psi}$ the adjoint spinor and γ^μ are the Dirac matrices. As in classical electromagnetism, the interaction with the electromagnetic field can be achieved by adding the potential A_μ like this:

$$\mathcal{L}_{\text{interaction}} = \bar{\Psi}(\gamma^\mu \partial_\mu - \gamma^\mu \alpha_{em} Q A_\mu - m)\Psi, \quad (2.3)$$

with Q being the charge of the fermion and α_{em} the finestructure constant. Classical electromagnetism is invariant under the following gauge transformation:

$$A_\mu \rightarrow A_\mu - \partial_\mu \chi(x), \quad (2.4)$$

where $\chi(x)$ is an arbitrary function of space-time. Introducing this transformation to relativistic quantum theory, the spinor Ψ must transform in the following way to achieve gauge invariance:

$$\Psi \rightarrow e^{i\alpha_{em} Q \chi(x)} \Psi. \quad (2.5)$$

This transformation is called a local gauge transformation. These types of transformations form Lie groups and are also referred to as gauge groups. In this case, the gauge transformation is a local U(1) phase transformation. Because U(1) is Abelian it follows that there is no direct interaction between the gauge bosons, in this case, the photons. The dynamics of the electromagnetic field also called the gauge field, and which is a massless spin 1 vector field, is added to the Lagrangian in a gauge invariant way using the electromagnetic field tensor $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$. In the quantum theory, A_μ is the photon field. Furthermore, one needs one interaction Lagrangian for each charged fermion. So the full QED-Lagrangian becomes:

$$\mathcal{L}_{\text{QED}} = \sum_n (\bar{\Psi}_n(\gamma^\mu \partial_\mu - \gamma^\mu \alpha_{em} Q_n A_\mu - m_n)\Psi_n) - \frac{1}{4} F_{\mu\nu} F^{\mu\nu}, \quad (2.6)$$

where the sum runs over all charged fermion species.

QCD

The Lagrangian of QCD is given by:

$$\mathcal{L}_{\text{QCD}} = \sum_q (\bar{\xi}_{q,c}(\gamma^\mu D_\mu - m_q)\xi_{q,c}) - \frac{1}{4} G_{\mu\nu}^i G^{i\mu\nu}, \quad (2.7)$$

¹⁰Throughout this thesis, the convention is used that repeated indices are summed (Einstein notation).

2. The standard model of particle physics

where the summation index q stands for the six quark flavours and c for the three colour charges quarks can have. The index i stands for the eight gauge fields necessary in QCD, and the covariant derivative D_μ contains the interaction potential. At first sight, this Lagrangian looks similar to the QED Lagrangian, but the small differences in the notation have huge observational consequences, which will be outlined in the following. In contrast to QED, the gauge group of QCD is SU(3). This leads to eight group generators instead of one, and the gauge transformation becomes

$$\xi \rightarrow e^{i\chi_i(x)T_i}\xi, \quad (2.8)$$

where $\chi_i(x)$ are eight arbitrary functions in space time and T_i are the group generators, described by the Gell-Mann matrices $\lambda_i = 2T_i$. The interaction potential described by the eight gauge fields is given as

$$D_\mu = \partial_\mu + ig_s T_i G_\mu^i, \quad (2.9)$$

where G_μ^i are the gluon fields and g_s the strong coupling. For the gauge transformation of the gluon fields one finds

$$G_\mu^i \rightarrow G_\mu^i - \frac{1}{g_s} \partial_\mu \chi_i(x) - f_{ijk} \chi_j G_\mu^k, \quad (2.10)$$

with f_{ijk} being the so-called structure constants. These are derived from the commutator relations of the group generators, as the SU(3) group is a non-Abelian group:

$$[T_i, T_j] = if_{ijk} T_k. \quad (2.11)$$

This non-Abelian character of the strong force also changes how the field strength tensor is described and introduces a self-coupling among the gauge fields:

$$G_{\mu\nu}^i = \partial_\mu G_\nu^i - \partial_\nu G_\mu^i - g_s f_{abc} G_\mu^b G_\nu^c. \quad (2.12)$$

This self-coupling is the reason for the specific behaviour of QCD [10–12]. Due to renormalisation, the coupling changes with the energy scale. This running coupling is nothing special, but in contrast to QED, where the coupling increases with the energy scale¹¹, for QCD, it decreases. On the one hand, this means the higher the energy, the more the quarks behave as free particles. This is called asymptotic freedom. On the other hand, at low energies, the coupling increases and perturbative QCD collapses at around 1 GeV. This leads to a property called confinement. The lower the energy, the stronger the force between quarks, which binds them together in colourless particles, like mesons, baryons and the recently discovered tetra- and penta-quarks [5]. The formation of these particles, starting with a single quark or gluon, is called hadronisation and is only described by phenomenological models, like the Lund string model [97] or cluster models [98]. The huge amount of hadrons produced during hadronisation and their subsequent decays leads to the formation of jets emerging from high energy QCD-interactions.

¹¹Which leads to a breakdown of perturbative QED at very high energies.

Electroweak unification

So far, there were two aspects not mentioned explicitly but assumed implicitly. One aspect is the fact that both, photons and gluons, the gauge fields of QED and QCD, are massless. On the one side, this is an observational fact, but from the theoretical side, the question remains: why is it the case? The answer is that massive gauge bosons spoil gauge invariance. The second assumption, based on observations, is C- and P-invariance. It can be shown that a renormalizable relativistic quantum field theory needs to be CPT invariant [99–101], but there is no need that the individual transformations (P=parity, C=charge conjugation and T=time reversal) leave the Lagrangian invariant. It turns out that the weak force is maximally parity violating [72, 73] with a vector-axialvector structure. Hence, the Lagrangian of the weak force needs to differentiate between left-handed and right-handed chiral states. Furthermore, the charged gauge bosons of the weak force need to be massive¹². From $\beta^\pm$ -decays, it is clear that the gauge bosons of the weak force must have a charge. Therefore, they are called $W^\pm$. Due to the massiveness of the two expected gauge bosons, the process $e^+e^- \rightarrow W^+W^-$ violates unitarity at high energies, meaning that the cross-section rises to infinity. This can be cured with a third neutral boson and a SU(2) gauge symmetry [69–71]. Taking the SU(2) symmetry and its generators, the three Pauli matrices $\sigma_{1,2,3}$, the fermion spinors ordered in doublets ϕ need to transform like this:

$$\phi \rightarrow e^{\frac{1}{2}ig_W\chi_i(x)\sigma_i}\phi. \quad (2.13)$$

The doublets contain only left-handed chiral states, while right-handed chiral states stay unaffected by the gauge transformation and are written in singlets. The doublets are the following ones, containing all twelve fermions:

$$\begin{pmatrix} \nu_e \\ e \end{pmatrix}_L, \begin{pmatrix} \nu_\mu \\ \mu \end{pmatrix}_L, \begin{pmatrix} \nu_\tau \\ \tau \end{pmatrix}_L, \begin{pmatrix} u \\ d \end{pmatrix}_L, \begin{pmatrix} c \\ s \end{pmatrix}_L, \begin{pmatrix} t \\ b \end{pmatrix}_L, \quad (2.14)$$

where the up-type particles have the third component of the weak isospin $I_3 = 1/2$, while the down-type particles have $I_3 = -1/2$. The twelve right-handed chiral singlets, which are not listed explicitly, have $I_3 = 0$. Hence, the weak Lagrangian for the fermions (for simplicity just the first generation is shown) $\mathcal{L}_{\text{weak}}$ becomes:

$$\mathcal{L}_{\text{weak}} = (\bar{\nu}_L, \bar{e}_L)\gamma^\mu D_\mu \begin{pmatrix} \nu_L \\ e_L \end{pmatrix} + (\bar{u}_L, \bar{d}_L)\gamma^\mu D_\mu \begin{pmatrix} u_L \\ d_L \end{pmatrix}, \quad (2.15)$$

where the covariant derivative has become:

$$D_\mu = (\partial_\mu + i\frac{g_W}{2}\sigma_i W_\mu^i). \quad (2.16)$$

As the weak force can change the flavour of a fermion, the basic fields W_μ cannot be those observed in nature, but must be linear combinations of them:

$$W_\mu^\pm = \frac{1}{\sqrt{2}}(W_\mu^{(1)} \mp iW_\mu^{(2)}). \quad (2.17)$$

¹²How massive gauge bosons can be incorporated into quantum field theory will be explained in the next section.

2. The standard model of particle physics

The field $W_\mu^{(3)}$ could be the field transferring the observed weak neutral currents [81, 82]. But as this gauge boson is electrically neutral, mixing with an underlying field called B_μ is possible. This field has an underlying U(1) gauge symmetry, being equivalent to QED. The conserved charge is called the weak hypercharge Y , and the coupling strength is called g' . Eventually, the photon and the Z boson emerge from those two fields as a rotation with the so-called Weinberg angle θ_W as follows:

$$A_\mu = +B_\mu \cos(\theta_W) + W_\mu^{(3)} \sin(\theta_W), \quad (2.18)$$

$$Z_\mu = -B_\mu \sin(\theta_W) + W_\mu^{(3)} \cos(\theta_W). \quad (2.19)$$

The rotation must give back the couplings of the Z and the photon to the leptons, where the photon couples to right-handed fermions the same way as to left-handed fermions, while the Z treats left-handed and right-handed fermions in a different way. A second condition must be fulfilled, too: the weak hypercharge Y must combine with the third component of the weak isospin to give back the electric charges of the fermions. These conditions yield

$$Y = 2(Q - I_W^{(3)}), \quad (2.20)$$

$$e = g' \cos(\theta_W) = g_W \sin(\theta_W). \quad (2.21)$$

So the full electroweak Lagrangian looks similar to Equation¹³ 2.15

$$\begin{aligned} \mathcal{L}_{\text{EW-fermion}} = & \sum_{\text{gen}=1,2,3} (\bar{\nu}_L, \bar{l}_L) i \gamma^\mu D_\mu^{(1)} \begin{pmatrix} \nu_L \\ l_L \end{pmatrix} + (\bar{u}_L, \bar{d}_L) i \gamma^\mu D_\mu^{(2)} \begin{pmatrix} u_L \\ d_L \end{pmatrix} \\ & + \bar{u}_R i \gamma_\mu D_\mu^{(3)} u_R + \bar{d}_R i \gamma_\mu D_\mu^{(4)} d_R + \bar{l}_R i \gamma_\mu D_\mu^{(5)} l_R + \bar{\nu}_R i \gamma_\mu D_\mu^{(6)} \nu_R, \end{aligned} \quad (2.22)$$

where the sum is over the three fermion generations with their corresponding up-type and down-type particles, and the covariant derivative from Equation 2.16 has become

$$D_\mu^{(1)} = \partial_\mu + \frac{ig_W}{2} \sigma_i W_\mu^i + \frac{ig'}{2} B_\mu, \quad (2.23)$$

$$D_\mu^{(2)} = \partial_\mu + \frac{ig_W}{2} \sigma_i W_\mu^i + \frac{ig'}{6} B_\mu, \quad (2.24)$$

$$D_\mu^{(3)} = \partial_\mu + \frac{2ig'}{3} B_\mu, \quad (2.25)$$

$$D_\mu^{(4)} = \partial_\mu + \frac{ig'}{3} B_\mu, \quad (2.26)$$

$$D_\mu^{(5)} = \partial_\mu + ig' B_\mu, \quad (2.27)$$

$$D_\mu^{(6)} = \partial_\mu, \quad (2.28)$$

¹³Neglecting the mass terms for fermions and bosons. For those, see the section on the BEH mechanism.

2.2. Summary of the standard model

depending on the way how the fermions couple to the electroweak gauge fields¹⁴. The field strength tensors are given by

$$W_{\mu\nu} = \partial_\mu W_\nu - \partial_\nu W_\mu - g_W \epsilon_{ijk} W_\mu^j W_\nu^k, \quad (2.29)$$

$$B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu, \quad (2.30)$$

with ϵ_{ijk} being the Levi-Civita-tensor (structure constants of SU(2)) and the gauge term of the Lagrangian is

$$\mathcal{L}_{\text{EW-gauge}} = -\frac{1}{4}(W_{\mu\nu}^i W^{i\mu\nu} + B_{\mu\nu} B^{\mu\nu}). \quad (2.31)$$

$$(2.32)$$

The Brout-Englert-Higgs-mechanism

To add masses to the standard model, two issues need to be resolved. One is that gauge bosons mass terms like $m_A^2 A_\mu A^\mu$ for any gauge field A_μ with mass m_A are not gauge invariant. For QED and QCD, this is not a problem, as the photon and the gluons are massless. This is different for the W and Z bosons. The second issue are fermion masses. While these are in principle possible in QED and QCD, they would spoil gauge-invariance in the electroweak part of the SM as presented above. The reason for this is the different coupling of right-handed and left-handed chiral states of the particles to the massive W and Z bosons. Both problems can be cured with the Brout-Englert-Higgs-mechanism (BEH-mechanism) [62–65], predicting the existence of a scalar boson, the Higgs boson. The starting point for this mechanism within the SM is a weak isospin doublet, with two complex scalar fields, giving four degrees of freedom¹⁵:

$$\phi = \begin{pmatrix} \phi_{\text{up}} \\ \phi_{\text{down}} \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix}. \quad (2.33)$$

The free Lagrangian for this field is

$$\mathcal{L}_{\text{Higgs}} = (\partial_\mu \phi)^\dagger \partial^\mu \phi - \mu^2 \phi^\dagger \phi - \lambda (\phi^\dagger \phi)^2. \quad (2.34)$$

To get the right behaviour to allow for spontaneous symmetry breaking, the potential must fulfil two conditions on the parameters μ^2 and λ , which are $\mu^2 < 0$ and $\lambda > 0$. This potential has infinite minima at $\phi^\dagger \phi = -\frac{\mu^2}{2\lambda} = \frac{v^2}{2}$, with v being the so-called vacuum expectation value. Choosing the so-called unitary gauge and expanding the field around its minima gives

$$\phi = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + h(x) \end{pmatrix} \quad (2.35)$$

¹⁴In principle, right-handed neutrinos could exist, but they would not couple to any gauge bosons.

¹⁵The up-type field has charge +1, and the down-type field is neutral.

2. The standard model of particle physics

where $h(x)$ is the expansion around the minima. Taking the covariant derivative of the electroweak model

$$D_\mu = \partial_\mu + \frac{ig_W}{2}\sigma_i W_\mu^i + \frac{ig'}{2}B_\mu \quad (2.36)$$

and writing out the resulting Lagrangian

$$\mathcal{L}_{\text{Higgs}} = (D_\mu \phi)^\dagger D^\mu \phi - \mu^2 \phi^\dagger \phi - \lambda(\phi^\dagger \phi)^2 \quad (2.37)$$

would lead to terms that are quadratic concerning the gauge boson fields. These terms are the mass terms of the gauge boson fields:

$$m_W = \frac{1}{2}g_W v, \quad (2.38)$$

$$m_Z = \frac{1}{2}v\sqrt{g_W^2 + g'^2}. \quad (2.39)$$

The fermion masses can be included via Yukawa couplings between the fermions and the Higgs field. Due to the weak isospin doublet structure, the Yukawa terms are slightly different for up-type and down-type particles:

$$\mathcal{L}_{\text{mass-down-type}} = -\frac{\sqrt{2}m_f}{v} \left((\bar{\psi}_{\text{up}} \bar{\psi}_{\text{down}})_L \begin{pmatrix} \phi_{\text{up}} \\ \phi_{\text{down}} \end{pmatrix} \psi_{\text{down,R}} + \bar{\psi}_{\text{down,R}} (\phi_{\text{up}}^* \phi_{\text{down}}^*) \begin{pmatrix} \psi_{\text{up}} \\ \psi_{\text{down}} \end{pmatrix}_L \right), \quad (2.40)$$

$$\mathcal{L}_{\text{mass-down-type}} = \frac{\sqrt{2}m_f}{v} \left((\bar{\psi}_{\text{up}} \bar{\psi}_{\text{down}})_L \begin{pmatrix} -\phi_{\text{down}}^* \\ \phi_{\text{up}}^* \end{pmatrix} \psi_{\text{up,R}} + \bar{\psi}_{\text{up,R}} (-\phi_{\text{down}} \phi_{\text{up}}) \begin{pmatrix} \psi_{\text{up}} \\ \psi_{\text{down}} \end{pmatrix}_L \right). \quad (2.41)$$

CP-violation in the standard model

From theoretical considerations it is known that there must be a source of CP violation in our universe [20]. Since 1964, it is known that there is CP violation in the weak decays of kaons [74]. To introduce CP-violation into the SM, it was found that this can be done with a matrix. This matrix is called the CKM matrix¹⁶ [13, 76]. It describes the relation between the weak eigenstates and the physical mass eigenstates that are observed. Within the SM, such a mixing only occurs in the quark sector¹⁷. To get a CP-violating phase introduced into the SM, at least a 3x3 matrix is needed. As this matrix is unitary it has four degrees of freedom: three angles and the CP-violating phase. So far, it was supposed that the weak eigenstates are the same as the physical mass eigenstates. As experimental observation shows, this is not true, but only affects those couplings that change the flavour structure, which are the couplings of the W bosons to the quarks. So taking the quark-sector from Equation 2.23 that is sensitive to flavour mixing, gives

$$\mathcal{L}_W = \sum_{\text{gen}=1,2,3} (\bar{u}_L, \bar{d}_L) \gamma^\mu D_\mu^{(2)} \begin{pmatrix} u_L \\ d_L \end{pmatrix}, \quad (2.42)$$

¹⁶CKM stands for Cabibbo-Kobayashi-Maskawa, the three inventors of that matrix.

¹⁷In the SM, neutrinos have no masses, and a flavour mixing cannot occur.

2.3. Limits of the standard model

with

$$D_\mu'^{(2)} = \frac{19W}{2} \sigma_i W_\mu^i, \quad (2.43)$$

where the index i only runs from 1 to 2, so only the fields $W_\mu^{(1)}$ and $W_\mu^{(2)}$ are of any interest. Expressing $\mathcal{L}_W$ in terms of the physical fields $W^\pm$ and taking the mixing into account gives

$$\mathcal{L}_W = -\frac{gW}{\sqrt{2}} \left(\bar{d}_L \gamma^\mu V^\dagger u_L W_\mu^- + \bar{u}_L \gamma^\mu V d_L W_\mu^+ \right), \quad (2.44)$$

where u_L and d_L are now vectors containing the three mass eigenstates and V being the CKM-matrix.

The full Lagrangian of the standard model

From the previous sections, one can construct the full SM Lagrangian $\mathcal{L}_{\text{SM}}$. For that, we need to take the expression from Equation 2.7 (ignoring the fermion mass terms), Equation 2.23, 2.37 and 2.41, while taking into account the changes Equation 2.42 introduces to Equation 2.23. The resulting SM Lagrangian can be divided into four independent parts as follows:

$$\mathcal{L}_{\text{SM}} = \mathcal{L}_G + \mathcal{L}_f + \mathcal{L}_\phi + \mathcal{L}_{\text{Yuk}}$$

where $\mathcal{L}_G$ describes the dynamics of the gauge fields, $\mathcal{L}_f$ the dynamics of the fermion fields and how they interact with each other via the gauge fields, $\mathcal{L}_\phi$ describes the Higgs potential and dynamics (including the interactions with the massive gauge fields) and $\mathcal{L}_{\text{Yuk}}$ describes the Yukawa coupling between the fermions and the Higgs field (excluding the neutrinos¹⁸).

2.3. Limits of the standard model

The SM, as described so far, is currently the best theory describing the interactions of the most fundamental entities known in nature. These are the quantum fields, whose excitations are called elementary particles. So far, this theory only describes three of the four known forces in nature, the electromagnetic, the weak and the strong force. Gravity is not included, as it turned out that quantising general relativity does not work with the common procedures to quantise classical fields. This is one of the shortcomings of the theory. It does not directly impact experimental particle physics, as so far, there are no observations that need a quantum version of gravity. But there are a few striking observations, some of them known for several decades, which cannot be explained by the standard model, which will be briefly outlined in this section. Interestingly, all these observations are from astrophysics/cosmology. There are more unsolved questions related to the SM than being discussed here, but most of them do not contradict experimental observations and hence are more of a “cosmetic” issue, e. g. the hierarchy problem.

¹⁸For neutrinos masses please look at Section 2.3.

2. The standard model of particle physics

Dark matter

The first hint for dark matter was found by Fritz Zwicky, observing the redshift of galaxies in the Coma-cluster [18]. As telescopes got better over time, it became possible to observe the rotational velocities of galaxies concerning the radial distance to the galaxy's centres [19]. The observed curves did reveal that there must be a large amount of matter not sending out light or general relativity would be wrong on galactic scales. In addition, the standard cosmological model predicts around 25% of the universe's energy content to be due to dark matter [5]. All ideas to explain dark matter with known particles like neutrinos or known astrophysical objects like black holes or brown dwarfs (failed stars) failed. So one needs to modify the SM by adding new, so far undetected, particles (e. g. using supersymmetry [21, 22]) or extra dimensions [23–25], to name only two of the wealth of different ideas. The other option is to modify general relativity to explain the observations (e.g. TeVeS-theory by Jakob Bekenstein [102]). So far, every modification of gravity has been ruled out, and hence the efforts are put into searches for dark matter particles either at particle colliders or with fixed target experiments, like Xenon1T [103] or DAMIC [104], to just name two of the numerous existing experiments.

Dark energy

In 1998/99 the accelerated expansion of the universe was detected [16, 17]. The force behind this could be a so-called cosmological constant, Λ , a constant that Einstein introduced into his field equations to get a static universe. As soon as the expansion of the universe was detected in 1929 [105], and thoroughly established in the following decades, this idea was dropped. With the detection of the accelerated expansion, it was reintroduced. However it is not yet clear whether a cosmological constant is the force behind the accelerated expansion or not. There are several alternatives [5], and hence the force is called dark energy. In the standard Λ -CMD (cold dark matter) model of the universe, dark energy is responsible for around 70% of the universe's energy content (leaving around 5% for ordinary SM matter).

CP-violation

As already discussed in Section 2.2.2, there must be CP-violation in our universe, and the SM has a mechanism of CP-violation in the flavour-changing interactions of the W bosons. But the amount of CP-violation observed so far does not explain the amount of baryonic matter seen in the universe [5]. So, there must be other CP-violating processes not described by the SM. The strong force theoretically would allow a term in its Lagrangian that would cause CP-violation, but the corresponding phase was measured to be 0 within current uncertainties [106, 107]. A possibility is the neutrino sector but also charged lepton flavour violation or SUSY models allowing flavour-changing neutral currents are being investigated, but so far with no positive result [5].

Neutrino masses

Since the discovery of neutrino oscillations [108] in 1998, it is known that neutrinos are not massless as assumed in the SM. However, it is unclear how the mass-generating mechanism works. In principle, there could be right-handed neutrinos, which would generate the neutrino masses with the left-handed neutrinos via the known Yukawa coupling of fermions to the Higgs field. Otherwise, they would not interact with any of the known gauge bosons and are therefore called sterile. But there are other options. As neutrinos do not have charge, the possibility exists that they are so-called Majorana fermions and not Dirac fermions, which means they are their antiparticles [109]. In principle, one could add Majorana mass terms with right-handed neutrinos to the SM-Lagrangian

$$\mathcal{L}_{\nu\text{-mass}} = -\frac{1}{2}M(\bar{\nu}_R^c\nu_R + \bar{\nu}_R\nu_R^c), \quad (2.45)$$

where ν_R^c is the CP-transformed neutrino spinor. The so-called seesaw mechanism [110] introduces Dirac- and Majorana mass terms to the SM-Lagrangian, predicting that each neutrino flavour has two mass states, a very light one and a very heavy one, where the light one is mostly left-handed, while the heavy one is mostly right-handed and therefore almost sterile. For this to be more than a hypothesis, the Majorana nature of neutrinos needs to be confirmed. So far, searches for neutrinoless double β decays have not revealed any hints of a Majorana nature of neutrinos. The strongest limits come from the KamLAND-Zen [111] and GERDA [112] experiments. But it could also be that the SM is an effective theory only valid on energy scales below a certain threshold, as Fermi theory [113] was an effective theory for the weak force. It allows the introduction of non-renormalisable terms to the SM-Lagrangian, which could create neutrino masses [5]. Beside the question of how neutrinos get masses, the flavour oscillations make it necessary to introduce a matrix similar to the CKM matrix. Again, as only three neutrino generations are known, this is a unitary 3x3 matrix, giving four degrees of freedom, three angles and one phase. In the case of neutrinos, being a mixture of Dirac and Majorana fermions, the matrix would have six angles and three phases, one Dirac phase and two Majorana phases. This matrix is called PMNS¹⁹-matrix [114, 115]. If the phase (or the phases) is (are) non-zero, this means that there is also CP-violation in the lepton sector. So far, the Dirac phase is poorly constrained, but the measurement hints towards being non-zero [116, 117].

2.4. The proton structure

The LHC is a pp collider, and protons are not elementary particles, but hadrons. Consequently, when bunches of protons collide, the interactions happen between the elementary constituents of the proton, the so-called partons. The proton itself is made up of three quarks, two up-quarks and a down-quark (the so-called valence quarks), which are

¹⁹Pontecorvo-Maki-Nakagawa-Sakat

2. The standard model of particle physics

bound together by the strong force, exchanging gluons. In addition, the gluons can split into $q\bar{q}$ pairs (sea-quarks), which annihilate back into a gluon. When protons collide, the actual interactions happen between the valence quarks, the sea-quarks and the gluons of the colliding protons. The resulting problem is that a priori it is not known, what energy and momentum the colliding particles have. The distribution of momentum within the proton is described by the so-called parton distribution functions (PDFs). They give the probability of finding a certain parton (quark or gluon) with a certain momentum in case the proton has a given momentum. These functions cannot be calculated as low-energy QCD is not perturbative. They need to be measured at a certain energy scale. But once that is done, the functions at higher energies can be calculated using the Dokshitzer-Gribov-Lipatov-Altarelli-Parisi (DGLAP) evolution equations [118–120]. The input measurements to derive the PDFs with the DGLAP evolution equations are from all types of experiments, like fixed target experiments, ep collisions, and other high-energy collisions as they happen at the LHC and did happen at the TEVATRON. Currently, there are several sets of PDFs derived by different collaborations [121–125]. With the knowledge of the PDFs one can calculate the cross-sections $\sigma_{pp \rightarrow X+Y}$ for certain processes like $t\bar{t}$ production using the factorisation theorem²⁰, which leads to following formula [127–129]:

$$\sigma_{pp \rightarrow X+Y} = \sum_{i,j} \int_0^1 dx_i \int_0^1 dx_j f(x_i, \mu_f) f(x_j, \mu_f) \sigma_{ij \rightarrow X+Y}(\sqrt{\hat{s}}, \mu_f, \mu_r), \quad (2.46)$$

where the sum runs over all possible initial state parton combinations, $f(x, \mu_f)$ are the parton distribution functions at a chosen factorisation scale and $\sigma_{ij \rightarrow X+Y}(\hat{s}, \mu_f, \mu_r)$ is the parton level cross-section at the given factorisation scale μ_f and renormalisation scale μ_r . Renormalisation is mathematical procedure to handle divergent integrals appearing in cross-section calculations. Both scales are generally chosen to be half of the transverse mass²¹ of all particles. $\sqrt{\hat{s}}$ is the centre-of-mass energy of the partons i and j , being $\hat{s} = x_i x_j s$ and $\sqrt{s}$ is the centre-of-mass energy of the colliding protons.

2.5. The top quark within the standard model

Within the SM, the top quark is the weak isospin partner of the b -quark (see Figure 2.1), with a predicted charge of $+2/3e$ and the third component of the weak isospin being $+1/2$. Its existence was predicted together with the existence of the b -quark in 1973 by Kobayashi and Maskawa [13]. A third quark generation was necessary to introduce a CP-violating mechanism. While the b -quark was discovered a few years later [87] it took until 1995 to discover the top quark too [88, 89], due to its large mass of around 173.3 GeV [130].

The top quark is of great interest, as it allows to probe the predictions of the SM in ways

²⁰The factorisation theorem is used with great success although it was not proven rigorously for all processes at hadron colliders and breaks down at higher orders [126].

²¹The transverse mass is defined as $m_T = \sqrt{m^2 + p_T^2}$.

2.5. The top quark within the standard model

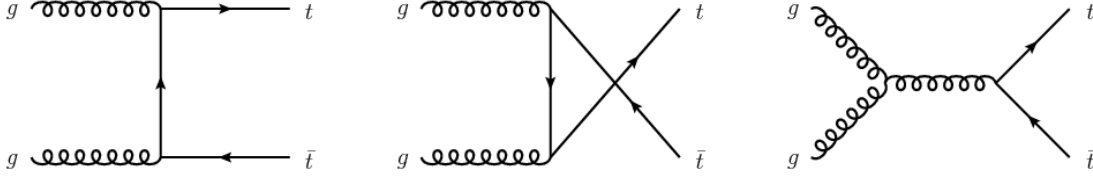


Figure 2.2.: The three leading order Feynman diagrams for $t\bar{t}$ pair production via gluon-gluon fusion. On the left, the t-channel diagram is shown, in the middle the u-channel diagram, and on the right the s-channel diagram.

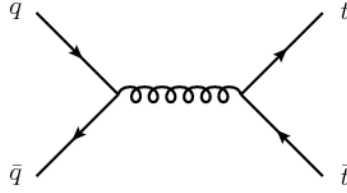


Figure 2.3.: The leading order Feynman diagram for $t\bar{t}$ pair production via quark-antiquark annihilation.

that are not possible with the other quarks. Mostly this is related to its decay width of around 1.35 GeV [131, 132], which prevents the top quark of forming bound states [133]. The large decay width transforms into a lifetime of the order of 10^{-25} s, while the time needed to hadronise is around 10^{-23} s.

2.5.1. Top quark production and decay at the LHC

Top quark pair production and decay

The LHC is a proton-proton collider (see Section 3.1), which means particles collide that have a colour charge (see Section 2.4) and can therefore interact via the strong force. Due to the relative strength of the strong force, in comparison to the electroweak force, most interactions at the LHC are induced by strong interactions. For the production of top quarks follows, that they are mainly produced in pairs ($t\bar{t}$) because the strong force is not flavour changing. Two production mechanisms are relevant: gluon-gluon fusion and quark-antiquark annihilation. Gluon-gluon fusion can happen as an s-channel, t-channel or u-channel process, depicted in Figure 2.2. In contrast, the $t\bar{t}$ pair production via quark-antiquark annihilation is only possible as an s-channel process shown in Figure 2.3. The gluon-gluon fusion production channel contributes around 90% of the production of $t\bar{t}$ pairs, while $q\bar{q}$ -annihilation contributes only 10%. The reason for this is given by the parton distribution functions of the proton (see Section 2.4). The higher the energy, the more gluons can be found inside the proton. A comparison of theoretical predictions and measurements from the LHC and the TEVATRON is shown in Figure 2.4.

2. The standard model of particle physics

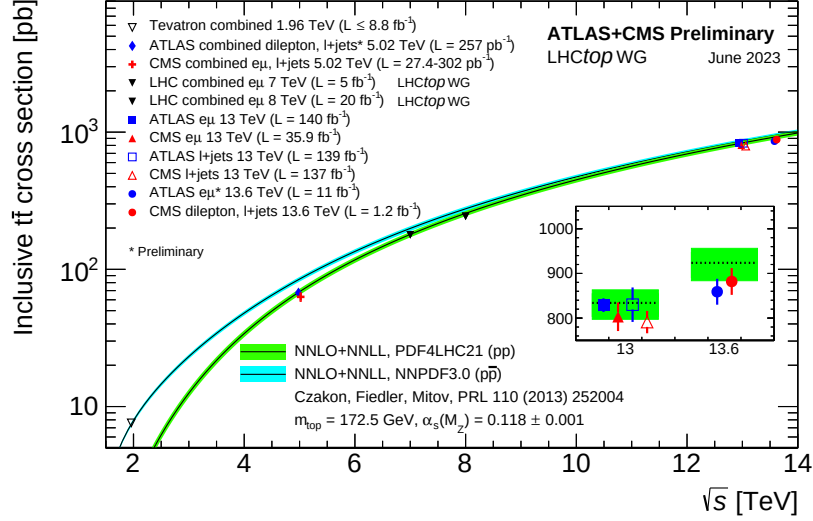


Figure 2.4.: Comparison of several $t\bar{t}$ cross-section measurements at the LHC and the TEVATRON at different centre-of-mass energies with the theoretical predictions at next-to-next-to-leading order (NNLO) in QCD with next-to-next-to-leading logarithmic (NNLL) resummation. The uncertainty band for the theory predictions takes scale uncertainties (factorisation and renormalisation), parton distribution function uncertainties and uncertainties on the strong coupling into account. Measurements at the same centre-of-mass energy are shown slightly shifted for better visibility [134].

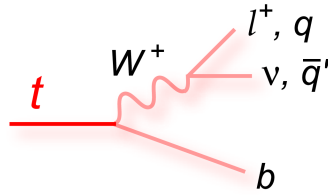


Figure 2.5.: The leading order Feynman diagram of a decaying top quark. The subsequent W^+ -decay is also shown with its two possibilities. The decay of an anti-top quark would look similar, as described above.

2.5. The top quark within the standard model

Due to the CKM matrix element $|V_{tb}|$ being close to 1, almost all top quarks decay into a W boson and a b -quark²². The W boson (independently of its charge) subsequently decays either hadronically and two jets (see Section 4.3.4) are created, or leptonically into a charged lepton and its corresponding neutrino. In Figure 2.5, the Feynman diagram of a top quark decay is shown. With the two top quarks this opens the possibility for three different decay modes²³:

- Fully hadronic channel: both W bosons decay into quarks and produce four jets. At LO, six jets (four light-jets²⁴ and two b -jets (see Section 4.3.4)) are expected to be detected.
- Lepton+jets channel: one W boson decays into two quarks and produces two jets, and the other W boson decays into a charged lepton and the neutrino (detectable via missing transverse energy E_T^{miss} ; see Section 4.3.5). At LO, two b -jets, one charged lepton and E_T^{miss} are expected to be detected.
- Dilepton channel: both W bosons decay into a charged lepton and its corresponding neutrino. At LO, two b -jets, two charged leptons, and E_T^{miss} are expected to be detected.

The different branching ratios for the W boson decay lead to different probabilities of the above-described decay channels commonly distinguished. The fully hadronic decay channel occurs in 45.7% of all $t\bar{t}$ events, the lepton+jets channel occurs in 43.8% of the cases, and the dilepton channel in 10.5% [5]. These numbers get slightly modified as most analyses only count muons and electrons when looking for leptons. As a consequence, only leptonic τ -decays are counted as leptonic W -decays, while the additional neutrino spoils a possible top-reconstruction. Hadronic τ -decays lead to migration of events between the different decay channels. The dilepton channel in particular is very clean and has only a small background, while the background becomes larger in the lepton+jets channel and is largest in the fully hadronic channel. In contrast, the number of expected events behaves in exactly the opposite way.

Single top quark production and decay

Via the weak interaction, it is also possible to get single top quarks (t or $\bar{t}$) produced at hadron colliders. The weak production of top quarks at the LHC is suppressed in comparison to the strong production of $t\bar{t}$ pairs. There are three production mechanisms, the s-channel, the t-channel, and Wt -production. The corresponding Feynman diagrams at LO are shown in Figure 2.6. As all three processes have different final state particles, it is possible to measure them individually. A comparison of different ATLAS measurements with theory predictions is shown in Figure 2.7. The decay of the top quarks is

²²To be precise, the top quark decays into a b -quark and a W^+ boson, while the anti-top quark decays into an anti- b -quark and a W^- boson.

²³Not taking the lepton charge into account.

²⁴Light-jets are jets not coming from a b -quark.

2. The standard model of particle physics

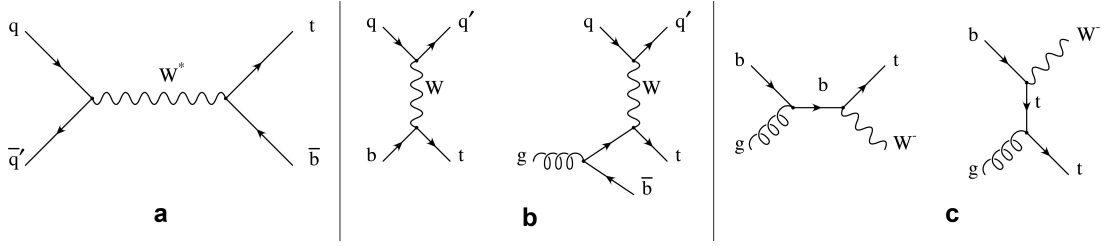


Figure 2.6.: Leading order Feynman diagrams for single top quark production via the s-channel (a), the t-channel (b) and the W -associated production (c).

the same as described above. Due to the different production mechanisms influencing the final state particles, the expectations which particles are detected differ as follows²⁵:

- t-channel:
 - Hadronic W -decay: one or two b -jets and three light-jets.
 - Leptonic W -decay: one or two b -jets, one charged lepton and E_T^{miss} .
 - One forward light jet.
- s-channel:
 - Hadronic W -decay: two b -jets and two light-jets.
 - Leptonic W -decay: two b -jets, one charged lepton and E_T^{miss} .
- Wt production: same final state particles expected as in $t\bar{t}$ pair production but only one b -jet expected instead of two.

In general, much larger backgrounds are expected than in $t\bar{t}$ pair production because the signal signatures are less clear (fewer jets and fewer b -jets), e.g. in comparison to W -boson+jets production.

2.5.2. Top quark properties

The top quark offers a variety of properties, which need to be measured, to see if the SM predicts them correctly. There are only a few parameters that are not predicted by the SM. Values for these parameters are needed to make all other predictions. One parameter is the top quark mass, or equivalently, its Yukawa coupling to the Higgs boson. The other parameters are the CKM matrix elements related to the top quark. Everything else is then predicted by the SM.

²⁵The hadronic decay happens in 67.6% of the cases, and the leptonic decay happens with a probability of 32.4% [5].

2.5. The top quark within the standard model

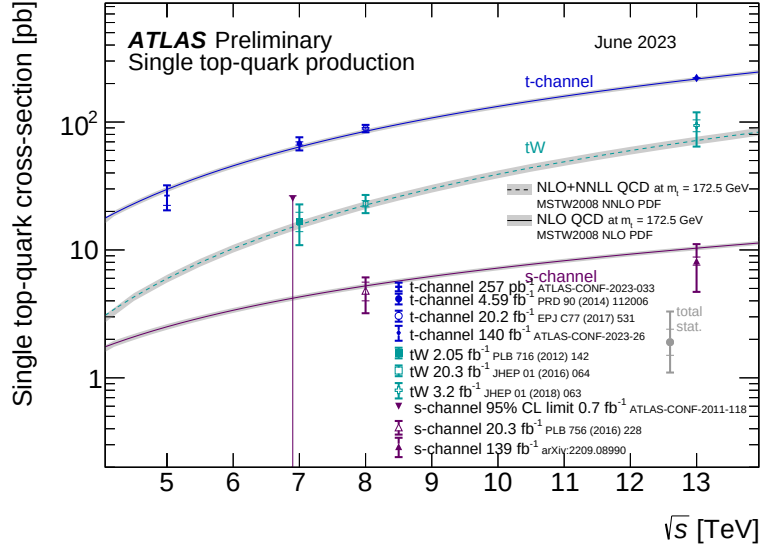


Figure 2.7.: Comparison of several single top quark cross-section measurements done by ATLAS at the LHC at different centre-of-mass energies with the theoretical predictions at NNLO in QCD with NNLL resummation. The uncertainty band for the theory predictions takes scale uncertainties (factorisation and renormalisation), parton distribution function uncertainties and uncertainties on the strong coupling into account. Measurements at the same centre-of-mass energy are shown slightly next to each other for better visibility [135].

Mass

There have been a lot of mass measurements performed at the TEVATRON and the LHC in all different decay channels with comparable results [5] around the world average of $174.34 \pm 0.76(\pm 0.36(\text{stat.}) \pm 0.67(\text{syst.}))$ GeV [130]. An overview of some of these measurements is given in Figure 2.8. Besides the direct mass measurements, it is also possible to use cross-section measurements to retrieve the mass because the cross-section is sensitive to it. A summary plot of these measurements is shown in Figure 2.9. Direct and indirect measurements are not directly comparable, due to theoretical considerations related to renormalisation and non-perturbative effects [5].

2. The standard model of particle physics

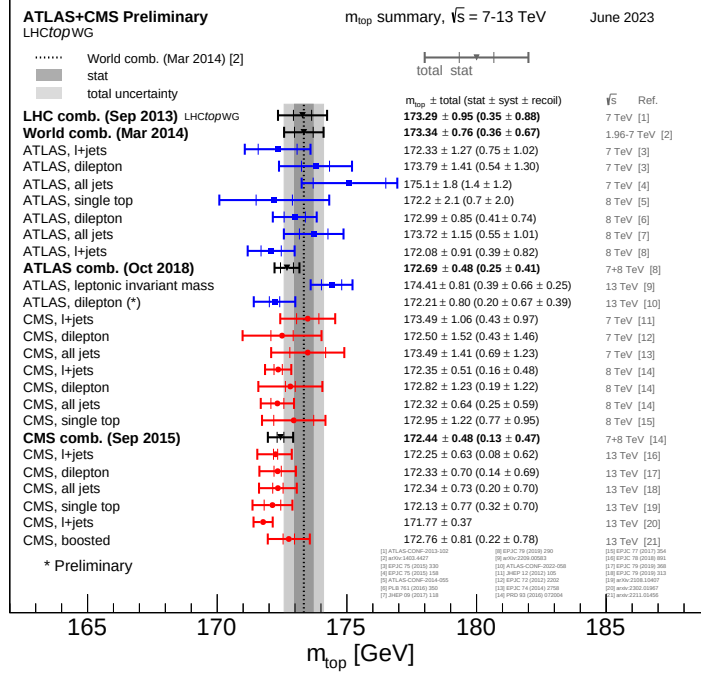


Figure 2.8.: Top quark mass measurements from ATLAS and CMS, as well as the relevant combinations [134].

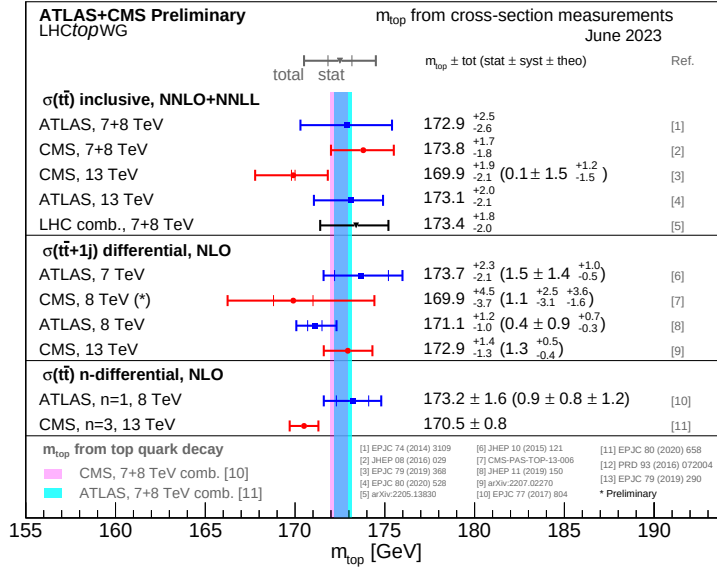


Figure 2.9.: Top quark mass measurements using cross-section measurements from ATLAS and CMS, as well as a combination [134].

2.5. The top quark within the standard model

CKM matrix elements

As mentioned above, the CKM matrix elements are not predicted by the SM and are therefore needed to make predictions. Because V_{tb} is quite large and V_{td} and V_{ts} are both very small, it becomes difficult to measure them precisely. It is possible to measure these parameters in both $t\bar{t}$ pair production and single top quark production. However, more precise results are obtained using single top quark measurements. If measurements showed that the CKM matrix is not unitary, this would hint at the existence of further quark generations, but so far, all measurements are consistent with unitarity. Again, a summary plot is shown in Figure 2.10.

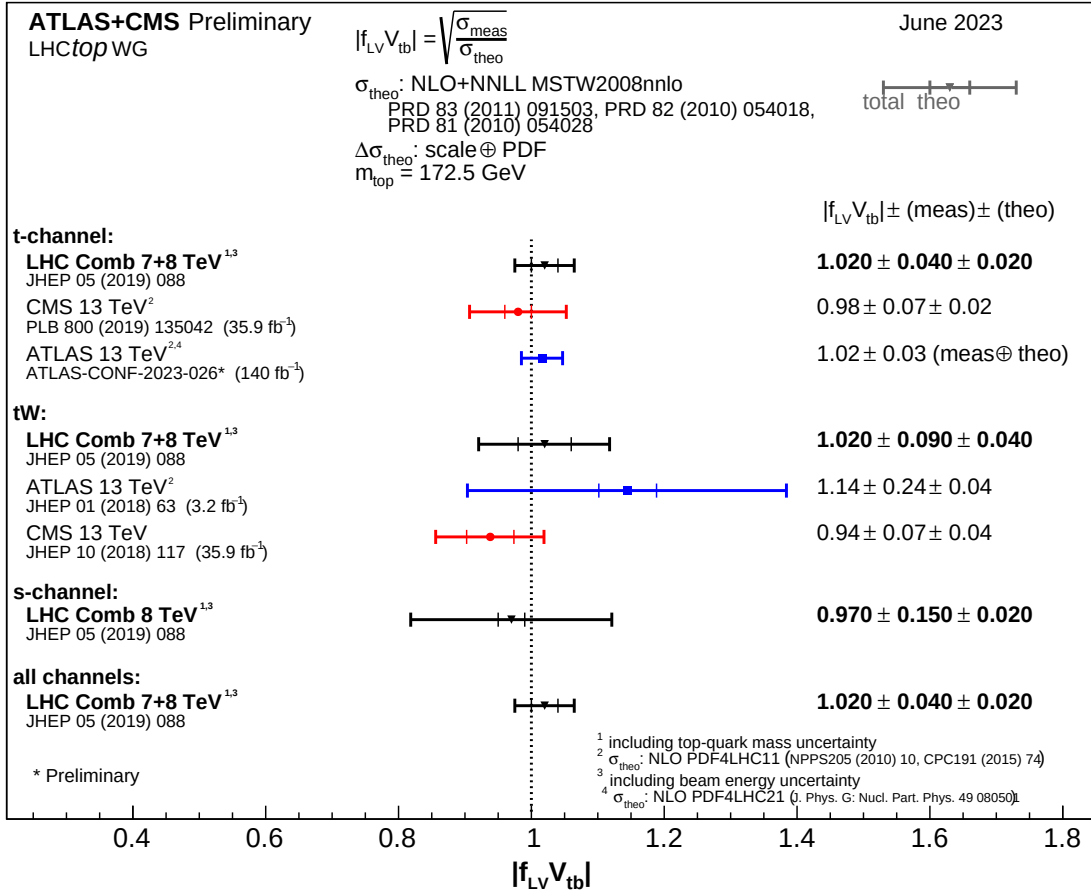


Figure 2.10.: Different results of V_{tb} measurements from ATLAS, CMS, and their combinations [134].

Top quark decay width

The top quark decay width depends on the different decay modes of the top quark. Through loop-corrections, it also depends on possible unseen particles. Hence, deviations

2. The standard model of particle physics

from the predictions can hint at new physics. The NNLO QCD prediction is 1.35 GeV for a top quark mass of 173.3 GeV [131, 132]. Direct measurements from ATLAS at $\sqrt{s} = 8$ TeV and $\sqrt{s} = 13$ TeV measure $\Gamma = 1.76 \pm 0.33(\text{stat.})_{-0.68}^{+0.79}(\text{syst.})$ GeV and 1.9 ± 0.5 GeV, respectively [136, 137]. An indirect measurement was made by CMS and yields $\Gamma = 1.36 \pm 0.02(\text{stat.})_{-0.11}^{+0.14}(\text{syst.})$ GeV [138]. No tension concerning the SM prediction was found.

Asymmetries

At the TEVATRON, a $p\bar{p}$ collider, the asymmetric initial state leads to a forward-backward asymmetry. In $q\bar{q}$ annihilation (the dominant $t\bar{t}$ production mechanism), the quarks mostly come from the proton beam and the anti-quarks mostly from the anti-proton beam. Due to charge asymmetric initial states in $t\bar{t}$ production at NLO, as $qg \rightarrow t\bar{t}$, and interference effects, the event kinematics become asymmetric. This can be visualised by the rapidity difference distribution $\Delta y = y_t - y_{\bar{t}}$ not being symmetric.

At LHC, the situation is different. Not only is $q\bar{q}$ annihilation less important than at the TEVATRON, but the initial state is symmetric, as protons are used for collisions. Hence, Δy is expected to be a symmetric distribution. However, there is another variable that becomes asymmetric, due to the aforementioned effects at NLO. The larger momentum fraction of the valence quarks in comparison to the sea quarks (see Section 2.4) and the fact that none of the valence quarks are anti-quarks, leads to a difference in the spread of the rapidity distribution for t and $\bar{t}$. Hence $\Delta|y| = |y_t| - |y_{\bar{t}}|$ becomes asymmetric. The asymmetry that follows is called charge asymmetry A_C and can be measured using

$$A_C = \frac{N(|y| > 0) - N(|y| < 0)}{N(|y| > 0) + N(|y| < 0)}. \quad (2.47)$$

This asymmetry is also transferred to the decay products. In the dilepton channel, instead of using the reconstructed top quarks, one can use the leptons instead (although the effect is a bit smaller) [139]. In general, the effect is predicted to be around 1% or even less [139]. In $t\bar{t}$ events with additional gauge bosons, these effects are expected to be enhanced, as $q\bar{q}$ production is less suppressed [140, 141]. Another asymmetry that can be measured is an energy-asymmetry [142]. All measurements are in agreement with the predictions [5, 143–146], but as a consequence of their large errors and the smallness of the effect at the LHC, all these measurements are also compatible with no asymmetry. This situation might not improve in further runs of the LHC, as a higher centre-of-mass energy results in a smaller asymmetry as gluon-gluon fusion will dominate $t\bar{t}$ production even more.

Polarisation, spin correlations and quantum entanglement

The top quark is a spin-1/2 particle and the spin information is transferred to its decay products, as it decays before it can hadronise. In principle, the way the spins of the two top quarks are correlated (being aligned parallel) is dependent on the production

2.6. Top quark pair production in association with a photon

mechanism. One can measure the following quantity [147]:

$$A = \frac{N(\uparrow\uparrow) + N(\downarrow\downarrow) - N(\downarrow\uparrow) - N(\uparrow\downarrow)}{N(\uparrow\uparrow) + N(\downarrow\downarrow) + N(\downarrow\uparrow) + N(\uparrow\downarrow)}. \quad (2.48)$$

Recent measurements have shown agreement with SM predictions [148, 149]. Further, one can measure the polarisation of the top quarks using the differential distribution of $\cos(\theta)$ of one of the fermions, where θ is the angle between the fermion and the spin basis in the rest frame of the top quark. Measurements show, in general, good agreement with the SM prediction [149].

Another property of top quark pairs is that they can form entangled states. The information about such a state is transferred to the decay products and encoded in the differential distribution of the angle between the leptons (in the dilepton channel) in the respective rest-frames of their corresponding top or anti-top quark [150]. The SM predicts entangled states in regions of phase space where the invariant mass of the $t\bar{t}$ system is approximately below 550 GeV. For detection of quantum entanglement in $t\bar{t}$ systems, a phase space cut on the invariant mass of 450 GeV is proposed [150]. Recently ATLAS published a first observation [151].

W-helicity

The helicity of W bosons from top quark decays is a property which probes the spin transfer from the top quark to its decay products, and in addition, the weak interaction vertex involved in the top-decay is probed as well. The helicity of the W boson is the projection of its spin onto its momentum. Because of its mass, the W boson can have three spin states $(0, \pm 1)$, resulting in three helicity states. If the spin and the momentum show in the same direction, it is called right-handed, while spin and momentum pointing in opposite directions is called left-handed. If the W boson is in its spin-0 state, the corresponding helicity state is called longitudinal. As the helicity is not a Lorentz-invariant quantity, it needs to be defined in a certain rest frame. The rest frame normally used to measure the W -helicity is the rest frame of its parent top quark. Therefore, the predictions are not $\sqrt{s}$ -dependent. Due to the event kinematics and the conservation of angular momentum not all three helicity states occur with the same probability. The predictions are $f_0 = 0.687 \pm 0.005$, $f_R = 0.0017 \pm 0.0001$ and $f_L = 0.311 \pm 0.005$ [152]. While ATLAS has measured the helicity fractions at all available centre-of-mass energies [153–155], CMS has measured them only at 7 and 8 TeV [156–158], and one combination exists [159]. The results are in agreement with the SM predictions.

2.6. Top quark pair production in association with a photon

2.6.1. General introduction to top quark cross-section measurements with associated boson production

The couplings between the fermions and the (gauge) bosons are essential predictions the SM makes about how quantum fields interact with each other. Measuring them,

2. The standard model of particle physics

and therefore probing the predictions of the SM, is one of the foremost tasks one can find in experimental particle physics. To specifically probe the couplings of the bosons to the top quark, $t\bar{t} + X$ processes are the relevant processes. In the production of single top quarks and the decay, the coupling to the W boson is already probed. $t\bar{t}W$ measurements only give additional insights about the production mechanisms and their relative fractions. The other two electroweak gauge bosons, the Z boson and the photon (on- and off-shell) do not play a relevant role in top quark production and/or decay. The same holds for the Higgs boson²⁶. So these couplings can be probed directly in events with top quarks and associated bosons, e.g. pair production of top quarks in association with a photon ($t\bar{t}\gamma$), a Z boson ($t\bar{t}Z$) or a Higgs boson ($t\bar{t}H$), as well as single top production in association with a Z boson and an additional quark (tZq), a photon and an additional quark ($t\gamma q$), a Higgs boson and an additional quark (tHq) or a Z and a W boson (tWZ). Recent summary plots of such measurements are shown in Figures 2.11 and 2.12.

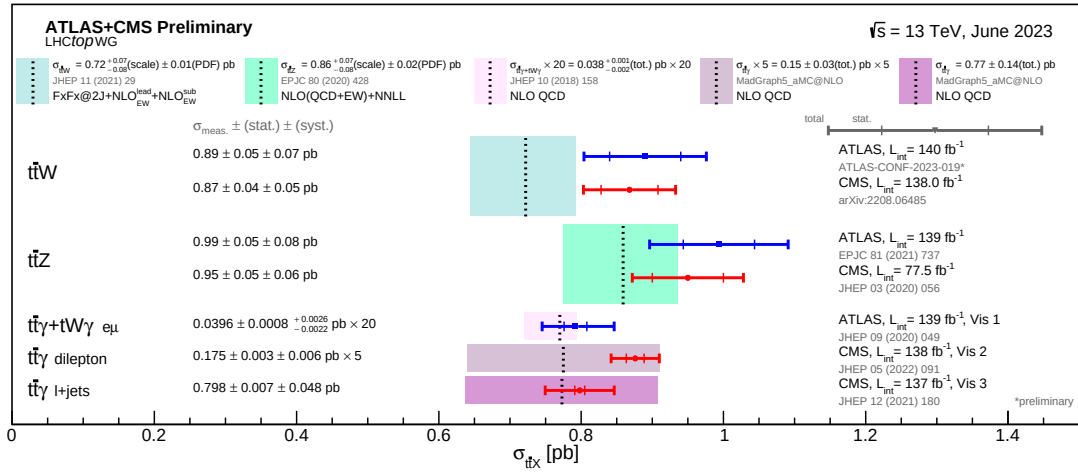


Figure 2.11.: Recent $t\bar{t}+X$ measurements from ATLAS and CMS [160].

2.6.2. The top-photon coupling

In the SM, the interaction Lagrangian between the top quark and the photon follows from Equation 2.6 as

$$\mathcal{L}_{t\gamma} = -\alpha_{em} Q_t \bar{\Psi}_t \gamma^\mu \Psi_t A_\mu, \quad (2.49)$$

where α_{em} is the fine structure constant and Q_t the charge, which is $2/3e$ in the case of the top quark. The QED interaction vertex factor Γ of the top-photon coupling becomes

$$\Gamma = i\alpha_{em} Q_t \gamma^\mu, \quad (2.50)$$

²⁶Although the top-Higgs coupling is also probed in Higgs boson production via gluon-gluon fusion and in 4-tops measurements.

2.6. Top quark pair production in association with a photon

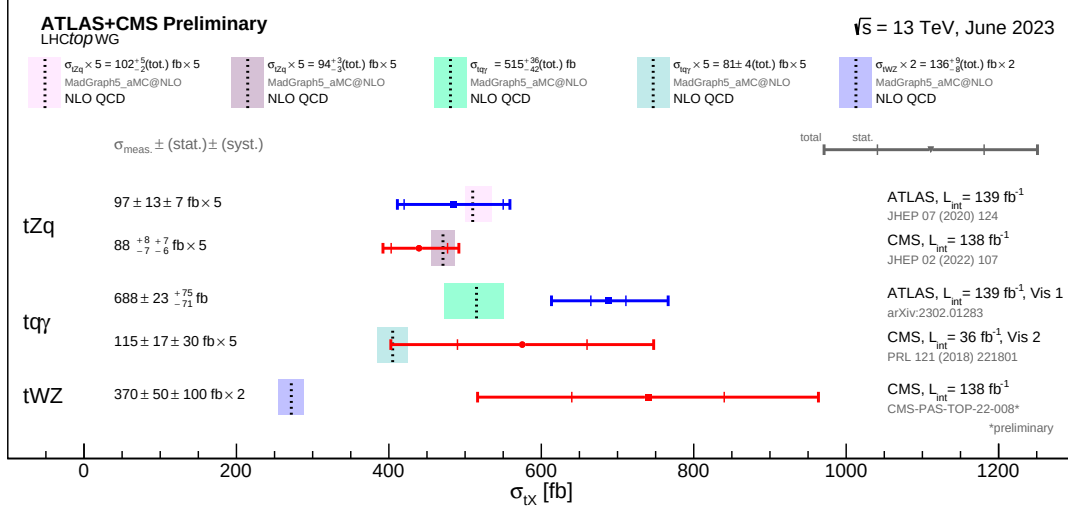


Figure 2.12.: Recent single top quark+X measurements from ATLAS and CMS [160].

which is a pure vector coupling, and therefore invariant under parity transformation P and charge conjugation C .

Several extensions to the SM can modify this coupling. The most general form of a fermion-gauge-boson coupling is given by [161–164]:

$$\Gamma = -ie \left(\gamma_\mu [F_1 + F_2 \gamma_5] + \frac{\sigma_{\mu\nu} (k + k')^\nu}{2m_f} [F_3 + F_4 \gamma_5] \right), \quad (2.51)$$

where F_i are so-called form factors, k and k' are the 4 momenta of the in- and out-going fermions (in this case top quarks), m_f the mass of the respective fermion and $\sigma_{\mu\nu} = \frac{1}{2}[\gamma_\mu, \gamma_\nu]$. The form factors do describe known physical quantities. $F_1 = Q_t = 2/3e$ is the predicted charge of the top quark. F_2 is predicted to be 0 in the SM as the photon has no axial-vector contribution. F_3 and F_4 are the magnetic and electric anomalous dipole moments. These do not appear at LO in the SM but are introduced through loop corrections [165] and therefore are very small. At LO, they can be introduced as higher-dimensional effective operators [166, 167], and if any contribution from them is measurable, this would indicate new physics, because these terms are too small to be measurable in the SM [168]. In general, there are a lot of extensions to the SM, e.g. supersymmetry, predicting such contributions to be measurable [163, 167–172].

The most general way to have model-independent interpretations of measurements is the framework of effective field theories. In effective field theories, the SM Lagrangian is extended by terms of higher dimension [173–175]

$$\mathcal{L}_{eff} = \mathcal{L}_{SM} + \frac{1}{\Lambda} \sum C_i O_i^{D=5} + \frac{1}{\Lambda^2} \sum C_i O_i^{D=6}. \quad (2.52)$$

The strength of these additional operators O_i is given by the so-called Wilson coefficients C_i [176], and Λ is the cut-off scale. $D = 5$ and $D = 6$ give the dimension of the operators.

2. The standard model of particle physics

These higher-order operators break renormalisation, but for an effective theory at a much lower scale, where the additional particles are tremendously off-shell, this does not matter. For example, Fermi theory was one of the first effective theories successfully used [113]. For anomalous contributions to the top-photon coupling giving sizable electric and magnetic dipole moments, only two dimension six operators are relevant [167]:

$$\delta F_3 = \frac{\sqrt{2}}{e} \text{Re} [c_w C_{uB\phi}^{33} + s_w C_{uW}^{33}] \frac{vm_t}{\Lambda^2}, \quad (2.53)$$

$$\delta F_4 = \frac{\sqrt{2}}{e} \text{Im} [c_w C_{uB\phi}^{33} + s_w C_{uW}^{33}] \frac{vm_t}{\Lambda^2}. \quad (2.54)$$

c_w and s_w are the cosine and sine of the Weinberg angle, v the vacuum expectation value of the Higgs field, and m_t the top quark mass. An additional operator might also modify the $t\bar{t}\gamma$ cross-section, which is related to the strong production, namely the $gt\bar{t}$ -vertex. The two coupling parameters related to it are

$$\delta F_5 = \frac{\sqrt{2}}{g_s} \text{Re} C_{uG\phi}^{33} \frac{vm_t}{\Lambda^2}, \quad (2.55)$$

$$\delta F_6 = \frac{\sqrt{2}}{g_s} \text{Im} C_{uG\phi}^{33} \frac{vm_t}{\Lambda^2}. \quad (2.56)$$

2.6.3. Probing the top-photon coupling

The best way to probe the top-photon coupling would be an e^+e^- -collider, as the production would happen at a well-defined centre-of-mass energy, which is not the case for hadron colliders (see Section 2.4). In addition, the production of top quark pairs would happen via the neutral gauge bosons and, therefore, would be directly sensitive to the top-photon coupling, taking interference with the top- Z coupling into account. A step in the so-called R-ratio of $e^+e^- \rightarrow \mu^+\mu^-$ production vs. $e^+e^- \rightarrow q\bar{q}$ production, would be visible at the $t\bar{t}$ production threshold of around 350 GeV and could be related directly to the charge. Another advantage would be the clean environment, without any hadronic pile-up. So, any background would be almost negligible. Unfortunately, synchrotron radiation made it impossible to build such a circular collider in the past, and the International Linear Collider (ILC) is not even confirmed to be built. The consequence is that the top-photon coupling must be probed in a hadronic environment [177]. In contrast to the $t\bar{t}$ production (see Section 2.5.1 and Figures 2.2, 2.3 and 2.5) only one additional photon is emitted at some point during production or decay. Example Feynman diagrams are shown in Figure 2.13 for photon emission during $t\bar{t}$ production and in Figure 2.14 during top quark decay. Only those processes are sensitive to the top-photon coupling, where the diagram shows a $t\gamma$ interaction vertex. It is mainly the production part (see Figure 2.13). The only other source of photons are initial state quarks. As long as one does not try to differentiate between gluon-gluon fusion and $q\bar{q}$ annihilation, these photons cannot be treated as background, because of interference effects with photon radiation from the top quark in $q\bar{q}$ initiated $t\bar{t}\gamma$ events. In the decay process of the top quark, a lot of photon production occurs that is not sensitive to the top-photon coupling.

2.6. Top quark pair production in association with a photon

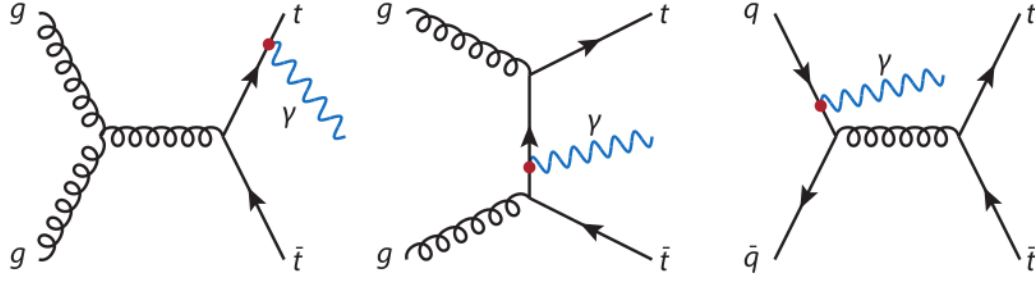


Figure 2.13.: Leading order Feynman diagrams showing how photons can be produced in the production of a $t\bar{t}$ pair. On the left, it comes from a real top quark, in the middle from a virtual top quark in t-channel production and on the right as initial state radiation from a quark in $q\bar{q}$ annihilation. The diagrams were taken from [179] under the CC BY-NC-ND 4.0 Deed license²⁷.

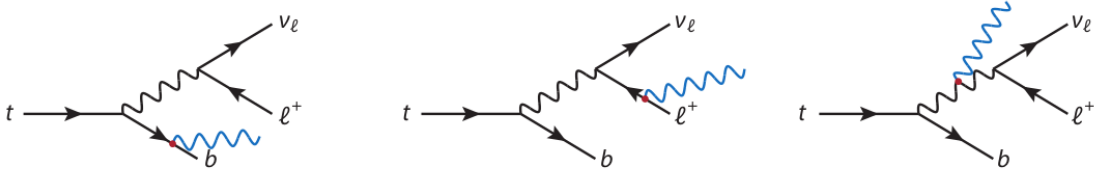


Figure 2.14.: Leading order Feynman diagrams showing how photons can be produced in the decay of a top quark. It can come either from the b -quark (left), from the charged decay products of the W boson (middle) or from the W boson itself. The diagrams were taken from [179] under the CC BY-NC-ND 4.0 Deed license²⁸.

In principle, there are interference effects between production and decay. These effects are negligible [178] and hence the so-called narrow-width approximation, which assumes $\Gamma_t \ll m_t$, can be applied to simplify the propagator of the top quark using

$$\frac{1}{(p_t^2 - m_t^2)^2 + m_t^2 \Gamma_t^2} \rightarrow \frac{\pi}{m_t \Gamma_t} \delta(p_t^2 - m_t^2) + \mathcal{O}\left(\frac{\Gamma_t}{m_t}\right). \quad (2.57)$$

It decouples the cross-sections for decay and production. Predictions of these decoupled cross-sections do only exist for the $e\mu$ decay channel [178], where a full NLO calculation in QCD exists [180]. Further predictions of the $t\bar{t}\gamma$ process do exist but assume stable top

²⁷The license terms can be found here: <https://creativecommons.org/licenses/by-nc-nd/4.0/>. No changes were made.

²⁸The license terms can be found here: <https://creativecommons.org/licenses/by-nc-nd/4.0/>. No changes were made.

2. The standard model of particle physics

quarks [181–184]. More accurate is a calculation using the narrow width approximation [185], but this calculation does not report the decoupled cross-sections for decay and production. Consequently, for this analysis, where it will be tried out for the first time to make use of separating the production contribution from the decay contribution for lepton+jets and the dilepton decay channel, only fiducial cross-section predictions from the generators for the MC samples are available, and they serve as theory predictions (see Section 4.2).

On the experimental side, evidence for the $t\bar{t}\gamma$ process was first found by CDF at the TEVATRON [186]. The first observation (at the LHC) was made by ATLAS at 7 TeV [26]. Both ATLAS and CMS have published several papers since then at centre-of-mass energies of 8 and 13 TeV [1, 2, 27–30]. All measurements did not differentiate between photons from production and photons from decay. No deviations from the SM predictions were found. The top quark charge was also measured indirectly by using $t\bar{t}$ events and the jet charge of the b -jets. An exotic top quark model, with a charge of $-4/3e$ instead of $+2/3e$ [187, 188], could be excluded, and the charge was measured to be SM-like with a very high significance [189–192]. Nevertheless, this measurement is not sensitive to the top-photon coupling where not only the charge plays a role but also possible anomalous electric/magnetic dipole moments, as described above. The $t\bar{t}\gamma$ measurements from ATLAS [1, 2, 26, 27] so far only measured cross-sections, but never extracted the charge or the coupling itself, albeit the negative results of no deviations concerning the SM predictions show that the charge is as predicted. CMS has made EFT interpretations at 13 TeV [29, 30] and set limits on the anomalous electric/magnetic dipole moments. This will be done for ATLAS within this analysis [3] for the first time as well and will constrain the top quark charge and possible modifications of the $t\gamma$ interaction vertex as precisely as possible²⁹.

²⁹The EFT interpretations were not made by the author.

3.1. CERN and the LHC

3.1.1. CERN

CERN (Conseil Européen pour la Recherche Nucléaire) is a large-scale research facility, which was founded in 1954 [193]. One purpose was to have a research facility on nuclear physics being independent of any particular government to guarantee the freedom of science in a field that was highly political at that time [193]. Therefore the CERN area is extraterritorial ground and has an observer role at the United Nations [194, 195]. From the beginning on it was also a mission for peace by bringing people together to work with each other motivated by their intrinsic curiosity to reveal the yet unknown. Another reason to build such a facility was the insight that fundamental research on nuclear physics and by now mainly on particle physics, is mostly not doable for one country alone. It needs a common effort to provide the necessary infrastructure like powerful accelerators and experimental halls. Ever since CERN had a leading role in exploring the frontiers of physical knowledge and making it available to society. Today it hosts the most powerful particle accelerator, that was built until now, the LHC [193, 196].

3.1.2. LHC

The LHC is the most powerful particle accelerator that has ever been built so far. The first ideas of a hadron machine making use of the big tunnel built for the Large-Electron-Positron-Collider (LEP) date back to 1984, even before LEP went into operation [197]. At that time, there were similar ideas for such machines to be built in the Soviet Union (UNK) and the United States (Superconducting Super Collider; SSC). Both projects

3. Experimental setup

were stopped during their building phase. In 1994 the LHC project was approved [198]¹.

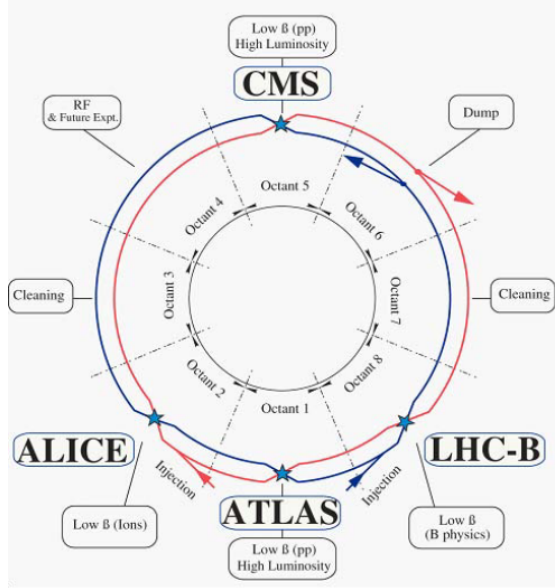


Figure 3.1.: Schematic overview of the LHC, the interaction points and the location of the technical infrastructure. The figure was taken from [200].

Design parameters²

The LHC was designed as a proton-proton collider (which can also be used as a heavy ion collider) with a circumference of approximately 27 km. Its original design centre-of-mass energy is 14 TeV with a peak luminosity of $10^{-34} \text{ cm}^{-2}\text{s}^{-1}$. The main factor limiting the beam energy are the superconducting magnets with a design field strength of 8.33 T, but there are also other factors limiting the beam energy as the particle density, the mechanical aperture and the energy stored in the beam and the magnets. The proton beams were planned to consist of bunches with around 10^{11} protons per bunch and a maximum number of bunches of 2808, which means that the collision rate is 40 MHz (collisions every 25 ns). At 4 interaction points the main experiments³ (ATLAS [203], CMS [204], LHCb [205] and ALICE [206]) are located, while beam injection happens at two of them. The machine is equipped with a beam dump, radio frequency (RF) cavities for acceleration and collimator systems for beam cleaning (see Figure 3.1). The RF cavities are working at 400 MHz. The field strength is 5.5 MV/m and with a 7 TeV beam energy the voltage will be 16 MV, while at the injection the voltage is 8 MV. The magnet system mainly consists of superconducting quadrupole magnets for the beam focus and superconducting dipole magnets for the beam bending, both being operated

¹More on the history of the LHC can be found in Chapter 1 in [199] and the references therein.

²If not stated otherwise, this section is based on information from [200].

³There are a few smaller experiments like TOTEM [201] or LHCf [202].

3.1. CERN and the LHC

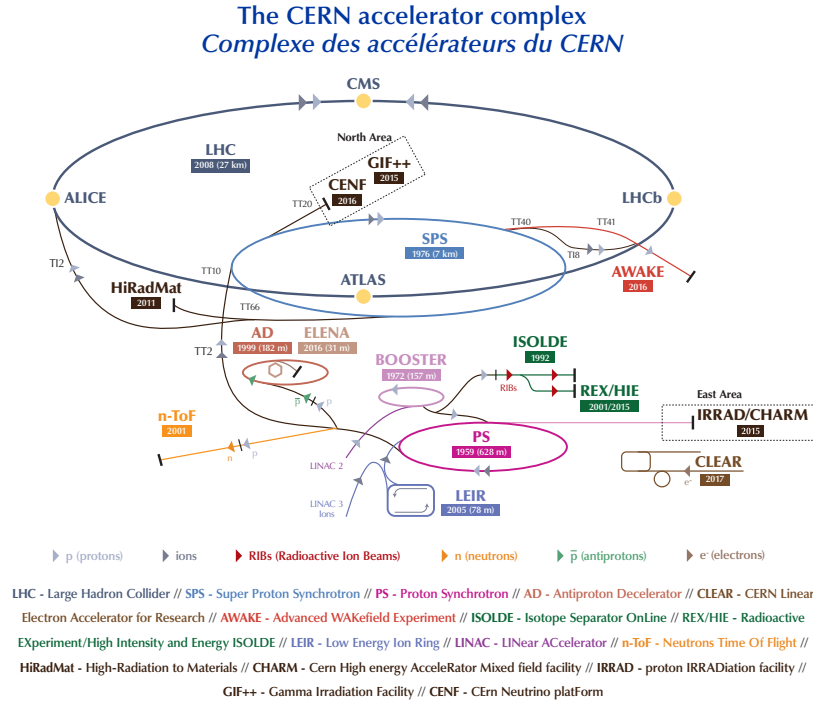


Figure 3.2.: Schematic overview of CERN's accelerator complex (copyright by CERN⁴).

at 1.9 K. The cooling is done with liquid Helium. Sextupole, octupole, decapole and 12-pole magnets are installed for corrections on the beam trajectories. One main feature is the twin aperture that allows to have one magnet and cryo-system for both beams. While running the LHC, the luminosity will decrease exponentially over time. The time to start a new run and the luminosity lifetime define the optimal run-time before a beam should be dumped and a new run will be started. To get the LHC filled with protons a whole chain of accelerators is needed. A schematic overview of the full accelerator complex at CERN is shown in Figure 3.2. At the beginning comes the LINAC 2, a linear accelerator where at the start Hydrogen atoms are used and the electrons are stripped off. LINAC 2 accelerates the protons up to 50 MeV. Afterwards, the Booster increases the energy to 1.4 GeV. The next step is the PS (Proton-Synchrotron). The protons will leave from here with an energy of 25 GeV and will be transferred to the SPS (Super-Proton-Synchrotron). Here they will get their injection energy for the LHC of 450 GeV. While accelerating step by step the preaccelerators also structure the proton beam as desired for the LHC (bunches of around 10^{11} protons at a distance of $c \cdot 25 \text{ ns} \approx 7.5 \text{ m}$ with the required beam shape). The last step of the acceleration occurs in the LHC where the protons reach their final energy.

⁴This figure was taken from <https://cds.cern.ch/record/2636343> (accessed on 2.10.2023).

3. Experimental setup

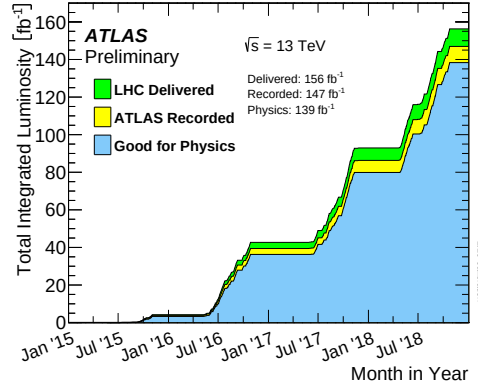


Figure 3.3.: The plot shows the integrated luminosity evolution in Run 2, showing the delivered luminosity, the recorded luminosity and the luminosity that is good for physics analysis. The numbers on this plot are determined from online luminosity estimates. A final Run 2 luminosity calibration gave a luminosity of 140.1 fb^{-1} instead of 139 fb^{-1} [210]. This figure was taken from [211].

LHC performance during Run 2⁵

Due to an incident in 2008, the LHC had a center of mass energy in Run 1 of 7 TeV in 2011 and 8 TeV in 2012 [208, 209]. For Run 2 the energy per beam was increased to 6.5 GeV, resulting in a centre-of-mass energy of 13 TeV. Over the four years of Run 2 (2015 to 2016), several other beam conditions changed and were improved to provide as much luminosity as possible to the experiments. Besides the change in the centre-of-mass energy, the focus in 2015 was to go from a bunch spacing of 50 ns to 25 ns, while having a secure operation. Mainly the higher beam energy made it necessary to be very careful to not cause any damage to the LHC. In 2016 the main advance was to have reached a β^* that was beyond the design level, resulting in a tighter beam focus than originally planned. Hence the peak luminosity was increased beyond the design value too. Step by step the luminosity was increased even more in 2017 and 2018. In the end, the luminosity could be doubled concerning the design goal. The full integrated luminosity delivered by the LHC is shown in Figure 3.3. 140.1 fb^{-1} of luminosity could be collected for physics analysis [210]. All these improvements also lead to an increase in pile-up making the environment much harsher for the detectors. The pile-up distribution for Run 2 is shown in Figure 3.4.

3.2. The ATLAS detector

The ATLAS detector is a multipurpose detector located at Point 1 at the LHC [203]. The experiment is one of the two multipurpose experiments currently operated at the LHC. The other one is CMS [204].

⁵If not stated otherwise all information in this section is from [207].

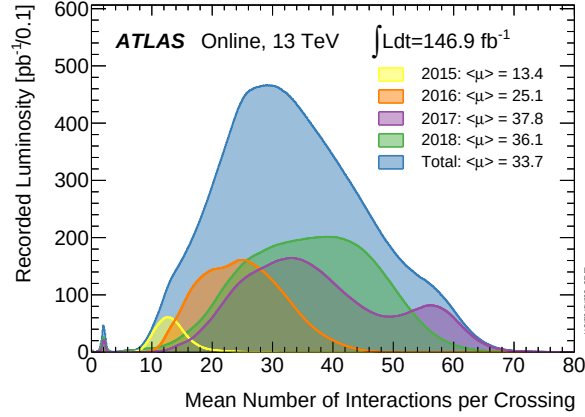


Figure 3.4.: The plot shows the pile-up profile in Run 2 for the different years [211].

The ATLAS detector was built to detect and reconstruct high-energy proton-proton collisions. Its general design follows the general shape of such detectors. By volume, it is the largest particle physics detector ever built, with a length of around 45 m and a height of about 25 m. The right-handed coordinate system of the detector is oriented such that the interaction point is the centre of it and the beam pipe is the z -axis. The $x - y$ -plane is oriented transversal to it. The x -axis points towards the centre of LHC, while the y -axis points upwards. Due to the cylindrical shape, instead of the x - and y -coordinates the azimuthal angle ϕ is used and the pseudorapidity $\eta = -\ln \tan(\theta/2)$. Closest to the beam pipe and the interaction point of the two proton beams is the Inner Detector (ID) located. It consists of a silicon pixel detector (PD), a silicon strip detector (semi-conductor tracker = SCT) and a transition radiation tracker (TRT). The ID is surrounded by a 2 T solenoid to bend the track and make p_T measurements possible in addition to the tracking. The TRT contributes to identifying particles. Outside the solenoid comes the electromagnetic calorimeter (E-CAL). It stops photons and electrons and records their showers to do energy measurements. The same happens in the hadronic calorimeter (H-CAL), which stops mesons and baryons and measures the particle shower produced in the stopping. The outermost part of the ATLAS detector is the muon spectrometer, equipped with toroidal magnets. All these detector systems are placed in a cylindrical shape around the beam pipe. To also cover the forward regions at the end of this cylinder, each detector system has also placed detector components there [203]. A schematic overview of the detector can be seen in Figure 3.5. The single detector systems and their subcomponents are described in more detail in the following sections.

3.2.1. The Inner Detector

The whole ID has a diameter of around 2.3 m and a total length of around 7 m (see Figures 3.6 and 3.7). Around the ID, a 2 T solenoid is placed to create a rather homogeneous magnetic field in the ID. The most important task for the ID is to contribute to vertex

3. Experimental setup

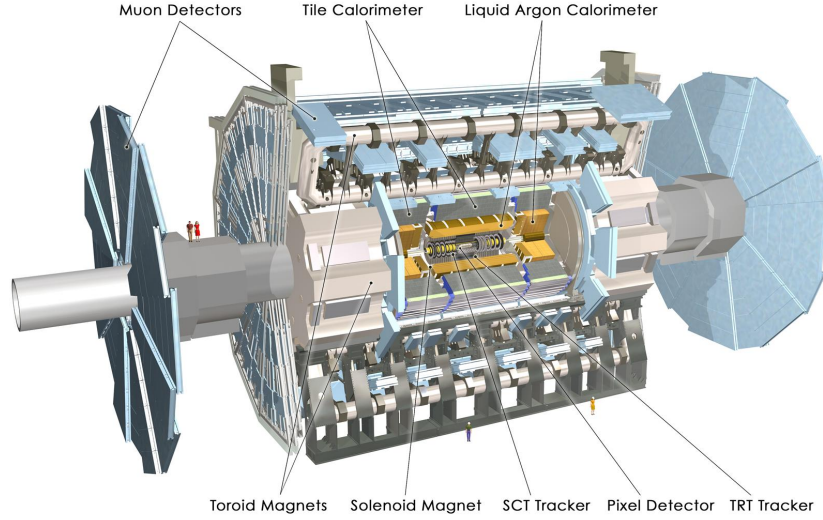


Figure 3.5.: The figure shows a schematic overview of the ATLAS detector [214]. One can also see schematic people that show the length scales of this experiment.

measurements (and therefore also contributes importantly to b -tagging) and p_T measurements (from the bent tracks due to the magnetic field). The insertion of IBL for Run 2 was a major improvement to the tracking information that is collected [203, 212, 213]. The ID also helps to do particle identification, making use of the transition radiation from the TRT [203]. In the following the three subsystems are described in more detail.

The pixel detector

The pixel detector is the innermost part of the ID [203]. A schematic view of the whole ID is shown in Figure 3.6 and 3.7. The pixel detector consists of 4 Layers concentrically placed around the beam pipe and 3 end-cap disks. The innermost layer is called IBL (insertable B-Layer) and was added in LS 1 (long shutdown 1) [212]. It is located 33.25 mm away from the beam pipe. The other three layers are called Layer 0, Layer 1 and Layer 2. Their distance to the beam pipe is 50.5 mm, 88.5 mm and 122.5 mm, respectively. In addition, there are 3 disks on each side to close the end-cap. Those disks are placed 495 mm, 580 mm and 650 mm away from the interaction point (in z -direction) respectively, while the barrel part extends on both sides up to 400.5 mm. With these specifications, the pixel detector can cover tracks up to $|\eta| < 2.5$ [203]. Most pixels have a size of $50 \times 400 \mu\text{m}^2$, while some pixels are a bit larger with $50 \times 600 \mu\text{m}^2$ (due to chip electronics). The intrinsic accuracy of the pixels is $10 \mu\text{m}^2$ in the $R - \phi$ plane, while it is $115 \mu\text{m}^2$ in z -direction [203]. The pixel size for IBL is smaller with $50 \times 250 \mu\text{m}^2$, which improved the already very good resolution in Run 1 [212]. In total, there are 1744 modules in the original layers and disks and IBL has 280 modules [203, 213].

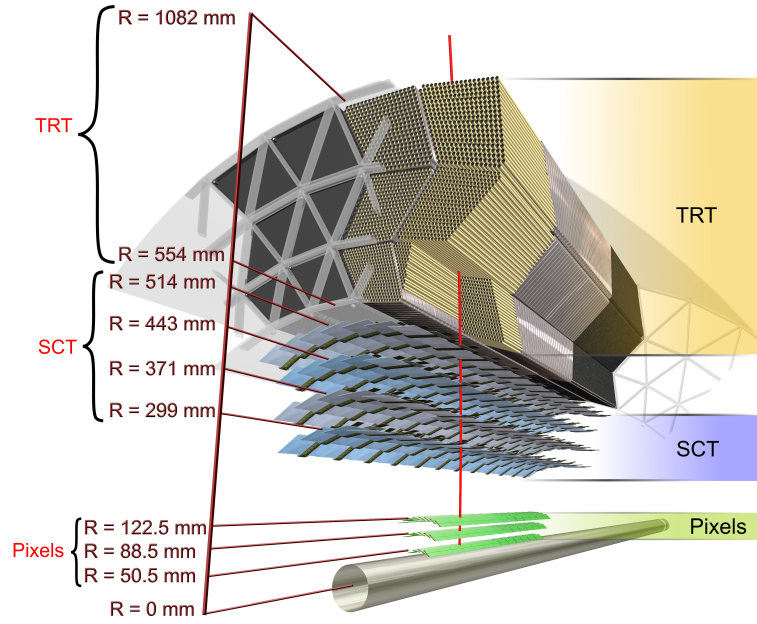


Figure 3.6.: The figure shows a schematic overview of the ATLAS Inner Detector [215]. The IBL layer is missing.

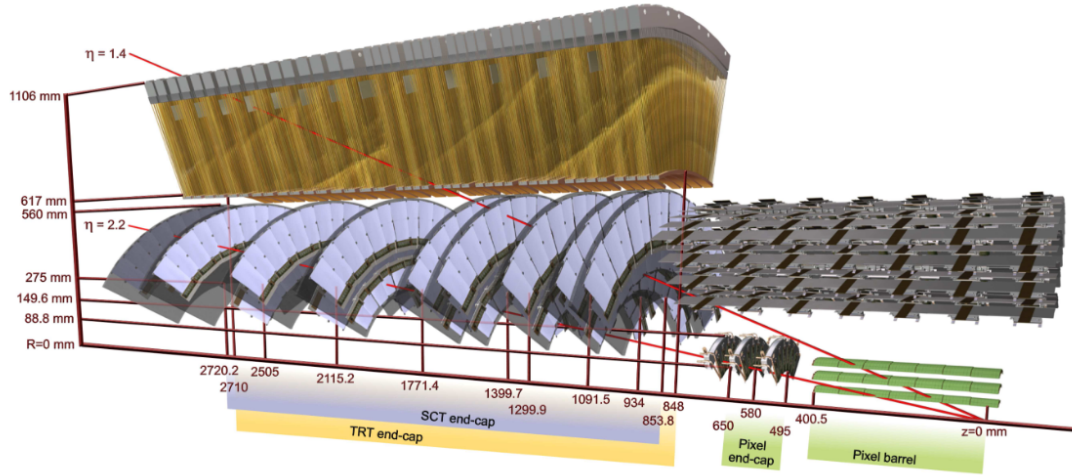


Figure 3.7.: The figure shows a schematic overview of the ATLAS Inner Detector end caps [203]. The IBL layer is missing.

3. Experimental setup

The semi-conductor tracker

Next to the pixel detector comes the SCT. As the pixel detector, it is a tracking detector, but as it is located further away from the beam pipe, it uses silicon strips instead of pixels. It has 4 barrel layers and up to 9 end-cap disks allowing the detector to cover the same η range of $|\eta| \leq 2.5$ as the pixel detector does. The layers are located 299 mm, 371 mm, 443 mm and 514 mm away from the beam pipe, respectively, as shown in Figure 3.6. The barrel part extends to 848 mm in positive and negative z -direction, respectively. The disks are placed in the range from 853.8 mm to 2720.2 mm in positive and negative z -direction (see Figure 3.7). The strips have a pitch of $80\mu\text{m}$. The intrinsic accuracy of the strips is $17\mu\text{m}$ in the $R-\phi$ plane and $580\mu\text{m}$ in the z -direction. In total, there are 4088 modules, 2112 in the barrel layers and 1976 in the end caps.

The transition radiation tracker

The TRT is the outermost part of the ID (see Figure 3.6). The barrel part extends from 554 mm to 1082 mm in the radial direction and to 780 mm in the z -direction. The end caps extend in the z -direction from 827 mm to 2744 mm, which only gives a coverage up to $|\eta| \leq 2.0$ (see Figure 3.7). The TRT is made from gas filled straws⁶ of 4 mm diameter, with an electric wire ($30\mu\text{m}$) in the middle. Charged particles will ionise the gas and the electric pulse that is created is registered as a hit (drift time measurement). The intrinsic accuracy is $130\mu\text{m}$ in the $R-\phi$ plane. Between the straws, polypropylene fibers are placed in the barrel part and polypropylene foils in the end-caps. When the particles traverse the boundary layer between the straws and the polypropylene they send out so-called transition radiation, which is used for particle identification.

3.2.2. The electromagnetic calorimeter

The first calorimeter that is placed outside the solenoid is the electromagnetic calorimeter. Its goal is to stop photons and electrons/positrons and record the electromagnetic shower produced by interactions of these particles with the material. It shares the same vacuum vessel with the solenoid to reduce material. The full calorimeter (including the H-CAL) is shown in Figure 3.8. The E-CAL is a sampling calorimeter using liquid Argon as active material and lead as passive material. The accordion structure assures full ϕ coverage. The η -coverage of the barrel part reaches $|\eta| \leq 1.475$ and the end-caps cover $1.375 \leq |\eta| \leq 3.2$. The length of the barrel part is 3.2 m and the radial distance reaches from 2.8 m to 4 m. The material has a depth of more than 24 radiation lengths in the barrel part and more than 22 radiation lengths in the end-caps. The energy resolution was measured to be

$$\frac{\sigma_E}{E} = \frac{10\%\sqrt{\text{GeV}}}{\sqrt{E}} \oplus 0.17\%. \quad (3.1)$$

A good position resolution is also assured [203].

⁶The gas mixture is: 70% Xe, 27% CO₂ and 3% O₂. The length of the straws is 144 cm in the barrel part and 37 cm in the end-caps.

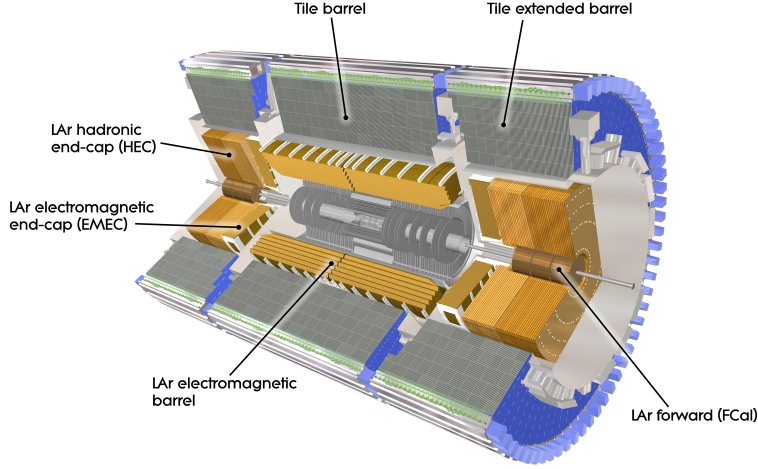


Figure 3.8.: The figure shows a schematic overview of the ATLAS calorimeters [216].

3.2.3. The hadronic calorimeter

Around the E-CAL, the H-CAL is placed (see Figure 3.8). It is again a sampling calorimeter using different active and passive materials. The barrel part is a tile calorimeter, where steel is used as a passive material and scintillator as the active medium. Its length is 5.8 m for the central barrel and 2.6 m for the extended barrels on both sides. It reaches an η -coverage of $|\eta| \leq 1.7$. At $\eta = 0$, a thickness of 9.7 radiation lengths is achieved. In the end cap parts, two different calorimeters are placed, the forward calorimeter and the hadronic end-cap calorimeter (see Figure 3.8). Both are using liquid Argon as the active material. The hadronic end-cap uses copper as the passive material, while the forward calorimeter uses tungsten and copper. The hadronic end-cap covers $1.5 \leq |\eta| \leq 3.2$ and the forward calorimeter covers $3.1 \leq |\eta| \leq 4.9$, allowing for forward jet measurements⁷. The thickness is again around 10 interaction lengths. The performance of the hadronic end-cap is about

$$\frac{\sigma_E}{E} = \frac{21.4\% \sqrt{\text{GeV}}}{\sqrt{E}}, \quad (3.2)$$

while for the forward calorimeter

$$\frac{\sigma_E}{E} = \frac{28.5\% \sqrt{\text{GeV}}}{\sqrt{E}} \oplus 3.5\%, \quad (3.3)$$

was measured. The tile calorimeter gives a resolution of

$$\frac{\sigma_E}{E} = \frac{56.4\% \sqrt{\text{GeV}}}{\sqrt{E}} \oplus 5.5\%, \quad (3.4)$$

giving the needed resolution for the harsh LHC environment [203].

⁷Although no flavour tagging possible as the tracker only covers $|\eta| \leq 2.5$.

3. Experimental setup

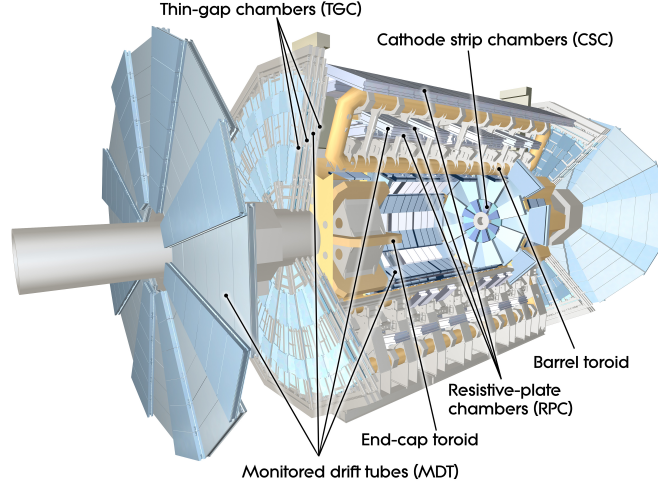


Figure 3.9.: The figure shows a schematic overview of the ATLAS muon spectrometer [217].

3.2.4. The muon spectrometer

The outermost part of the ATLAS detector is the muon spectrometer (MS). It consists of several gaseous detector systems and makes use of a large superconducting toroidal magnet system (see Figure 3.9). As muons are minimal ionizing particles in the context of the LHC, they traverse the calorimeters only losing a bit of their energy. All other particles are stopped in the calorimeters and punch-throughs (hadronic particles escaping the H-CAL) are very rare. To measure the muons a large tracking system was built to be able to measure the bending radius of the muons created by the magnetic field from the toroidal magnets. The central part is equipped with monitored drift tubes (MDT), while at large $|\eta|$ cathode strip chambers (CSC) are used, placed in wheels for the end caps. The muon trigger system uses resistive plate chambers (RPC) for the barrel part and thin gap chambers (TGC) for the end-cap wheels. These detectors do also support the precision tracking done with the MDTs and CSCs. While the trigger system extends to $|\eta| \leq 2.4$, the tracking is done up to $|\eta| \leq 2.7$. The MDTs have a spatial resolution of $35 \mu\text{m}$ and the CSCs achieve $40 \mu\text{m}$. The RPCs and TGCs only achieve a few mm but their signal response comes within 15 to 25 ns, which allows the correct assignment of the signals to the beam-crossing. To benefit from the good spatial resolution in tracking, a very precise alignment of the MDTs and CSCs within a few tens μm is needed. The gas mixture is Argon with a few per cent of CO_2 in the MDTs and 20% CO_2 in the CSCs. The gas mixture in the RPC is $\text{C}_2\text{H}_2\text{F}_4/\text{Iso-C}_4\text{H}_{10}/\text{SF}_6$ with 94.7%, 5% and 0.3%, respectively. The TGCs use CO_2 with n-pentane in a ratio 55% to 45% [203].

One of the key signatures of the ATLAS detector is the toroidal magnet system. The magnets are superconducting (Nb/Ti/Cu, stabilized with Al) and have an air core. They create B-fields of 0.15 to 2.5 T for the barrel and 0.2 to 3.5 T for the end-caps, which means the energy stored in them is 1.6 GJ. The B-field is known to have a precision of

a few mT, which in combination with the tracking precision leads to a muon transverse momentum resolution of 10% for 1 TeV muons [203].

3.2.5. The ATLAS trigger system

Running a detector at 40 MHz, with an event size of about 1.3 MB [203], does not allow storing each bunch crossing. The trigger for Run 2 underwent substantial refurbishment concerning Run 1 [218]. One major difference is that the Level-2 (L2) and Event-Filter (EF) computing farms were merged, to use the computing power more efficiently. As a consequence, the trigger became a two-level trigger, with the Level 1 (L1) trigger being hardware-based and the so-called high-level-trigger (HLT) being software-based. The L1 trigger uses data from the calorimeters and the muon spectrometer to identify candidates of physics objects. Subsequently, the information is processed in the Central Trigger Processor (CTP). It also manages the dead time of the detector. Its primary goal is to reduce the incoming data rate from 40 MHz to 100 kHz. If an event gets accepted by L1, the HLT will request the data from all subdetectors to process the event further. The L1 trigger also defines regions of interest (RoI), which are passed on to the HLT to be investigated. The HLT can decide within a few ms to accept the event or reject it. If it is accepted, the data is passed on to the offline reconstruction and goes into permanent storage [219]. Although the system was designed to store at a rate of 1 kHz [218], it was capable to achieve 1.2 kHz [219].

To define if an event passes the trigger selection or not, several trigger chains are defined, setting the requirements on the different physics objects. If these requirements are fulfilled the event will be recorded. As different analyses are looking for different signatures there are a lot of trigger chains implemented. The events, depending on which triggers they pass are processed in six different types of streams:

- the physics stream for events used in physics analyses,
- the express stream (a small subset of the physics stream) to have fast offline monitoring to ensure data quality,
- the debug streams for events where the HLT failed to make a decision,
- the calibration streams to check subdetectors more closely (only partial event information needed),
- the trigger-level-analysis streams to look at very specific objects (only partial event information needed)
- the monitoring streams to allow for online detector monitoring (no data will be recorded).

In special cases, additional streams can be added. During Run 2, a data quality efficiency of around 95% averaged over the four years could be achieved [219].

3. Experimental setup

3.2.6. Luminosity measurements

To compare the number of measured events N with a predicted cross-section σ , the so-called luminosity $\mathcal{L}$ needs to be known, as following equation holds:

$$N = \sigma \mathcal{L}. \quad (3.5)$$

While the cross-section can be calculated, the luminosity is a feature of the particle accelerator and needs to be measured independently. The most precise luminosity measurement for ATLAS for Run 2 gave a total luminosity of $140.1 \pm 1.2 \text{ fb}^{-1}$ [210]. The absolute luminosity is determined with van-der-Meer-Scans (vdM-Scans) [220, 221]. To account for small changes in the beam conditions, separate forward detectors are used for relative luminosity measurements, hence following an approach outlined in [222]. The main detector for this is the LUCID-2 detector [203, 223], which is supported by the beam conditions monitor (BCM) and offline measurement of the number of reconstructed charged particles and calorimeter information [210]. The LUCID-2 detector is a Cherenkov detector, which uses quartz as the medium and photomultiplier tubes. It is located $\pm 17 \text{ m}$ away from the interaction point along the beam pipe (z -axis) and covers $5.561 \leq |\eta| \leq 5.641$. In principle, LUCID-2 measures the number of interactions per bunch crossing by measuring inelastic pp collisions. The probability P is given by [210, 222, 223]

$$P = \frac{N_{\text{hit}}}{N_{\text{PMT}} N_{\text{BC}}} = 1 - e^{-\epsilon \mu}, \quad (3.6)$$

where ϵ is the efficiency of the photomultiplier tubes and μ the number of collisions per bunch crossing. N_{hit} is the number of all hits found in all photomultiplier tubes, N_{PMT} is the number of photomultiplier tubes and N_{BC} the number of bunch crossings. As the luminosity is given as [210, 222, 223]

$$\mathcal{L} = \frac{f}{\sigma} \sum_{i=1}^{n_b} \mu_i, \quad (3.7)$$

with f being the revolution frequency, σ the total inelastic cross-section and n_B the number of bunch crossings. So it comes down to a relative measurement of μ and an absolute measurement with vdM-Scans.

From data and MC simulations to the final event selection

4.1. Data

For this analysis, the pp data recorded by ATLAS during Run 2 is used, which results in a dataset having an integrated luminosity of 140 fb^{-1} , with an overall uncertainty of 0.83% [210]. Not every single LHC run can be used for physics analysis, as wrongly working or non-working detector components can degrade the quality of the recorded data. Those runs, where data taking happened under conditions being sufficiently good (see Section 3.2.5), were used to build the good run list (GRL). This analysis makes use of the full GRL from Run 2, spanning over the years from 2015 to 2018.

4.2. Monte Carlo simulations

Several Monte Carlo simulations for the signal and background processes are used (see the description below and the full list in Appendix A). The event generation was done with different matrix element generators interfaced with different shower algorithms. The parton shower is needed to describe which particles already decay in the beam pipe and which particles enter the detector. The ATLAS detector simulation was done either with GEANT4 for a full detector simulation or with ATLFAST-II, where hadronic showers in the calorimeters are parametrised to save computation time [224]. A pile-up reweighting was also applied to match the conditions during data-taking in the MC samples. This accounts for additional interactions during the same bunch crossing. Technically, this is done by creating so-called minimum-bias (inelastic) events and superimposing them with the hard scattering event after the parton shower. As some samples contain a prompt photon from the matrix element (called dedicated samples, e.g. samples for the $t\bar{t}\gamma$ signal process), a sample overlap removal procedure was applied to the corresponding inclusive samples (e.g. the $t\bar{t}$ sample), to avoid counting parts of the phase space twice,

4. From data and MC simulations to the final event selection

as the parton shower adds finale state radiation. The procedure is explained in Section 4.2.2. Not all reconstructed photons¹ are prompt photons. Either electrons or positrons may be misidentified as a photon. These are called $e \rightarrow \gamma$ fakes. Furthermore, single hadrons, like pions, can be misidentified as photons. These are called $h \rightarrow \gamma$ fakes. To categorise such events correctly and distinguish them from prompt photons, the MCTruthClassifier² information was used. It is based on a ΔR matching between so-called truth particles and the reconstructed photon. That means the truth particle and the reconstructed photon are very close to each other in the η - ϕ plane. The information about the truth particles comes from the simulation and describes all particles before the detector simulation.

4.2.1. MC samples

$t\bar{t}\gamma$ production

The nominal simulation of the photon emission during the production stage (using the narrow width approximation, see Section 2.6.3) is a $2 \rightarrow 3$ process, simulated at next-to-leading order in QCD. The generator used was MadGraph5_aMC@NLO 2.7.3 [225] with the NNPDF3.0NLO [226] set for the PDFs. The choice for both, renormalisation and factorisation scale, is $H_T/2$, which gives the sum over all “transverse masses”³ of all particles in the final state. For the modelling of the showering, hadronisation, fragmentation and the underlying event, Pythia 8.240 [227] was used with the A14 [228] tune and the NNPDF2.3LO PDF-set. The top quark decay, which is decoupled from the photon emission happening in the production, was done at LO preserving spin-correlations (MadSpin) [229, 230]. Bottom- and charm-quark decays were done with EvtGen 1.6.0. [231]. Kinematic cuts are applied to avoid singularities (infrared and collinear) in the photon radiation. Therefore, a p_T -threshold of 20 GeV was set for leptons and 1 GeV for quarks, while it is 15 GeV for photons. For the photon, “Frixione isolation” is used with $\delta_0 = 0.1$, $\epsilon_\gamma = 0.1$ and $n = 2$ [232]. The sample is normalised to the NLO cross-section given by the simulation. The sample provides renormalisation scale, factorisation scale and PDF variations. The variations are needed to account for the systematic uncertainties of these inputs to the simulation. The simulation is done with a different scale choice (factor of 2 up and down) or a systematically modified PDF, which leads to reweighted events.

In addition, one alternative sample was simulated, where Herwig 7.2.1 [233] was used instead of Pythia 8.240 [227] for the showering. For the uncertainty on the initial state radiation (ISR) a second alternative sample was produced, that modifies the appearance of ISR. Both are normalised to the same cross-section as the nominal sample.

¹For the photon definition, see Section 4.3.

²This is a tool used in the event reconstruction (see Sections 4.2.2 and 4.4) to retrieve the truth object, which was reconstructed as a photon. Information on this tool can be found here: <https://twiki.cern.ch/bin/view/AtlasProtected/MCTruthClassifier>. The page is not publicly accessible.

³The transverse mass is defined as: $m_T = \sqrt{m^2 + p_T^2}$.

$t\bar{t}\gamma$ decay

To simulate the emission of a photon during the decay of a top quark (again making use of the narrow width approximation), first, the $2 \rightarrow 2$ process of $t\bar{t}$ pair production is simulated at LO. Subsequently, the decay of the top quarks is simulated with the requirement of one additional photon. The same choice of MC generators as for the $t\bar{t}\gamma$ production sample is used. As it is a LO-sample, an inclusive K-factor⁴ $K_{NLO,decay}$ was derived for the lepton+jets and dilepton channel. It was calculated by comparing the normalisations from this sample and the $t\bar{t}\gamma$ production sample with the LO $t\bar{t}\gamma$ $2 \rightarrow 7$ inclusive sample (corrected by the K-factor $K_{NLO,2 \rightarrow 7}$ from [1] based on the procedure given in [185]) using

$$K_{NLO,2 \rightarrow 7} \sigma_{LO,2 \rightarrow 7} = K_{NLO,decay} \sigma_{LO,decay} + \sigma_{NLO,production} , \quad (4.1)$$

where σ are the corresponding cross-sections and K the corresponding K-factors. The inclusive LO $t\bar{t}\gamma$ $2 \rightarrow 7$ sample was used for Ref. [1]. Details on the simulation are found therein. The K-factor $K_{NLO,decay}$ was found to be 1.50592. The sample was normalised to its LO cross-section multiplied with the K-factor.

The variations to describe uncertainties on ISR, as well as for renormalisation and factorisation scales and PDFs, are included in the nominal sample. An alternative sample was produced with Herwig 7.2.1 instead of Pythia 8.240 for the uncertainty on the showering algorithm. The sample is normalised the same way as the nominal sample.

 $tW\gamma$

The $tW\gamma$ process is also divided into two samples. The separation is based on the narrow width approximation, corresponding to the photon emission either during production or during decay. This process is in principle sensitive to the $t\gamma$ coupling too, but taken as background in this analysis, because the interest was on $t\bar{t}\gamma$ only. The samples were produced at LO the same way as the $t\bar{t}\gamma$ samples, but using a 5-flavour PDF scheme, which allows for b -quarks in the initial state. The samples are normalised to the generator cross-sections. For modelling uncertainties, the samples provide variations on both scales (renormalisation and factorisation), and alternative samples with Herwig 7.2.1 for the showering (instead of Pythia 8.240) were generated.

 $V\gamma$

NLO QCD Monte Carlo samples for $V\gamma$ ($V = W, Z$), with a prompt photon generated in the matrix element, were produced with Sherpa 2.2.8 [234] using again the NNPDF3.0NLO PDF-set [226]. One additional jet was simulated at NLO and up to 2 more jets at LO. These samples were normalised to the cross-section given by the generator. The whole chain of the event generation up to the final hadronisation is done with Sherpa. The showering in Sherpa is based on Catani-Seymour dipoles [235–237].

⁴A K-factor describes the ratio between the NLO cross-section and the LO cross-section to normalise events generated at LO to their NLO cross-section, assuming no shape dependence.

4. From data and MC simulations to the final event selection

The matching of the matrix element with the parton shower is done in the MEPS@NLO prescription [238]. Virtual NLO QCD corrections are coming from OpenLoops [239, 240].

$t\bar{t}$

The generation of inclusive $t\bar{t}$ events was done with Powheg-Box-v2 [241–243] at NLO in QCD. The showering was done with Pythia 8.230. The same PDF sets and b/c-quark decay algorithms as before were used, as well as the showering tune and the corresponding PDF sets (see the details on the $t\bar{t}\gamma$ sample generation). The cross-section is normalised to a calculation done with Top++2.0 at NLO in QCD, adding NNLL resummations [244]. Alternative samples were produced to have a variation on the h_{damp} parameter. This parameter controls the scale of the first additional emission in the parton shower. Its nominal value is 1.5 times the top quark mass, while the variation uses three times the top quark mass. Further, there is an additional sample for ISR variations and an additional sample for showering (use Herwig 7 instead of Pythia 8). Scale variations are provided with the nominal sample.

$t\bar{t}V$

The samples for $t\bar{t}$ pair production in association with a heavy vector boson (Z or W) are generated at NLO in QCD, making use of the same MC generators that were used for the $t\bar{t}\gamma$ samples. The samples are normalised to NLO QCD and EW predictions [245].

Single-top quark

The three different production modes of single-top production (s-channel, t-channel, and Wt) are all simulated in dedicated samples. The simulation was done at NLO in QCD with PowhegBox [241–243] using the NNPDF2.3LO PDF set. The rest of the simulations (hadronisation, b/c-quark decays) were done as they were done for the $t\bar{t}$ sample. The samples are normalised to the NNLO cross-section using K-factors [246–248].

Scale variations for the Wt production which are provided with the nominal sample. The showering uncertainty is taken into account by an alternative sample using Herwig 7 instead of Pythia 8.

V +jets

The inclusive samples for heavy vector boson production ($V = W, Z$) in association with jets were done with Sherpa 2.2.1 [234] at NLO in QCD. Up to two additional jets are simulated at NLO and up to three more jets are simulated at LO. Everything else is done in the same way as for the $V\gamma$ samples. The cross-section is taken from an NNLO prediction in QCD and NLO in EW [249].

Diboson

Events with two heavy vector bosons, namely WW , ZZ and ZW events, are generated with Sherpa 2.2.2 [234], in case both bosons decay leptonically and Sherpa 2.2.1 for the other samples. The standard PDF set NNPDF3.0NLO is used. The whole chain of the event generation up to the final hadronisation is done within Sherpa. The samples are normalised to NLO cross-sections in QCD [250].

4.2.2. Sample overlap removal

There are several samples where the simulation is done without the photon already present at matrix element level. This means that photon radiation is simulated in the parton shower. These samples are also called inclusive samples. In this analysis, the inclusive samples are the following ones: $t\bar{t}$, Wt , V +jets, single top s-channel and t-channel, $t\bar{t}V$ and diboson. Because single top s- and t-channel, $t\bar{t}V$ and diboson are small backgrounds (see Section 4.4.3) only inclusive samples were taken. For the signal process and the major backgrounds, so-called dedicated samples exist. In these samples, the photon is simulated at the matrix element level, which is more accurate than producing photons via FSR and ISR in the showering. This leads to the situation that the phase space of the dedicated and inclusive samples do overlap but as the phase space covered by the inclusive samples is larger one cannot simply use the dedicated sample only. On the contrary, one cannot use the inclusive samples only as they are less accurate. To avoid double-counting of events from the dedicated and inclusive samples, which fall into the same region of phase space, a so-called overlap removal procedure is applied. This procedure is provided as a central algorithm and called VGammaORTool⁵. The algorithm only acts on the inclusive samples, where the region of phase space overlapping with the dedicated samples shall be removed. This assures that the more accurate samples are fully taken into account, while the inclusive samples only cover the region of phase space not covered by the dedicated samples.

First, the overlap region must be defined, which is given by the phase space cuts of the dedicated samples, which are:

- Photon $p_T \geq 15$ GeV,
- $\Delta R(l, \gamma) \geq 0.2$.

Next, the photons that need to be taken into account, need to fulfill the following criteria:

- The photon must be a true photon and not a misidentified particle that was reconstructed as a photon.
- The photon needs to be a stable final state particle, hence not a virtual photon.
- $p_T \geq 3$ GeV.

⁵Details on this tool can be found here: <https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/VGammaORTool>. The page is not publicly accessible.

4. From data and MC simulations to the final event selection

- The photon needs to be a prompt photon from the event generation and not a photon from the detector simulation.
- The photon shall not originate from hadrons or τ -leptons.

The information about the “true” origin of a reconstructed photon is retrieved with the MCTruthClassifier tool, using generator-level information.

For the leptons taken into account, the following requirements are imposed:

- The lepton is a real lepton.
- The lepton must come from the hard interaction.
- If the lepton is a τ -lepton, none of its decay products must be a τ -lepton. This ensures that the τ is taken in the state before it decays.
- For electrons and muons the parent particle cannot be a τ .

First, the photons are checked to have $p_T \geq 15$ GeV. If a photon fulfils this criterion, the ΔR criterion is checked for every lepton taken into account. If both criteria are fulfilled the photon is considered to be in the overlap region and therefore the event will be discarded. The fraction of all generated events that do survive the overlap removal is about 5%.

4.3. Object definitions

If a bunch crossing leads to a hard interaction, most particles, especially particles like top quarks, b -quarks, W and Z bosons, decay in the beam pipe. Decay and hadronisation products enter the detector, leaving tracks and energy deposits. Not all of these particles, entering the detector, are stable. Some particles like pions or kaons will decay in the detector. The digital signals that the different detector components record, need to be reconstructed. The reconstructed particles are the objects that will be used for any analysis. They represent the final state particles from the hard scattering event.

Due to the hadronic environment of the pp collisions, one needs to separate which detector signals correspond to the hard event and which signals are pile-up contamination and the underlying event. To do so, primary vertices are identified from tracks that can be interpolated to the interaction region. The primary vertex that defines the hard event is the vertex with at least 2 tracks and the highest p_T -sum formed by all tracks with $p_T > 500$ MeV.

The principal objects that are reconstructed and connected to the primary vertex are the following ones: charged leptons (e, μ, τ), photons, muons, jets and E_T^{miss} . Everything else, like reconstructed top quarks or neutrinos, for example, is derived from these main objects. For the preparation of data and simulation the AnalysisTop⁶ software was

⁶Information on this tool can be found here:

<https://gitlab.cern.ch/atlas/athena/-/tree/main/PhysicsAnalysis/TopPhys/xAOD>.

used, which takes the reconstructed objects, calibration information, and scale factors to produce so-called ntuples. To allow for the combination with the $t\bar{t}Z$ analysis [251], the same object definitions, with the same uncertainty schemes were used. The details of object reconstruction are described in the following. As τ -leptons do not play a role in this analysis, their reconstruction is not further explained.

4.3.1. Photons

The analysis aims explicitly at analysing the coupling between the top quark and the photon. Consequently, the photon is one of the main objects to be studied. The most important piece of information is the energy deposits in the ECAL. If these cannot be connected to any track that photon is labelled as an unconverted photon. In case of a photon undergoing conversion into an e^+e^- pair, one needs two showers in the calorimeter which can be connected to individual tracks in the ID and both tracks must have an intersection point where the pair creation happened. To have a high background rejection⁷ in the photon reconstruction and identification, isolation and identification criteria are applied. Identification criteria are based on variables describing the measured shower, so-called shower shape variables. Two sets of criteria are available, a so-called tight set and a loose set, where the tight set gives a better background rejection. This analysis uses tight identification. The full list of criteria for tight and loose identification and more details on the photon reconstruction can be found in [252]. For the isolation the following is required: $E_T^{iso}|_{\Delta R < 0.4} < 0.022 \cdot E_T(\gamma) + 2.45 \text{ GeV}$ and $p_T^{iso}|_{\Delta R < 0.2} < 0.05 \cdot E_T(\gamma)$ [252]. The first criterion defines how much transverse energy E_T^{iso} , which does not belong to the photon, is allowed to be in a predefined ΔR cone around the photon. The second criterion is a similar requirement for the transverse momentum p_T^{iso} . The energy and transverse momentum are reconstructed and calibrated as described in [253]. Corresponding scale factors are derived. These scale factors do account for inefficiencies in the photon reconstruction and are derived by data-driven methods [253]. In general scale factors are applied to account for missmodelling in the simulations. A photon candidate is only accepted as a photon in case it has $p_T > 15 \text{ GeV}$, $|\eta| < 2.37$.

4.3.2. Electrons

The reconstruction of electrons (e^+ and e^-) is done as a combination of tracking information and calorimeter information. Only if a shower can be matched to a track, a possible electron candidate is found. While the bending radius of the considered electron track is used to measure p_T , the corresponding energy deposits from the calorimeter that can be matched to the electron track give the E . The track and the location of the shower give the information on η and ϕ . From different shower shape variables, a likelihood is created in bins of E_T and η . The value of the likelihood defines if the electron candidate is accepted or rejected, depending on the threshold that is imposed. For the identification, the medium threshold is used, while a tight and a loose option is also available. More

⁷More details on fake photos are given in the following sections of this chapter.

4. From data and MC simulations to the final event selection

details on this procedure can be found in [254]. The isolation requirement is chosen to be loose, using the PLV (prompt lepton veto) method. This is a boosted decision tree (BDT) to discriminate between prompt leptons and non-prompt leptons (from weak decays in a jet for example) [255]. For the calibrated electrons corresponding scale factors are derived [253]. Hard cuts are applied to the electron candidates on the transverse momentum, the impact parameter d_0 and the z -axis difference between the track and primary vertex. The cuts are $p_T > 7$ GeV, $|d_0|/\sigma_{d_0} < 5$ and $|\Delta z \sin(\theta)| < 0.5$ mm. For the pseudorapidity $|\eta| < 2.47$ is required.

4.3.3. Muons

Muons are reconstructed from track information from the ID and the muon spectrometer, which are required to match each other (combined muons). The background is coming mainly from muons coming from charged hadron decays (soft muons). Track fits and momentum measurements in general only fit poorly together for the ID and the MS. This leads to a good differentiation between prompt muons and soft muons. Corrections due to the energy loss in the calorimeters are applied. The criteria required by the analysis for muons are medium identification [256] and loose PLV isolation [255]. Corresponding scale factors are taken into account [256]. The cuts are $p_T > 7$ GeV, $|d_0|/\sigma_{d_0} < 3$ and $|\Delta z \sin(\theta)| < 0.5$ mm. For the pseudorapidity $|\eta| < 2.5$ is required.

4.3.4. Jets

Jets are sprays of hadrons and leptons formed in the process of hadronisation, as bare quarks do not exist⁸ due to confinement. These collimated sprays of particles are merged into jets as the object which is being used. The algorithm to merge the single particles into a jet is the anti- k_t -algorithm with a radius of $R=0.4$ [257]. The information used by the algorithm are so-called particle-flow objects, which already take the information from the ID and the calorimeters into account. These objects represent the particles, that form the jet [258]. A calibration procedure to the reconstructed jets is applied [259]. This is necessary as the measurement of the jet's 4-vector is skewed by many detector effects which this calibration procedure takes into account, to give better measurements. The jet candidates need to fulfill $p_T > 25$ GeV and $|\eta| < 2.5$. To reject jets from pile-up events only jets associated with the primary vertex of the hard event are taken into account. Therefore, a procedure called jet-vertex-tagging is used and corresponding scale factors are taken into account [260]. The algorithm constructs a likelihood that discriminates between jets from the primary vertex of the hard events and other primary vertices (pile-up). For this analysis, the likelihood value is required to be above 0.59 for jet $p_T < 60$ GeV and $|\eta| < 2.4$. For other jets fulfilling the general jet definition used in this analysis, pile-up rates are in general low enough.

As top quarks preferentially decay into b -quarks and W bosons, an identification of jets coming from b -quarks is very useful to suppress background, especially from events with

⁸Here top quarks are not taken into account as they decay too fast and do not form jets.

W and Z bosons, as their decays into b -quarks are relatively rare. Because the decay of B -hadrons is CKM suppressed, they can travel several mm in the beam-pipe before they decay. This creates secondary vertices and hence the information about them is quite useful for b -tagging. As track information from the ID is required for it, this limits b -tagging to jets within $|\eta| < 2.5$ and therefore forward jets are not taken into account. To tag b -jets a multiclass neural network was trained called DL1r and different working points were calibrated and defined [261, 262]. Besides the working points with a fixed efficiency, also the full distribution with 5 bins is calibrated, where each bin contains those events that passed the previous working point but not the next tightest working point. This is called pseudo-continuous b -tagging (PCBT).

4.3.5. E_T^{miss}

Collisions in the LHC happen as head-to-head collisions. This means the colliding partons (quarks and gluons within the proton) have negligible momentum in the transverse plane. From energy-momentum conservation, it follows that the measured transverse energy/momentum should also be close to 0. If the event contains particles that the detector does not detect (e.g. neutrinos or hypothetical BSM particles) these can be identified by measuring an amount of missing transverse energy E_T^{miss} . This quantity is computed by taking the negative vectorial sum of all reconstructed physics objects and from the track and calorimeter signals that were not taken into account for any of the reconstructed objects (soft term) [263]

$$E_T^{miss} = - \sum_{\text{electrons}} p_T^e - \sum_{\text{photons}} p_T^\gamma - \sum_{\tau\text{-leptons}} p_T^\tau - \sum_{\text{muons}} p_T^\mu - \sum_{\text{jets}} p_T^j - \sum_{\text{unused tracks}} p_T^{\text{track}}. \quad (4.2)$$

The algorithm is described in [263, 264]. It is guaranteed that no detector signal is counted twice. This is implemented by a procedure that takes the different objects in a predefined way into account, which is reflected in the formula given above. As a consequence, the object definitions have a crucial role, especially the jet definition, as reconstructed jets undergo a dedicated calibration procedure (see above). This has a direct impact on the result of the E_T^{miss} calculation and the corresponding uncertainties. The resolution of the E_T^{miss} calculation is determined with MC simulations comparing the reconstructed E_T^{miss} and the true E_T^{miss} [263, 264].

4.3.6. Object overlap removal

In principle, there is the possibility that either the same track information or the same calorimeter information is used for the reconstruction of two different objects. To avoid this double-counting a so-called overlap removal procedure is applied. The procedure follows several steps:

- In case an electron and a muon are assigned to the same track, the electron is removed and the muon is kept.

4. From data and MC simulations to the final event selection

- In case electrons and jets are close to each other the following two-step procedure is applied. First, it is checked if the jet is closer to the electron than $\Delta R = 0.2$. If that is the case, the jet is removed. If this is not the case, it is checked if the electron is around the jet within a distance of $\Delta R = 0.4$. If that is the case, the electron is removed.
- A similar strategy is used for jets and muons. If the jet has only one or two tracks and is within a distance of $\Delta R < 0.2$ to the muon, the jet is removed. Otherwise, the same requirement as for electrons applies.
- Next, if the photon is closer than $\Delta R = 0.4$ to the lepton, it is removed.
- The last step is to check if there is a jet within $\Delta R = 0.4$ to a photon. Such a jet will be removed.

In case in one of the steps an object is removed, it is not taken into account in the following steps.

4.4. Event reconstruction and selection

For each event in the data and the MC samples, the object reconstruction is done as described above but not all events are used for further analysis. To store an event a so-called pre-selection is done. First of all, only events for which a single-lepton trigger (electron or muon) fired are considered. This is independent of the decay channel, being either the lepton+jets channel or the dilepton channel.

For the lepton+jets channel the following additional requirements were made:

- At least four jets are required.
- Leptons need to have at least 25 GeV of p_T and exactly one of them is required. Events containing further leptons below that energy threshold are removed.
- Photons need to have at least 20 GeV of p_T and only events with exactly one such photon are accepted. Similarly to the leptons events containing additional photons below the energy threshold are removed.

In the case of the dilepton channel, the requirements are slightly different:

- At least two jets are required.
- The triggered lepton needs to have $p_T \geq 25$ GeV, while the second lepton is required to have $p_T \geq 20$ GeV. Exactly two leptons are required in any combination (ee , $\mu\mu$ and $e\mu$) with opposite charge.
- The requirements of the photons are the same as in the lepton+jets channel.

These rather loose requirements for the pre-selection were chosen to have the possibility to study backgrounds (see Chapter 5.2). The final signal selection will be given in Section 4.4.3.

4.4.1. Photon categorisation

In Section 4.3.1 it was already mentioned that reconstructed photons can be other particles that were misidentified by the reconstruction algorithm. There are two sources of such mis-reconstructed photons. One source are electrons and the other source are hadrons, like pions or even jets. Misidentified electrons are called $e \rightarrow \gamma$ fakes, while hadrons/jets are called $h \rightarrow \gamma$ fakes. To identify such events in the MC samples to get a prediction for them, MC truth information is used⁹. It is checked if the reconstructed photon can be associated with a true photon at particle level. If this is the case, the photon is labelled as a prompt photon. In case the reconstructed photon cannot be matched to a true photon but to a true electron it is labeled as an $e \rightarrow \gamma$ fake. A photon is labelled as a $h \rightarrow \gamma$ fake if the reconstructed photon can be connected to a truth jet. These labels are given to the MC events without taking into account the process in which the fake photon appeared. While processes with prompt photons are described sufficiently well by the MC simulations, the number of events with photon-fakes is in general underestimated. Therefore further background studies were done and corresponding data-driven scale factors to correct for the underestimation were derived (see Chapter 5.2).

4.4.2. Event categorisation

In the following, it will be described how the MC events are categorised in general depending on the process and the photon label (prompt, $e \rightarrow \gamma$ fake or $h \rightarrow \gamma$ fake):

- $t\bar{t}\gamma$ from production: event with a prompt photon found in the $t\bar{t}\gamma$ from production MC samples. Sometimes abbreviated as $t\bar{t}\gamma$ (prod.) or $t\bar{t}\gamma$ production.
- $t\bar{t}\gamma$ from decay: event with a prompt photon found in the $t\bar{t}\gamma$ from decay MC samples. Sometimes abbreviated as $t\bar{t}\gamma$ (dec.) or $t\bar{t}\gamma$ decay.
- Other γ : event with a prompt photon found in the following background processes:
 - $Wt\gamma$
 - Wt (event was not rejected in sample overlap removal)
 - $t\bar{t}$ (these are most likely photons from the showering and hence there is no sensitivity to the top-photon coupling)
 - Single top (s- or t-channel)
 - $t\bar{t}V$
 - Diboson
 - $V\gamma$
 - V +jets (event was not rejected in sample overlap removal)

⁹The same procedure was already used in [1, 2].

4. From data and MC simulations to the final event selection

Table 4.1.: Final selection requirements for the e +jets and μ +jets channel.

e +jets	μ +jets
1 photon with $p_T > 20$ GeV, $ \eta < 2.37$ and $ \eta \notin [1.37, 1.52]$	
1 e with $p_T > 25$ GeV $ \eta < 2.47$ and $ \eta \notin [1.37, 1.52]$	1 μ with $p_T > 25$ GeV $ \eta < 2.5$
≥ 4 jets with $p_T > 25$ GeV and $ \eta < 2.5$	
1 b -tagged jet at the 70% working point	
$ m(e, \gamma) - m_Z > 5$ GeV ¹⁰	-
$\Delta R(e \text{ or } \mu, \gamma) > 0.4$	

- $e \rightarrow \gamma$ fakes from all processes mentioned above, sometimes abbreviated as e-fakes or e-fake γ .
- $h \rightarrow \gamma$ fakes from all processes mentioned above, sometimes abbreviated as h-fakes or h-fake γ .
- Lepton fakes (only relevant in lepton+jets channel; see Section 5.2.4), sometimes abbreviated as Lep. fakes.

4.4.3. Final event selection

Both channels, the lepton+jets channel and the dilepton channel, have advantages and disadvantages and so there are reasons to look at both. The lepton+jets channel provides much more statistics but also a larger background, while the dilepton channel in principle does the contrary. It has lower statistics but is much purer. In the end, both channels will be measured in a combined way to profit from both. In contrast, the all-hadronic channel was not considered as its multijet QCD background is too large to get a comparable precision with the other two channels.

Lepton+jets channel

The final event selection in the lepton+jets channel is given in Table 4.1. While the η -cuts on the different objects are given by the detector geometry, the lepton p_T -cuts are imposed by the triggers. The number of jets is given by their expected number at LO (see Section 2.5.1). The b -tagging working point [261, 262], which describes the efficiency to tag a real b -jet, was studied and it was found that in comparison to Ref. [1], the statistics allow to select a tighter working point, namely at 70% instead of 77%. This allows for a better background rejection of mis-tagged light jets in processes without expected b -jets. The ΔR between photon and lepton imposed in [1] was loosened as the separation of $t\bar{t}\gamma$ from production and decay is done with a neural network. The Z -mass window in the e +jets channel suppresses $e \rightarrow \gamma$ fakes from Z +jets events. The yields of the different categories are given in Table 4.2 and a breakdown of the photon fake yields from the corresponding processes is given in Table 4.3.

4.4. Event reconstruction and selection

Table 4.2.: Event yields and their uncertainties (statistical + systematic (see Section 5.4)) for the different event categories in the lepton+jets channel.

Process	Yields
$t\bar{t}\gamma$ from production	12450 ± 740
$t\bar{t}\gamma$ from decay	13400 ± 3100
Other γ	6300 ± 1700
$e \rightarrow \gamma$ fakes	6920 ± 980
$h \rightarrow \gamma$ fakes	3600 ± 1200
Lepton fakes	640 ± 100
Total	43900 ± 4600
Data	47767

Table 4.3.: Photon fake yields per process in the lepton+jets channel. Dedicated and inclusive processes are listed separately. The individual uncertainties were not derived but only the total, as presented in Table 4.2.

Process	$e \rightarrow \gamma$	$h \rightarrow \gamma$
$t\bar{t}\gamma$ from production	45.7	11.0
$t\bar{t}\gamma$ from decay	34.8	15.1
$t\bar{t}$	6026.5	3280.4
$Wt\gamma$	3.7	1.3
Wt	189.7	108.3
$W\gamma$	0	2.8
W +jets	0.5	67.8
$Z\gamma$	24.6	0.5
Z +jets	487.6	24.9
$t\bar{t}V$	79.6	16.5
single top (s-,t-channel)	1.6	29.5
VV	27.5	5.0

Dilepton channel

The final event selection in the dilepton channel is given in Table 4.4. The selection is based on the selections made in [1, 2]. The reasons for the η and p_T cuts are the same as in the lepton+jets channel. For the second lepton, the p_T cut was loosened a bit, to take more statistics into account. Both leptons need to have opposite charges as they are expected to come from a W^+ - and a W^- -boson. The minimum number of required jets matches again the number of expected jets at LO. For the b -tagging working point, it was decided to use the one with 85% efficiency, which gave the best trade-off between signal efficiency and background rejection. The Z -mass window cuts on the invariant masses help to reduce again Z +jets events. The cut on E_T^{miss} accounts for the two neutrinos that are expected and the lower boundary on the dilepton invariant mass helps to suppress background from low mass resonances. The yields for each category are given in Table 4.5 and a breakdown of the photon fake yields from various processes is given in Table 4.6.

¹⁰ $m_Z = 91.19$ GeV.

4. From data and MC simulations to the final event selection

Table 4.4.: The final selection requirements for the ee , $\mu\mu$ and $e\mu$ channels.

ee	$\mu\mu$	$e\mu$
1 photon with $p_T > 20$ GeV, $ \eta < 2.37$ and $ \eta \notin [1.37, 1.52]$		
1 e with $p_T > 25$ GeV 1 e with $p_T > 20$ GeV $ \eta < 2.47$ and $ \eta \notin [1.37, 1.52]$	1 μ with $p_T > 25$ GeV 1 μ with $p_T > 20$ GeV $ \eta < 2.5$	1 μ or 1 e with $p_T > 25$ GeV 1 e or 1 μ with $p_T > 20$ GeV $ \eta < 2.47$ and $ \eta \notin [1.37, 1.52]$ for e $ \eta < 2.5$ for μ
Both leptons with opposite electric charges.		
≥ 2 jets with $p_T > 25$ GeV and $ \eta < 2.5$		
1 b -tagged jet at the 85% working point		
$\Delta R(e \text{ or } \mu, \gamma) > 0.4$		
$ m(l, l) - m_Z > 5$ GeV		-
$ m(l, l, \gamma) - m_Z > 5$ GeV		-
$E_T^{miss} > 30$ GeV		-
$m(l, l) > 15$ GeV		

Table 4.5.: Event yields and their uncertainties (statistical + systematic) for the different event categories in the dilepton channel.

Process	Yields
$t\bar{t}\gamma$ from production	2390 ± 110
$t\bar{t}\gamma$ from decay	3110 ± 640
Other γ	1120 ± 220
$e \rightarrow \gamma$ fakes	57.9 ± 7.1
$h \rightarrow \gamma$ fakes	219 ± 82
Total	6900 ± 720
Data	7379

Table 4.6.: Photon fake yields per process in the dilepton channel. Dedicated and inclusive processes are listed separately. The individual uncertainties were not derived but only the total, as presented in Table 4.5.

Process	$e \rightarrow \gamma$	$h \rightarrow \gamma$
$t\bar{t}\gamma$ from production	5.1	0.8
$t\bar{t}\gamma$ from decay	2.5	0.8
$t\bar{t}$	14.4	193.0
$Wt\gamma$	0.5	0.2
Wt	1.4	6.6
$W\gamma$	0	0
W +jets	0	0.1
$Z\gamma$	0	3.3
Z +jets	0.9	10.1
$t\bar{t}V$	17.3	2.8
single top (s-,t-channel)	0	0.1
VV	15.1	1.1

4.5. Fiducial phase space at particle level

Differential cross-sections, as described in Section 6.4.1, are measured at particle level. It is the stage of the event after hadronisation and before the particles enter the detector. For doing such a measurement, a procedure called unfolding (correcting for the detector effects; for the details see Section 6.4.1) is applied. For that, one needs to define the fiducial phase space within which the unfolding happens. It should be as close as possible to the phase space at the reconstruction level to minimise the dependence of the extrapolation on the simulation. Hence, the object definitions at the particle and reconstruction level are also as similar as possible. Here are the definitions of the objects at the particle level:

- For the photon a p_T threshold of 20 GeV is applied as well as an η -requirement of $|\eta| < 2.37$. To have a prompt and isolated photon, the photon must not come from a hadron decay and p_T from other particles around the photon within a cone of $\Delta R = 0.2$ must be smaller than 5% of the photon's p_T .
- For electrons and muons, the p_T and η requirements are $p_T > 25$ GeV and $|\eta| < 2.5$. Additionally, there is a lower p_T requirement, that lepton candidates need to have at least 7 GeV. The leptons take photons that are within a cone of $\Delta R = 0.1$ and do not come from hadrons into account. The leptons themselves are also required not to come from hadron decays.
- Jets are required to have $p_T > 25$ GeV and $|\eta| < 2.5$. They are built with the anti- k_t -algorithm with $R = 0.4$ and non-interacting particles and muons are not considered for the clustering.
- b -jets are defined as follows: a jet contains a B-hadron with $p_T > 5$ GeV. If a jet contains such a B-hadron is checked via ghost matching [265].
- Overlap removal is done between muons and jets, electrons and jets, and photons and jets. In each case, the radius taken into account is $R = 0.4$ and if the two objects are closer than this radius the jet is always removed.

Following the object definitions at the particle level, the fiducial phase space needs to be defined. In the lepton+jets channel one requires exactly one photon, exactly one lepton, at least four jets and out of those at least one jet is a b -jet. In the dilepton channel, the requirements are exactly one photon, two leptons with opposite charges, at least two jets and again at least one of them must be a b -jet.

5.1. General outline

In the previous chapters the theory, the experimental setup and the technical setup were described. This sets the scene for the actual analysis. Therefore, this chapter will now serve as an overview of the different steps towards the final results. The author's contributions to these results were listed at the beginning of the thesis.

To be able to measure the signal process ($t\bar{t}\gamma$ from production) it is necessary to have an accurate description of the background. This is achieved with different methods, of which some are partially or completely data-driven (see Section 5.2).

As soon as a good background description is available, one can work on the separation of signal ($t\bar{t}\gamma$ from production) from the background. This analysis uses neural networks for this purpose. In the case of the lepton+jets channel, a multiclass classification was implemented, while for the dilepton channel, a binary classification was sufficient enough. The difference in the approach is related to the signal-to-background ratio (see Section 5.3). The neural network output is then used to define the regions (signal regions (SR) and control regions (CR)) for the profile likelihood fit to extract the inclusive cross-sections (see Section 5.3 and Sections 6.1 to 6.3). The model of systematic uncertainties is described in Section 5.4. The inclusive cross-section is measured for $t\bar{t}\gamma$ from production alone and for $t\bar{t}\gamma$ from production and decay together¹. These two measurements are done in the lepton+jets channel, the dilepton channel and a combined measurement in both channels. The profile likelihood fit allows the extraction of the so-called signal strength $\mu = \sigma_{\text{measured}}/\sigma_{\text{SM}}$. With the prediction of the cross-section σ_{SM} in the fiducial phase space, the measured cross-section σ_{measured} is extracted.

Besides the inclusive cross-sections, which are the main results of this thesis, the whole

¹This allows to have a comparison to previous measurements, which did not separate $t\bar{t}\gamma$ from production and decay.

5. Analysis Strategy

analysis, as presented in Reference [3] and shown here for completeness, also aims at measurements of differential cross-sections (see Section 6.4.1). While the inclusive cross-section is mostly sensitive to the charge, the differential cross-sections have sensitivity to the structure of the coupling. Both, the inclusive and differential cross-section measurements and their uncertainties only reveal how standard-model-like the measured process is. An additional EFT-fit is done to quantify how large the deviations might be (see Section 6.4.2). As a last step, the analysis presents results of a combined $t\bar{t}\gamma$ from production and $t\bar{t}Z$ EFT-fit and a ratio measurement of the two processes. The reason for this is the complementary information regarding the electroweak couplings that are provided by the two analyses (see Section 6.4.3).

5.2. Background modelling

Several background processes were studied to ensure their reliable modelling. These backgrounds are the $V\gamma$ processes, the photon fakes ($e \rightarrow \gamma$ and $h \rightarrow \gamma$), and the lepton fakes. The following sections will describe the treatment of these backgrounds². For all backgrounds but the lepton fakes, the same treatment was also used for the $t\bar{t}\gamma$ charge asymmetry analysis [146].

5.2.1. Prompt photon background

One of the main backgrounds are events with prompt photons. Processes such as diboson, $t\bar{t}V$ or single-top (s- and t-channel) only contribute a small number of expected events, due to their distinct signatures. $Wt\gamma$, $W\gamma$ and $Z\gamma$ events are the main sources of the prompt photon background in the lepton+jets and dilepton channels. While for $Wt\gamma$ modelling uncertainties are taken into account, the modelling of $W\gamma$ and $Z\gamma$ was checked in dedicated control regions described below. Scale factors for the fake backgrounds were applied accordingly in both control regions. Besides $Wt\gamma$, the $t\bar{t}$ process was not investigated with a dedicated control region, although a contribution to the signal region is expected. The reason is again that dedicated modelling uncertainties are taken into account (see Section 5.4), while the other sources of prompt photon events get only overall normalisation uncertainties assigned.

$W\gamma$ control region

For the $W\gamma$ control region ($W\gamma$ CR), which needs to be statistically independent from the $t\bar{t}\gamma$ signal region³, there are two competing factors. On the one hand, it is desirable to have a control region that is pure, but on the other hand, the kinematic properties of the studied background in that region should be similar to those in the signal region. The

²The studies on the $V\gamma$ background were done by the author of this thesis, while the other backgrounds were studied by other members of the analysis team. Their results are reported here for completeness.

In general, Chapter 5.2 is based on the ATLAS internal note of this analysis.

³In general, control regions need to be statistically independent to signal regions.

reason is that one needs to understand the behaviour of the respective background in the signal region without using it. In addition, the statistics should be sufficient. Purity for the $W\gamma$ -process is achieved by requiring fewer jets, as this removes events with high jet multiplicities like $t\bar{t}\gamma$, where already at LO four jets are expected. Therefore, the jet multiplicity requirement was set to three jets or fewer. To keep the kinematics of the events at least similar to those in the $t\bar{t}\gamma$ signal region some requirements on the flavour-tagging of the jets are necessary. To have enough statistics in the control region and a sizable number of b -jets the b -tagging working point was set to 85% efficiency. To be statistically independent of the signal region, the jets must not pass the working point with an efficiency of 70%. For the e +jets channel the Z -mass window was increased to [70,100] GeV. All other cuts are the same as for the $t\bar{t}\gamma$ signal region. The control regions were divided into e +jets and μ +jets regions to see possible mismodelling.

In general, a clear normalisation offset of about 25% is observed for all variables that were investigated. As examples, Figures 5.1 to 5.4 are shown, which show the photon p_T and η distributions⁴, as well as the transverse W -mass $m_T = \sqrt{m^2 + p_T^2}$ and the E_T^{miss} distributions. Even for variables where lepton-fake events do accumulate in a certain interval, like E_T^{miss} for example, this behaviour was observed (see Figure 5.4). Further plots are shown in Appendix C. First, this shows that the lepton-fake estimation (see Section 5.2.4) seems to be quite robust and second, this is a strong sign that the interpretation of the data-MC disagreement to be a normalisation issue is correct. A direct fit of the $W\gamma$ CR was tested and did not show an improvement of sensitivity compared to assigning a large Gaussian prior of 50% to the normalisation uncertainty of the $W\gamma$ contribution. Hence, it was decided to assign the large prior, to keep the fit simple (see Chapter 5.4 and Section 6.2).

$Z\gamma$ control region

For the $Z\gamma$ control region ($Z\gamma$ CR), similar considerations were made, about purity, statistics and kinematic similarity between signal and control region. Because of the second lepton, it is much easier to construct a pure $Z\gamma$ CR by requiring not more than 3 jets and zero b -tags. Furthermore, no cuts on the Z -mass and on E_T^{miss} (no neutrinos are expected) were applied. All other cuts are identical to the $t\bar{t}\gamma$ signal region. A good data-MC agreement is observed throughout all investigated variables. A few of them are presented in Figures 5.5 to 5.8, showing the photon p_T and η distributions as well as a distribution of ΔR between the photon and the leading lepton and the invariant mass distribution of the photon and the nearest lepton. More plots are available in Appendix C. Consequently, no corrections are applied to this background in the $t\bar{t}\gamma$ signal region.

⁴These distributions are so called pre-fit distributions (as any other distribution labeled as pre-fit) as they show uncertainties before any potential fit was done, which could affect them.

5. Analysis Strategy

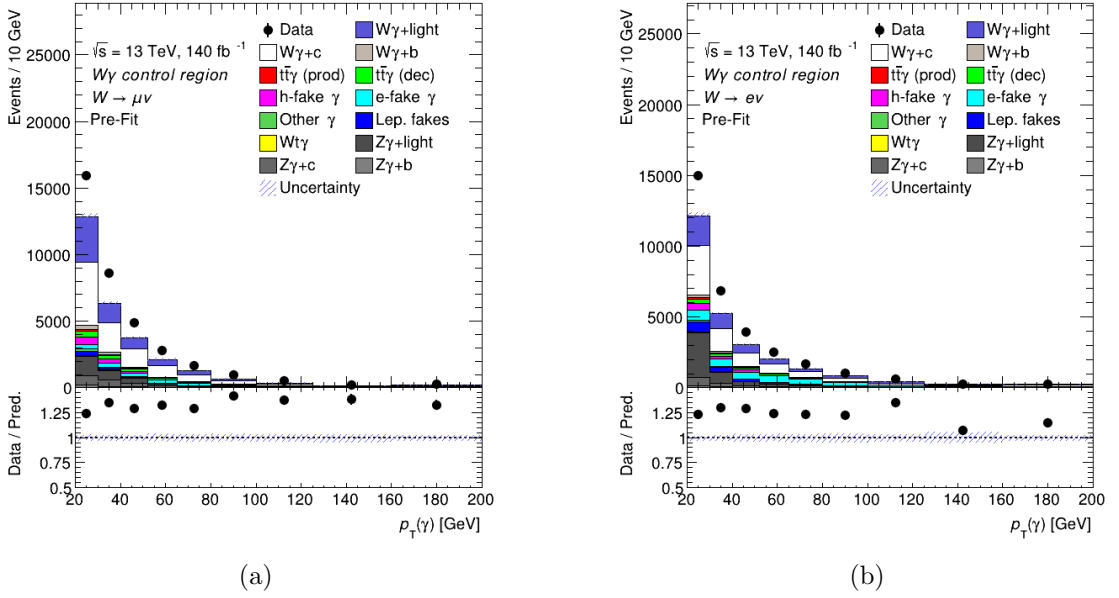


Figure 5.1.: Distributions for the $W\gamma$ CR, showing the p_T distributions of the photon for the μ +jets (a) and e +jets (b) subchannels. The uncertainty band shows statistical uncertainties only. In addition to the different event categories defined before, here the $W\gamma$ and $Z\gamma$ process are shown individually and a further subdivision depending on the presence of heavy quarks is made. This holds for all further plots on the $W\gamma$ CR in this section.

5.2. Background modelling

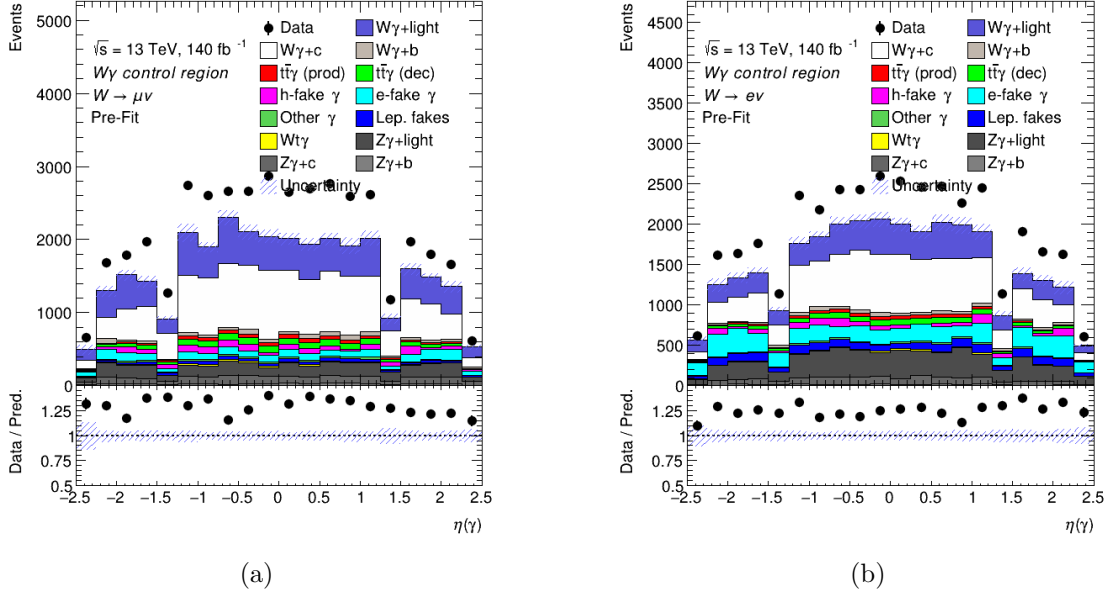


Figure 5.2.: Distributions for the $W\gamma$ CR, showing the η distributions of the photon for the μ +jets (a) and e +jets (b) subchannels. The uncertainty band shows statistical uncertainties only.

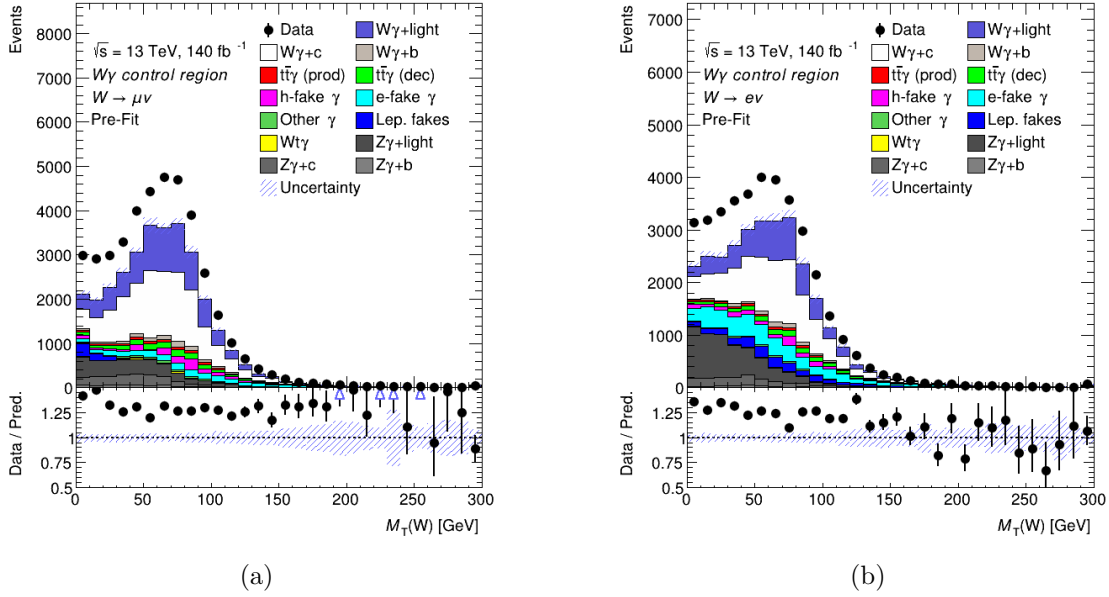


Figure 5.3.: Distributions for the $W\gamma$ CR, showing the M_T distributions of the leptonic W -boson for the μ +jets (a) and e +jets (b) subchannels. The uncertainty band shows statistical uncertainties only.

5. Analysis Strategy

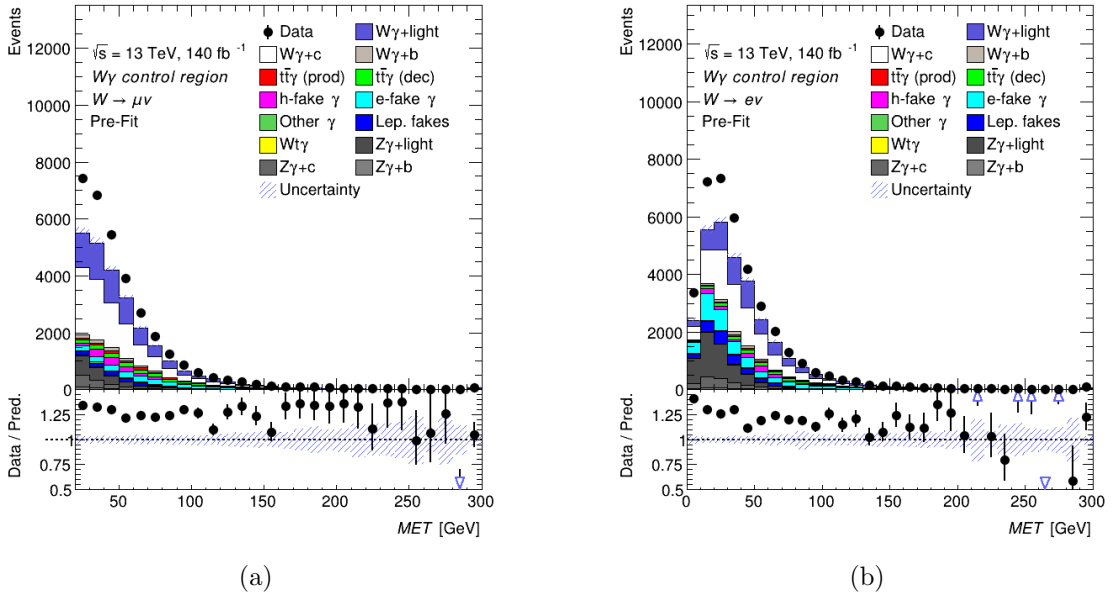


Figure 5.4.: Distributions for the $W\gamma$ CR, showing the E_T^{miss} distributions for the μ +jets (a) and e +jets (b) subchannels. The uncertainty band shows statistical uncertainties only.

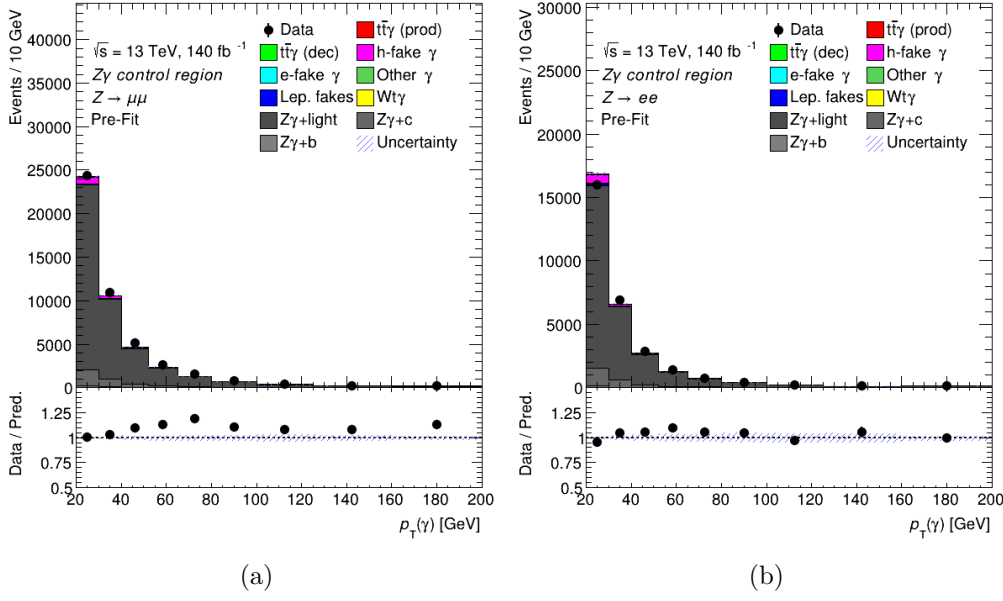


Figure 5.5.: Distributions for the $Z\gamma$ CR, showing the p_T distributions of the photon for the $\mu\mu$ (a) and ee (b) subchannels. The uncertainty band shows statistical uncertainties only. In addition to the different event categories defined before, here the $Z\gamma$ process is shown individually and a further subdivision depending on the presence of heavy quarks is made. This holds for all further plots on the $Z\gamma$ CR in this section.

5. Analysis Strategy

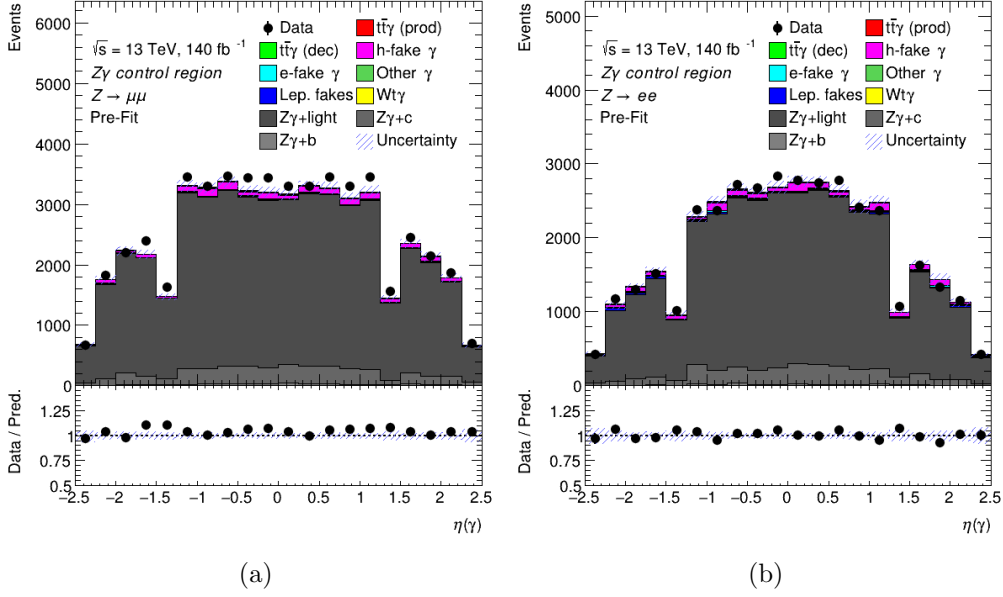


Figure 5.6.: Distributions for the $Z\gamma$ CR, showing the η distributions of the photon for the $\mu\mu$ (a) and ee (b) subchannels. The uncertainty band shows statistical uncertainties only.

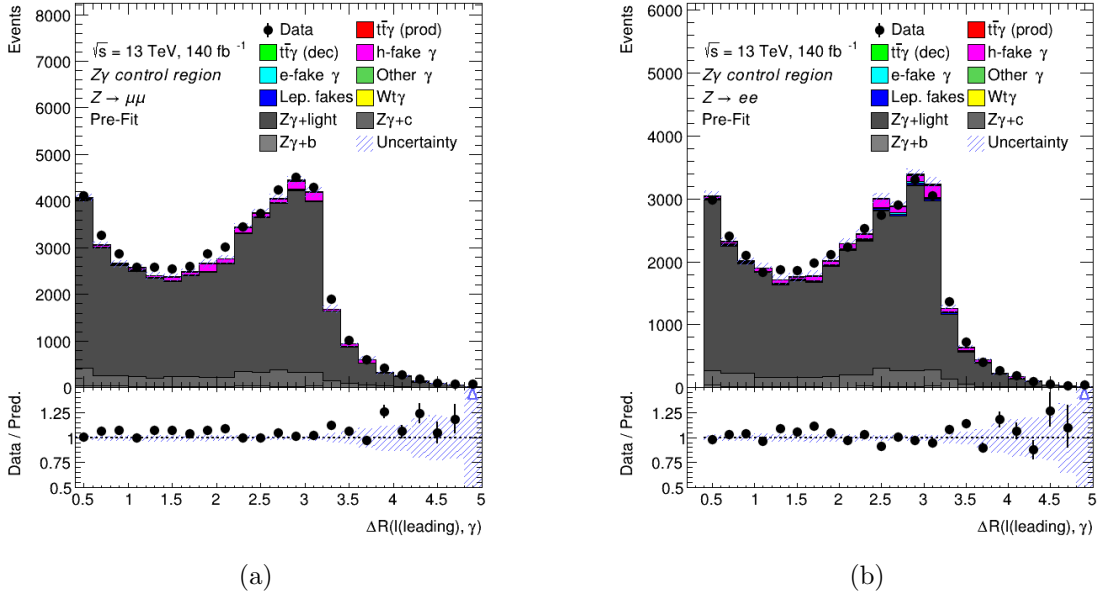


Figure 5.7.: Distributions for the $Z\gamma$ CR, showing the ΔR distributions between the photon and the first lepton for the $\mu\mu$ (a) and ee (b) subchannels. The uncertainty band shows statistical uncertainties only.

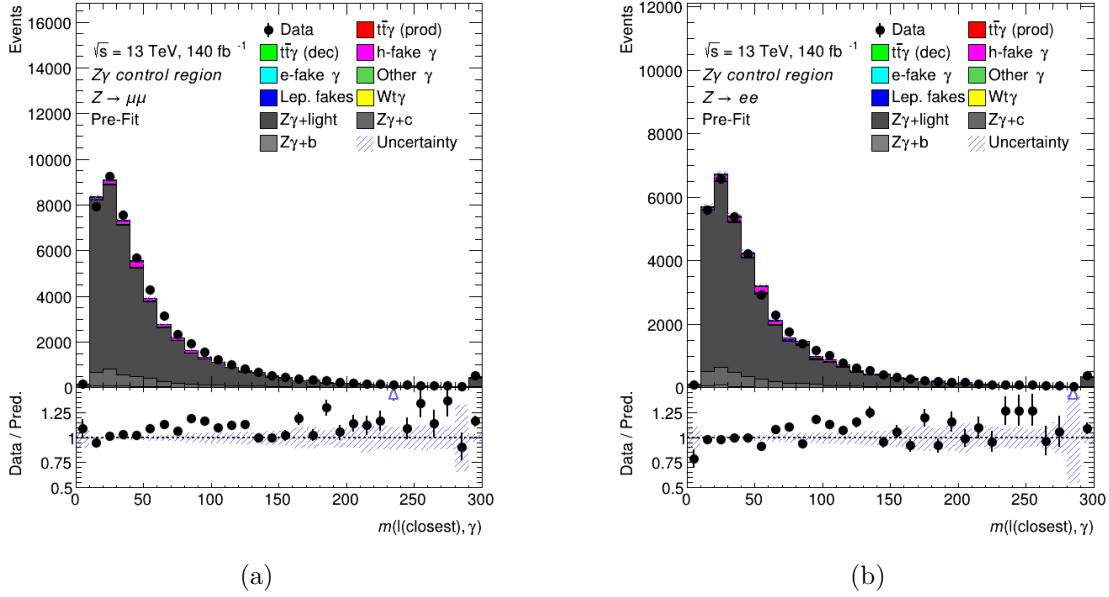


Figure 5.8.: Distributions for the $Z\gamma$ CR, showing the invariant mass distributions for the photon and the nearest lepton in the $\mu\mu$ (a) and ee (b) subchannels. The uncertainty band shows statistical uncertainties only.

5.2.2. $e \rightarrow \gamma$ fake background

Electrons that are mis-reconstructed as photons (called $e \rightarrow \gamma$ fakes) are an especially relevant background in the l +jets channel, while their contribution to the dilepton signal region is rather negligible. A large contribution of such events comes from $Z \rightarrow ee$ events. Therefore a Z -mass window selection cut is applied in the e +jets channel (see Table 4.1), which suppresses this background contribution strongly, but still, a large fraction of these events appears in the $t\bar{t}\gamma$ signal region (coming from dilepton $e^+e^- t\bar{t}$ events; see Table 4.3). In general fake contributions are badly modeled in MC simulations. Therefore a method that takes data into account is needed. For the $e \rightarrow \gamma$ fakes the so-called tag-and-probe method is used. The following sections describe in detail the definition of the control regions, the sources of $e \rightarrow \gamma$ fakes and the final extraction of the scale factors and their uncertainties, which are then applied to $e \rightarrow \gamma$ fake events in the $t\bar{t}\gamma$ signal region to correct for the rather poor modelling obtained by MC simulations. The same scale factors were also used in [146] and the same procedure was already applied in [1].

Control regions

To measure the fake rate of $e \rightarrow \gamma$ fakes, it is necessary to define a control region that is dominated by the $e \rightarrow \gamma$ fakes. This region was found by the usage of a back-to-back pair of a photon and an electron in $Z \rightarrow ee$ events. The concrete definition of this $e\gamma$

5. Analysis Strategy

Table 5.1.: Definitions of the two control regions for the tag-and-probe method to extract scale factors for $e \rightarrow \gamma$ fake events.

Object/Requirement	$e\gamma$ control region	ee control region
Tag e	> 25 GeV $ \eta < 2.47$	> 25 GeV $ \eta < 2.47$
Probe e or γ	> 20 GeV $ \eta < 2.37$	> 20 GeV $ \eta < 2.37$
Identification Isolation	Tight Tight	Tight Tight
Trigger	e must have a trigger match	e (tag) is the one with the trigger match (if both e have a trigger match tag and probe are chosen randomly)
$\Delta\phi$ (tag,probe) m (tag,probe)	150° $[m_Z - 50 \text{ GeV}, m_Z + 50 \text{ GeV}]$	150° $[m_Z - 50 \text{ GeV}, m_Z + 50 \text{ GeV}]$

control region is given in Table 5.1. A second control region (called ee control region; see also Table 5.1) is used where the photon is replaced by a second electron (with opposite charge), called the probe-electron. Comparing both control regions will give the scale factors (see calculation below).

Extraction of scale factors

The scale factors are extracted in slices of η and p_T and differentiate between unconverted and converted photons. In MC simulations, the fake rate is based on the following calculation using the number of $e \rightarrow \gamma$ fakes in $Z \rightarrow ee$ events, where a tag electron was found and a probe photon detected. For the $e\gamma$ control region, the following equation holds:

$$N_{e,\gamma}^i = N_{\text{true}}^i \times \epsilon_{\text{reco}}^e \times \epsilon_{\text{other}}^e \times p_{e \rightarrow \gamma}^i \times \epsilon_{\gamma(F)}^i, \quad (5.1)$$

where $N_{e,\gamma}^i$ is the number of observed events in bin i , N_{true}^i is the number of generated $Z \rightarrow ee$ events with the probe electron in bin i , ϵ_{reco}^e and $\epsilon_{\text{other}}^e$ are the reconstruction efficiency and other selection efficiencies of the tag electron respectively, $p_{e \rightarrow \gamma}^i$ is the probability of misidentifying an electron as a photon in bin i and $\epsilon_{\gamma(F)}^i$ is the selection efficiency of the fake photon. In the ee control region, a similar equation holds:

$$N_{ee}^i = N_{\text{true}}^i \times \epsilon_{\text{reco}}^{e_{\text{tag}}} \times \epsilon_{\text{other}}^{e_{\text{tag}}} \times \epsilon_{\text{reco}}^{e_{\text{probe},i}} \times \epsilon_{\text{other}}^{e_{\text{probe},i}}, \quad (5.2)$$

where N_{ee}^i is the number of observed events in bin i , N_{true}^i is the number of generated $Z \rightarrow ee$ events with the probe electron in bin i , ϵ_{reco}^e and $\epsilon_{\text{other}}^e$ are the reconstruction efficiency and other selection efficiencies of the tag electron respectively and $\epsilon_{\text{reco}}^{e_{\text{probe},i}}$ and $\epsilon_{\text{other}}^{e_{\text{probe},i}}$ are the reconstruction efficiency and other selection efficiencies of the probe

5.2. Background modelling

electron respectively. The fake rate FR_{MC}^i in MC-simulations in bin i is then simply given as

$$FR_{MC}^i = \frac{N_{e,\gamma}^i}{N_{ee}^i}. \quad (5.3)$$

In data the fake rate FR_{data}^i in bin i is defined as

$$FR_{data}^i = \frac{N_{e,\gamma}^{i,data} - N_{e,\gamma}^{i,non-Z} - N_{real\ \gamma}^i}{N_{ee}^{i,data} - N_{ee}^{i,non-Z}}, \quad (5.4)$$

where $N_{e,\gamma}^{i,data}$ are the number of events in the $e\gamma$ control region, $N_{e,\gamma}^{i,non-Z}$ are background events, $N_{real\ \gamma}^i$ are the events with real photons (MC-estimate), $N_{ee}^{i,data}$ are the number of events in the ee control region and $N_{ee}^{i,non-Z}$ are the background events in the ee control region. As the background is almost negligible it cannot be estimated from MC. To account for it the ee or the $e\gamma$ invariant mass spectra in the range from 60 to 120 GeV are fitted. The signal is fitted as a doubled-sided crystal ball function⁵ CB, which has a Gaussian core and a power law for the tails [266, 267]:

$$CB(x) = A \begin{cases} e^{-\frac{(x-\bar{x})^2}{2\sigma^2}} & x \in [\bar{x} - a, \bar{x} + a] \\ B(C - \frac{x-\bar{x}}{\sigma})^{-n} & x \notin [\bar{x} - a, \bar{x} + a] \end{cases} \quad (5.5)$$

where x is the variable to describe, $\bar{x}$ the mean of the Gaussian, σ its width and $a = \frac{x-\bar{x}}{\sigma}$ defines the point where the Gaussian turns into the power law. A , B and C are parameters dependent on $\bar{x}$, σ , a and n . So the function has 4 parameters to fit. The background is modelled using a Bernstein-polynomial⁶ $B(x)$ of order 4 [268]

$$B(x) = \sum_{i=0}^4 b_i \binom{4}{i} x^i (1-x)^{4-i}, \quad (5.6)$$

where x is again the variable to describe and b_i are the coefficients to fit. The subtraction of $N_{real\ \gamma}^i$ is done before the fit is done. The final scale factor SF^i in bin i , to scale simulated $e \rightarrow \gamma$ events correctly to data in the signal region of the $t\bar{t}\gamma$ analysis, is given by

$$SF^i = \frac{FR_{data}^i}{FR_{MC}^i}. \quad (5.7)$$

The binning used in η and p_T is the same for converted and unconverted photons:

- for η [0, 0.6, 1.37] and [1.52, 1.8, 2.37],
- and for p_T [25, 35, 45, 1000] (in GeV).

⁵In principle the Z mass spectrum is a Breit-Wigner function (as any resonant mass spectrum), but at the reconstruction level detector effects come into play. The crystal ball function has proven to describe the tails of the distribution quite well.

⁶A polynomial is used, because there is no known analytic function of the background (again the variety of detector effects are the reason for this), so a Taylor expansion of the unknown real function is used.

5. Analysis Strategy

Systematic uncertainties

As sources of systematic uncertainties, the following effects are taken into account:

- The shape of the signal function was changed from the double-sided crystal ball function to the MC template (using smoothing to reduce fluctuations).
- The fitting range was shortened to 65 to 115 GeV.
- Instead of a 4th-order polynomial, a Gaussian function was used for the background estimation.

These sources of systematic uncertainties are taken to be uncorrelated and the total uncertainty on the scale factors is given by adding the individual uncertainties in quadrature. As no impact of the jet multiplicity on the scale factors was observed, the scale factors are used in both the lepton+jets and the dilepton channel.

5.2.3. $h \rightarrow \gamma$ fake background

A second source of fake photons are jets or single hadrons that are misreconstructed as photons ($h \rightarrow \gamma$ fakes). The fake rate is again estimated by comparing MC predictions to data in dedicated control regions to derive scale factors⁷. For this type of background, the so-called ABCD method was chosen. For this method to work, two variables are needed that are relatively uncorrelated. In this case, these variables are photon isolation and identification criteria⁸. Both variables help to suppress photon fakes. The method uses the ratios between prompt photons and fake photons in 4 regions with complementary cuts concerning the two chosen variables. Three regions, A, B, and C, are dominated by fakes, while the fourth region, the signal region D, is dominated by prompt photons. Figure 5.9 shows a sketch to show the way the regions are defined.

Control regions

The four regions that are necessary to use the ABCD method are defined as follows:

- Region A: Photons are required to obey tight isolation criteria but fail tight identification requirements. Failing tight identification requirements means fulfilling the requirements for a loose photon identification but not the additional requirements for tight identification listed in [252]. These criteria are shower shape variables, as prompt photons have narrower showers than fake photons.
- Region B: Photons are required to fail tight identification requirements as in Region A. In addition, they also fail tight isolation criteria, again following the description in [252].

⁷The extraction of the scale factors was still done with a luminosity of 139 fb^{-1} .

⁸All other cuts and requirements listed in Section 4.4.3 for the lepton+jets channel are the same for all four regions.

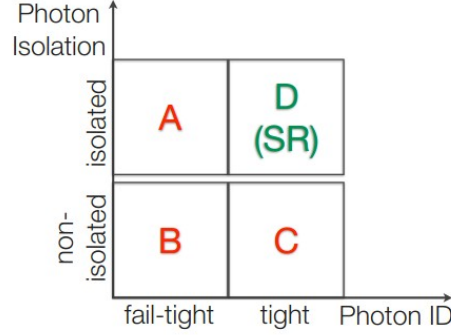


Figure 5.9.: A sketch that shows the different regions defined for the ABCD method by using different isolation and identification criteria for the photons. The sketch is taken from the internal documentation of the analysis.

- Region C: Photons are required to have tight identification requirements but loose isolation criteria [252].
- Region D: Photons are required to pass tight identification requirements and tight isolation criteria, as in the signal region [252].

In this case, Region D is the signal region of the lepton+jets channel, where such $h \rightarrow \gamma$ fakes have quite a substantial impact. Due to the very small impact in the dilepton channel, the scale factors derived here for the lepton+jets channel are also applied to the dilepton channel, assuming that the lepton and jet multiplicities have no relevant impact on the scale factor estimation.

Extraction of scale factors

The main assumption is that photon isolation criteria and identification requirements are mainly uncorrelated for $h \rightarrow \gamma$ fakes. If that is the case in the four regions, the following equation should hold for $h \rightarrow \gamma$ fake events:

$$\frac{N_A^{h \rightarrow \gamma}}{N_B^{h \rightarrow \gamma}} = \frac{N_D^{h \rightarrow \gamma}}{N_C^{h \rightarrow \gamma}}. \quad (5.8)$$

5. Analysis Strategy

Any deviations from the assumption of uncorrelated variables can be corrected by the following double ratio estimated from MC:

$$\Theta_{MC} = \frac{\frac{N_{D,MC}^{h \rightarrow \gamma}}{N_{C,MC}^{h \rightarrow \gamma}}}{\frac{N_{A,MC}^{h \rightarrow \gamma}}{N_{B,MC}^{h \rightarrow \gamma}}}. \quad (5.9)$$

This gives for the $h \rightarrow \gamma$ fake estimation in data for Region D

$$N_D^{h \rightarrow \gamma} = \frac{N_A^{h \rightarrow \gamma} \times N_C^{h \rightarrow \gamma}}{N_B^{h \rightarrow \gamma}} \times \Theta_{MC}. \quad (5.10)$$

To get the estimations of the $h \rightarrow \gamma$ fakes in each of the three control regions the background is subtracted. The scale factors for $e \rightarrow \gamma$ fakes from the previous section are taken into account. This gives:

$$N_i^{h \rightarrow \gamma} = N_i^{data} - N_i^{prompt \gamma, MC} - N_i^{e \rightarrow \gamma, MC} \times SF^{e \rightarrow \gamma}, \quad (5.11)$$

where i is either A, B or C, N_i^{data} is the number of data events in the respective region, $N_i^{prompt \gamma, MC}$ is the number of MC events with a prompt photon⁹ and $N_i^{e \rightarrow \gamma, MC} \times SF^{e \rightarrow \gamma}$ being the $e \rightarrow \gamma$ fake contribution in those three regions. Inserting the results from Equations 5.9 and 5.11 into Equation 5.10 gives the result for $N_D^{h \rightarrow \gamma}$. The scale factors are finally derived as

$$SF_{h \rightarrow \gamma} = \frac{N_D^{h \rightarrow \gamma}}{N_{D,MC}^{h \rightarrow \gamma}}. \quad (5.12)$$

The scale factors are derived in slices of $|\eta|$ and p_T of the photon for converted and unconverted photons separately with the following binning:

- for $|\eta| : [0, 0.25, 0.5, 1.37]$ and $[1.52, 2.37]$,
- and for $p_T : [0, 50, \infty]$ (in GeV).

Systematic uncertainties

The systematic uncertainties taken into account for Θ_{MC} are only $t\bar{t}$ modelling uncertainties because approximately 90% of all $h \rightarrow \gamma$ fakes are from $t\bar{t}$ events. Experimental uncertainties are not only small in comparison to the modelling uncertainties but should also cancel out in the double ratio in Equation 5.9. Therefore those uncertainties are not considered. The $t\bar{t}$ modelling uncertainties taken into account are the following ones¹⁰:

⁹From processes like $t\bar{t}\gamma$, $Wt\gamma$, $V\gamma$ etc.

¹⁰The alternative samples used to describe the modelling uncertainties are the same that are listed in Appendix A.

5.2. Background modelling

- The uncertainty on the parton shower model is estimated by comparing Pythia 8.230 (used for the nominal sample) and Herwig 7.2.1 (given by an alternative sample; see Section 4.2) both interfaced to the Powheg matrix element generator.
- Renormalisation and factorisation scale uncertainties are estimated by varying the scales independently up and down by a factor of 2. Variations are implemented in the nominal Powheg+Pythia 8 setup.
- The uncertainty on ISR due to α_s is evaluated by varying the var3c parameter in the A14 tune up and down within its uncertainty dictated by the uncertainty on α_s itself at the Z -mass pole: 0.115 and 0.140. This variation is implemented within the nominal Powheg+Pythia 8 sample.
- The effect of the scale of the first emission is modelled with the so-called h_{damp} parameter. Its variation is modelled using an alternative sample (see Section 4.2).

For equation 5.11 the following systematic uncertainties are taken into account:

- $e \rightarrow \gamma$ scale factor uncertainty
- 50% normalisation uncertainty on the prompt photon background $N_i^{\text{prompt } \gamma, MC}$

Systematic uncertainties are taken to be uncorrelated and Gaussian error propagation (including statistical uncertainties) is used to calculate the total uncertainty.

5.2.4. Lepton fake background

The last background that needed to be studied carefully is the so-called fake lepton background. This background comes from non-prompt leptons from heavy-flavour decays of b/c -hadrons or from particles like pions or kaons that are misreconstructed as leptons. In the lepton+jets channel, these events mainly come from QCD multijet events in association with a (fake) photon. Again, this is not well modelled in the MC simulations and hence data-driven estimates are necessary. For this, the matrix method as used for nominal $t\bar{t}$ production is applied, following the description in [269]. For the dilepton channel, the main source of events with a fake lepton are lepton+jets $t\bar{t}\gamma$ or $W\gamma$ events, where the second lepton is a fake lepton. This contribution is negligible (below 0.5%), as well as events with two fake leptons. As a consequence, the fake lepton background is only evaluated for the lepton+jets channel.

The matrix method

The matrix method needs two regions. For one region the tight lepton criteria (isolation and identification) as in the signal region are used (see Sections 4.3.2 and 4.3.3). For the second region, the loose criteria are applied. For electrons, this means loose identification criteria [254] and no isolation requirements. For muons, medium identification criteria

5. Analysis Strategy

[256] is used and again no isolation requirements. For the number of events N in those two regions, the following equations hold:

$$N^{\text{loose}} = N_{\text{real}}^{\text{loose}} + N_{\text{fake}}^{\text{loose}} \quad (5.13)$$

$$N^{\text{tight}} = N_{\text{real}}^{\text{tight}} + N_{\text{fake}}^{\text{tight}}. \quad (5.14)$$

One can define the probability of a fake or a real lepton to pass the tight selection criteria as ϵ_{fake} and ϵ_{real} , respectively. This changes Equation 5.14 to

$$N^{\text{tight}} = \epsilon_{\text{real}} N_{\text{real}}^{\text{loose}} + \epsilon_{\text{fake}} N_{\text{fake}}^{\text{loose}}. \quad (5.15)$$

Combining Equations 5.13, 5.14 and 5.15, gives the following matrix equation

$$\begin{pmatrix} 1 & 1 \\ \epsilon_{\text{real}} & \epsilon_{\text{fake}} \end{pmatrix} \begin{pmatrix} N_{\text{real}}^{\text{loose}} \\ N_{\text{fake}}^{\text{loose}} \end{pmatrix} = \begin{pmatrix} N^{\text{loose}} \\ N^{\text{tight}} \end{pmatrix}. \quad (5.16)$$

Assuming that $\epsilon_{\text{fake}} \neq \epsilon_{\text{real}}$ ¹¹, the matrix is invertible¹². This allows to convert Equation 5.16 into

$$\epsilon_{\text{fake}} \begin{pmatrix} N_{\text{real}}^{\text{loose}} \\ N_{\text{fake}}^{\text{loose}} \end{pmatrix} = \begin{pmatrix} \epsilon_{\text{fake}} N_{\text{real}}^{\text{loose}} \\ N^{\text{tight}} \end{pmatrix} = \frac{\epsilon_{\text{fake}}}{\epsilon_{\text{real}} - \epsilon_{\text{fake}}} \begin{pmatrix} -\epsilon_{\text{fake}} & 1 \\ \epsilon_{\text{real}} & -1 \end{pmatrix} \begin{pmatrix} N^{\text{loose}} \\ N^{\text{tight}} \end{pmatrix}. \quad (5.17)$$

From Equation 5.17 one gets directly

$$N_{\text{fake}}^{\text{tight}} = \frac{\epsilon_{\text{fake}}}{\epsilon_{\text{real}} - \epsilon_{\text{fake}}} (\epsilon_{\text{real}} N^{\text{loose}} - N^{\text{tight}}). \quad (5.18)$$

While N^{tight} and N^{loose} are known from the data, ϵ_{real} and ϵ_{fake} need to be measured (see next section).

With the knowledge of ϵ_{real} and ϵ_{fake} one can define event weights w_i using

$$w_i = \frac{\epsilon_{\text{fake}}}{\epsilon_{\text{real}} - \epsilon_{\text{fake}}} (\epsilon_{\text{real}} - \delta_i), \quad (5.19)$$

where $\delta_i = 1$ in case the event contains a lepton that passes loose and tight lepton criteria and $\delta_i = 0$ in case the event contains a lepton that only passes the loose criteria. These weights are applied to real data events that fulfil all selection criteria of the signal region (see Section 4.4.3) but loose identification and isolation criteria are required for leptons. Adding up the weights of all these events gives the lepton fake estimation in the signal region.

The values for ϵ_{real} and ϵ_{fake} depend on the lepton kinematics. Therefore the efficiencies are derived in slices of p_T and m_{T_W} ¹³(see next section).

¹¹This is justified, as tight identification and isolation cuts are defined such that they suppress fake leptons.

¹²Otherwise Equations 5.13 and 5.14 would not be linearly independent and the matrix method would not make sense.

¹³The transverse mass of the leptonic W boson.

Extraction of the fake efficiency

To estimate the fake efficiency ϵ_{fake} the following two control regions are defined using the general object definitions described in Section 4.3 but for leptons either the loose or the tight definitions are applied. The following cuts are placed:

- Exactly one loose/tight lepton.
- Two or three jets.
- At least one b -tagged jet at the 70% working point.
- $E_T^{\text{miss}} < 30$ GeV.

The fake efficiency ϵ_{fake} can then be calculated as

$$\epsilon_{\text{fake}} = \frac{N_{\text{fake}}^{\text{tight}}}{N_{\text{fake}}^{\text{loose}}} = \frac{N^{\text{tight}} - N_{\text{real}}^{\text{tight}}}{N^{\text{loose}} - N_{\text{real}}^{\text{loose}}}, \quad (5.20)$$

using the values of these two control regions. For the estimation of $N_{\text{real}}^{\text{tight}}$ and $N_{\text{real}}^{\text{loose}}$ the MC estimates for the real leptons are used¹⁴, which come mainly from $t\bar{t}$, W/Z +jets and single-top events. The values for ϵ_{fake} were estimated separately for e +jets and μ +jets events, because the underlying production mechanism is different (QCD-multijet events vs. soft muon decays from B-hadrons). The following binning of lepton- p_T and m_{W_T} was chosen:

- lepton p_T : [27,35,45,62,1000] GeV for e +jets and [27,35,45,52,60,1000] GeV for μ +jets
- m_{T_W} : [0,7,15,25,40,1000] GeV

Extraction of the real efficiency

The real efficiency ϵ_{real} is estimated from MC predictions as

$$\epsilon_{\text{real}} = \frac{N_{\text{real}}^{\text{tight}}}{N_{\text{real}}^{\text{loose}}}. \quad (5.21)$$

In this case, the same two control regions already used for the fake efficiency were used. The same binning for lepton p_T and m_{T_W} as for ϵ_{fake} is used.

¹⁴To define a prompt lepton, truth matching between reconstruction-level and particle-level is done.

5. Analysis Strategy

Systematic uncertainties

Several sources of systematic uncertainties are taken into account.

The extraction of ϵ_{real} was repeated in a region where almost no fake leptons are present by choosing $E_T^{\text{miss}} > 120$ GeV. The difference to the nominal extraction is around 1% and therefore negligible compared to the other sources of systematic uncertainties taken into account.

For ϵ_{fake} the $t\bar{t}$ normalisation uncertainty, which is about 5%, is taken as a source of systematic uncertainty. The size of this uncertainty was estimated by reducing the $t\bar{t}$ normalisation by 5% and reextracting ϵ_{fake} . The difference between the nominal values and this extraction of ϵ_{fake} gives the size of the uncertainty, which is symmetrised. Both efficiencies ϵ_{fake} and ϵ_{real} were also extracted in two other control regions. One region only requires at least one jet instead of limiting the control region to events with two or three jets. The other control regions use a different E_T^{miss} range from 30 to 50 GeV. Again the differences are taken as uncertainties.

All these sources of systematic uncertainties are taken into account as one cumulative up/down variation of the event weight as given in Equation 5.19. For this, the impact of the different sources is taken as uncorrelated and added up in quadrature.

Two further systematic variations are taken into account independently from each other. These are two different choices of the parametrisation. While the nominal parametrisation uses slices of lepton p_T and m_{W_T} with the binning given above, two other parametrisations were also used. One uses the lepton p_T and lepton $|\eta|$ in the following binning:

- lepton p_T : [27,35,45,62,1000] GeV for e +jets and [27,35,45,52,60,1000] GeV for μ +jets
- lepton- $|\eta|$: [0,0.5,1.0,1.5,2.0,2.5]

The other parametrisation uses lepton- p_T and $\Delta\phi(l, E_T^{\text{miss}})$ in the following binning:

- lepton p_T : [27,35,45,62,1000] GeV for e +jets and [27,35,45,52,60,1000] GeV for μ +jets
- $\Delta\phi(l, E_T^{\text{miss}})$: [0,0.5,1,1.5,2,3.2]

Again in both cases, the differences in the nominal parameter choice are used as an up/down variation.

5.3. Neural network

One of the main challenges and one of the main tasks of the author of this thesis was to discriminate, within the $t\bar{t}\gamma$ process, those events where the photon was emitted during $t\bar{t}$ production and where the photon was emitted by $t\bar{t}$ decay products. For more details on what production and decay mean see the corresponding descriptions of the MC-samples in Section 4.2.1. As outlined in Section 2.6.3 this separation is based on the narrow width approximation. Due to the similarities of the kinematics, such a

separation needs the usage of advanced methods based on machine learning. For this purpose, it was decided to use artificial neural networks (ANNs). To perform the neural network training the computational framework from the recently published $t\bar{t}Z$ analysis by ATLAS [251] was used. The framework uses the Python library Keras [270] as a front-end and Tensorflow [271] as its back-end. The idea of ANNs and a short introduction are given in the next section, while the concrete application of such networks for this analysis is outlined in the subsequent sections. Those sections are based again on the description of the analysis in the internal ATLAS documentation.

5.3.1. Introduction to neural networks

Under the term of ANNs, a lot of different machine learning methods are summarised. Machine learning itself is an area of computer science, where different sophisticated algorithms are used to learn general features from various forms of input, for example, pattern recognition. The learned information can then be used to analyse so far unseen input and draw conclusions on it. ANNs are graphs, where each node and each connection correspond to a mathematical operation. The output of these operations depends on weights. These weights are optimised with respect to a certain measure, called loss-function. More details on the weights and the loss function follow below. One of the simplest ANN is a so-called feed-forward neural network (FNN) used for classification tasks. Classification means that the trained FNN gives an output value based on several input features. Depending on that value the input is classified into predefined categories, also called classes. An FNN itself contains the input layer, where all the features are given to the network, several hidden layers and an output layer. A schematic view of such a network is shown in Figure 5.10. The information is passed on from the input layer through the hidden layers up to the output layer. Each layer of the network has a certain number of nodes. In the case of FNNs, these nodes are fully connected to the previous layer and the subsequent layer, which means one node receives information from all nodes from the previous layer and passes the output to all nodes of the subsequent layer¹⁶.

Mathematically the following happens: the vector $\mathbf{x}$, which contains the information of the previous layer is given a weight w_i for each vector entry x_i such that the multidimensional input is mapped to the space of real numbers as

$$f(\mathbf{x}) = z = \sum_{i=0}^N w_i x_i + C, \quad (5.22)$$

where N is the dimension of the vector and C is a constant offset, also called bias term. These linear models were not able to learn the exclusive OR¹⁷ [272]. The solution are

¹⁵The website was accessed on November 8th 2023.

¹⁶In graph theory this is called a complete multipartite graph.

¹⁷Exactly one out of two arguments equals a certain condition.

5. Analysis Strategy

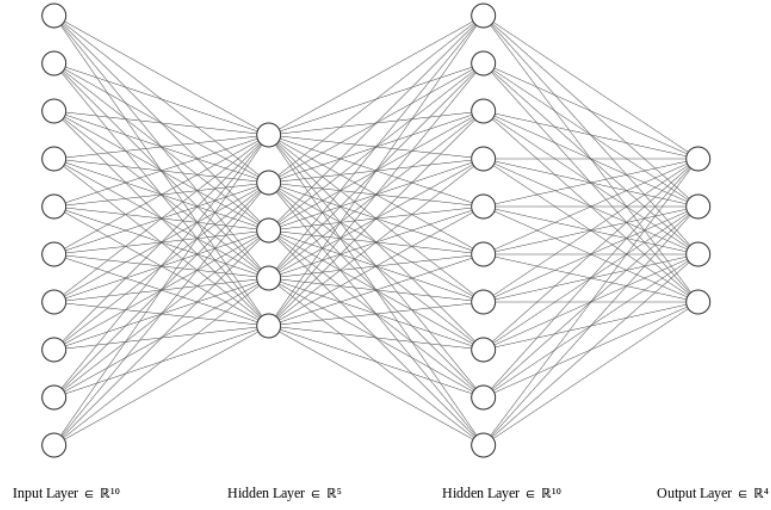


Figure 5.10.: An example of a fully connected network with 10 input nodes, a first hidden layer with 5 nodes, a second hidden layer with 10 nodes and an output layer with 4 nodes. The graph was created using <https://alexlenail.me/NN-SVG/index.html>¹⁵.

non-linear activation functions¹⁸ [273]. The output $\hat{x}$ of a node becomes

$$\hat{x} = h(f(\mathbf{x})) = h(z). \quad (5.23)$$

There are different options for these activation functions, like $\tanh(z)$, sigmoid or the rectifier linear unit (ReLU). For the output layer, the softmax function is a very common choice, as it allows to interpret the output y as a Bayesian conditional probability of the input vector $\mathbf{x}$ to belong to a certain class $\mathcal{C}$, $y = P(\mathcal{C}|\mathbf{x})$. To optimise the weights for each node backpropagation [274] is used to calculate the gradient of the loss function. The negative gradient gives the direction in the hyperspace of weights to minimise the loss function $L(\mathbf{w})$. A common choice for binary classification is the so-called binary cross-entropy, while for multiclass classification a generalised form, the categorical cross-entropy CCE, is used. The lower its output value, the better the classification. The change of weights in the weight vector $\mathbf{w}$ after each recalculation of the gradient, is given as

$$\mathbf{w}_{\text{new}} = \mathbf{w}_{\text{old}} + \eta \nabla L(\mathbf{w}_{\text{old}}). \quad (5.24)$$

The value η is called the learning rate. If η is not constant but depends on the differences in the loss function between two iterations, one has an adaptive learning rate, which can help to find a good balance between fast but maybe unstable learning (when η is chosen to be large) and slow but computationally expensive learning, which is more accurate

¹⁸Such nodes are called perceptrons and FNNs are therefore also called multilayer perceptrons (MLPs).

(when η is chosen to be small). The adaptive algorithm used in this thesis is the Nadam-algorithm [275–278].

The training process is done in the following way. The events are sorted into three statistically independent subsets, used for training, validation and testing. How this is done exactly is described in the next section. The training set is further divided into smaller batches. After each batch was processed by the network the weight optimisation happens. To avoid overtraining, the optimised weights are checked with the validation set. Only if the loss function improves for the validation set too, the improvement is valid. In the end, the best set of weights are those that give the minimal loss function for the validation set. The better the training the larger is also the area under the curve (AUC) of the receiver-operator-characteristic curve (ROC curve) (for more details see next section). To have a further test that the trained model can be generalized, a comparison between the training and the testing set is done. The more similar they are (ROC curves, classifier output distributions), the better the model generalises. In this analysis, K -fold cross-validation is used with $K = 5$ folds. The input is divided into K subsets. One subset is used for testing while $K - 1$ subsets are used for training and validation. The proportion of training and validation can still be varied and is not fixed a priori. This method allows to increase the training data by the cost of statistical independence. To minimise the dependence on a specific node and its connections, randomly a few nodes can be removed from the training. This technique is called dropout [279]. As from one epoch to the next, other nodes and their connections are removed, the network gets stable against being dependent on a single node. This helps to prevent overtraining. A second technique to mitigate overtraining is batch normalisation. This technique normalises the distributions before those are propagated to the next layer.

5.3.2. Event classification in the lepton+jets channel

Before going into detail about the variable selection, the neural network training and the usage of the the classification results, a short section about reconstructing top quarks is given, as the resulting variables are important for the neural network training.

Top quark reconstruction

From the author’s master thesis [280] it was already known on parton level that reconstructing the top quarks’ invariant mass might be very helpful to distinguish $t\bar{t}\gamma$ from production and $t\bar{t}\gamma$ from decay by the usage of neural networks [280]. As this analysis tries to apply a similar approach at the reconstruction level to measure cross-sections, a reconstruction of the top quarks was useful. At the parton level, all decay products of the top quark are present and it is known which jets are the b -jets. Hence, there is no ambiguity in the top quark reconstruction. At the reconstruction level, the correct assignment of the different objects to their corresponding top quark is no longer known. Therefore the following, rather simple algorithm was applied in the lepton+jets channel: 1. The two b -jets were identified by ordering all jets concerning the pseudo-continuous b -tagging score. As jets might end up in the same bin of the pseudo-continuous b -tagging

5. Analysis Strategy

distribution [262] a second criterion was needed to distinguish between them. For this purpose, the p_T of the jet was used by taking the jet with the highest p_T .

2. From all other jets the permutation of those two jets giving the invariant mass closest to the W -pole mass was used as the reconstructed W -boson, which decays hadronically.

3. To reconstruct the leptonically decaying W -boson, first the neutrino was reconstructed. Imposing that the W -boson was produced at its pole mass (80.4 GeV [5]), one can derive a quadratic equation (see Appendix B.1) for the unknown neutrino $p_{z,\nu}$. If the measurements of the detector were exact, this equation would give only one solution. Due to resolution effects, this is not always the case. In the case of two solutions, the more physical one is taken, which was in general the solution with the smaller absolute value. In case the equation has no real solution, the discriminant of the quadratic equation was set to 0. Under the assumptions that the E_T^{miss} is completely the result of the non-detected neutrino and a negligible neutrino mass, one can use the measured lepton four-vector, E_T^{miss} , $\phi(E_T^{miss})$ and the calculated neutrino $p_{z,\nu}$ to reconstruct the W -boson.

4. The next step is to combine the two jets defined as being the b -jets with the hadronic and leptonic W -bosons. This opens the possibility for 2 permutations. Including the photon gives another two permutations, so one ends up with 4 permutations:

- $b_1 + W_{had}$ and $b_2 + W_{lep} + \gamma$
- $b_2 + W_{had}$ and $b_1 + W_{lep} + \gamma$
- $b_1 + W_{lep}$ and $b_2 + W_{had} + \gamma$
- $b_2 + W_{lep}$ and $b_1 + W_{had} + \gamma$

To order these permutations the following quantity was calculated for each permutation:

$$\Delta m^2 = (m_1 - m_{top})^2 + (m_2 - m_{top})^2, \quad (5.25)$$

where m_1 and m_2 are the invariant masses of the two reconstructed top quarks and m_{top} the top quark pole mass, which was taken to be 172.5 GeV (the standard value used for MC simulations [5]). As described in Chapter 7 the following variables were used in the neural network: $m_1 + m_2$ from all 4 permutations and Δm^2 from all 4 permutations. The validity of this approach was checked by comparing normalised distributions at the reconstruction level and the particle level. These are shown in Appendix B.2.

Variable selection

One of the main goals of this analysis is to separate between the contributions from $t\bar{t}\gamma$ from production and $t\bar{t}\gamma$ from decay. In the lepton+jets channel (in most Figures this is abbreviated with l+jets) contributions from other backgrounds than $t\bar{t}\gamma$ from decay are significant. One can group them into two major classes: the prompt photon background and the fake photon ($e/h \rightarrow \gamma$) background. Lepton-fake events were not considered in the training, because there are too few of them. Hence, it is reasonable to

define four classes for a multiclass classification. These classes are: $t\bar{t}\gamma$ from production, $t\bar{t}\gamma$ from decay, photon fakes and other photons. To be able to distinguish the four classes, it is mandatory to have enough variables, which provide a reasonable capability to statistically separate the input classes. The separation power S , which quantifies this capability, can be defined as

$$S = \frac{1}{2} \sum_{i \in \text{bins}} \frac{(s_i - b_i)^2}{s_i + b_i}, \quad (5.26)$$

where the sum runs over all bins and s and b are the signal and background histograms, respectively¹⁹. Because S depends highly on the binning, only with a suitable choice of it, S gives a realistic impression of the real separation power. Besides the good separation power of the input variables, the simulation data should also give a good description of real data. This is checked by the data-MC agreement of the different input variables. In Figure 5.11a one can see the distribution for the photon p_T as an example, while Figure 5.11b shows the separation between different categories. In general, the data-MC agreement is good. Calculated χ^2 probabilities, including uncertainties before the fit, range from slightly below 10% to around 90%. There is an offset of about 10% between data and prediction by MC simulations. However the shape description by MC is good, as the ratio between data and MC events is mostly flat. Only in a few cases, one sees shape differences in the tails of the distributions. The full list of plots for all variables can be found in Appendix D.1. In the following all used variables are listed in the order of their separation power between $t\bar{t}\gamma$ from production and $t\bar{t}\gamma$ from decay:

- Invariant mass of photon and lepton
- Photon p_T
- Invariant mass of photon and the jet with the highest PCBT-score (PCBT; see Section 4.3.4); if two b -jets are in the same PCBT-bin they are ordered with respect to their p_T
- ΔR between photon and lepton
- Invariant mass of reconstructed top quark and anti-top-quark (all four permutations; see previous section)
- Photon E
- Δm of reconstructed top quark and anti-top-quark (all four permutations; see Equation 5.25 in previous section)
- Invariant mass of all jets, the lepton and the photon
- H_T (scalar sum of all jet p_T)

¹⁹Histograms are normalized to unity.

5. Analysis Strategy

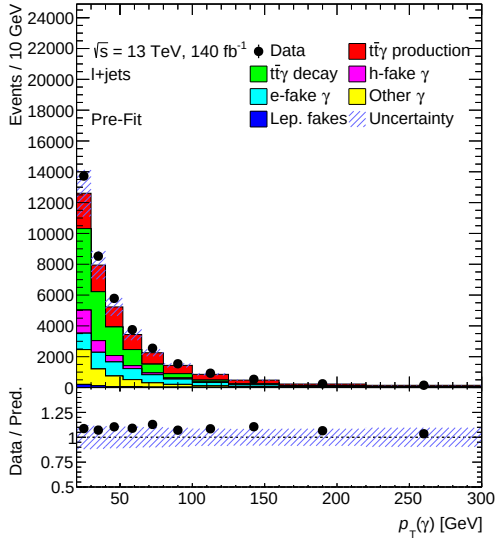
- p_T of leptonic W boson (reconstructed from E_T^{miss} and lepton p_T)
- p_T and E of the first four jets (p_T ordered)
- ΔR between photon and closest b -jet
- E_T^{miss}
- Invariant mass of lepton and closest b -jet
- Transverse mass of leptonic W boson
- Number of jets
- ΔR between lepton and closest b -jet
- Difference between reconstructed W boson mass (leptonic and hadronic) and the W boson pole mass
- Photon η
- B-tagging output for the first four jets (PCBT-ordered)
- Photon conversion type (converted or unconverted)
- Number of b -jets (tagged with 70% efficiency)
- Photon ϕ

There are different reasons to use these variables. Those related to the top quark reconstruction and the photon kinematics are mainly used to separate between $t\bar{t}\gamma$ from production and $t\bar{t}\gamma$ from decay, due to the differences in photon kinematics and the photon's presence in the decay or the production. The other variables are related to the full event kinematics and help to separate the other two background classes. The large number of variables is justified by the complexity of the classification task. In general variables with a high separation power should have higher priority to be used. But also the correlations among variables are very important. The neural network is expected to be able to learn correlations in a high-dimensional space, which unfortunately cannot be visualised. So even variables with a low separation power can improve the performance of the network. Especially for the separation of photon fakes it turned out that all variables are important (see Figure 5.20 in the next Section).

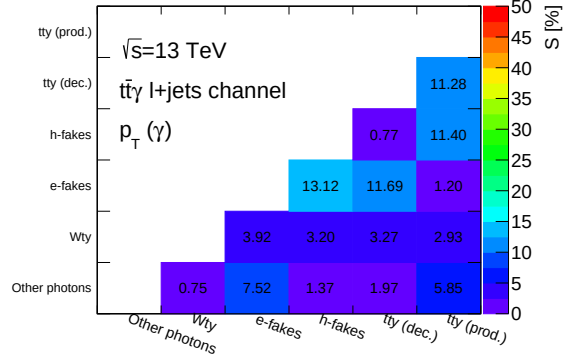
Training, optimisation and results

The initial neural network training was done with the following setup:

- four hidden layers with 50 neurons each
- one dropout layer following the first hidden layer



(a) Distribution for the p_T distribution of the photon in the lepton+jets channel with full pre-fit uncertainties (statistical and systematic). $Wt\gamma$ -events are subsumed in the Other γ category.



(b) Separation power for the p_T of the photon in the lepton+jets channel for different categories with respect to each other. In the neural network training $e \rightarrow \gamma$ fakes and $h \rightarrow \gamma$ fakes are grouped together as the “photon fake” class and $Wt\gamma$ is included in the “other photon” class.

Figure 5.11.: Distribution (left) and separation power (right) for the photon p_T .

5. Analysis Strategy

- no batch normalisation layers
- learning rate: 0.01
- activation function: rectifier linear unit
- patience: 100 epochs
- epochs: 5000
- batch size: 100000
- validation size: 5% of the fold used for training
- dropout probability: 0.2
- output activation: softmax
- metrics: accuracy
- loss: categorical cross-entropy

The MC events used in the training were weighted with the same weights used to produce the variable distributions, which means to take different (object) scale factors into account, as well as the general MC event weight, the cross section and the luminosity²⁰. The reason was to take the events into account in the same way how they contribute to the theoretical expectation. One problem is the treatment of negative weights coming from events generated at next-to-leading order. These cannot be taken into account by ANNs, but they have a physical meaning because of the NLO calculations they are based on. It was decided to drop them and adjust the positive weights such that the overall sum of weights, which gives the total MC prediction, stays the same. Other possibilities were discarded, like dropping those events fully or taking the absolute value of the weights, as this would have been even more unphysical. This approach changed the shape of some distributions slightly but not to an extent that a variable would not have been usable in the training anymore. Such a comparison²¹ is shown in Figure 5.12 for the photon p_T . It shows the nominal distributions for signal ($t\bar{t}\gamma$ from production) and background (everything else). Besides this, it shows how the distributions would change if one would only take events with positive weights into account and how the distributions look after the correction of the normalisation. The ratios of these distributions with respect to the nominal distributions are shown in the ratio panel below. As the ratios of the corrected distributions are close to unity and deviations are typically smaller than uncertainties, the described approach is justified.

Before the training the variables were all transformed into the interval $[0,1]$. So the highest value of a certain variable gets the value 1, while the lowest value is set to 0.

²⁰The event weighting for the neural network training was done still taking the luminosity to be 139 fb^{-1} .

²¹Further variables are shown in Appendix D.1.

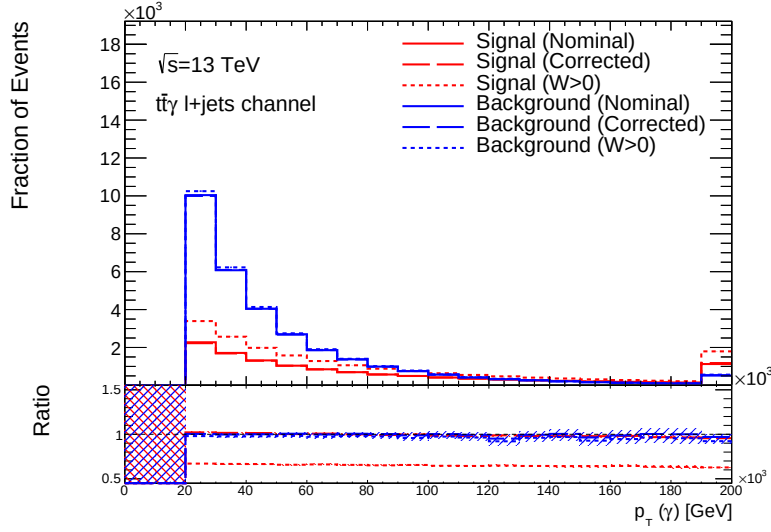


Figure 5.12.: Comparison of distributions for the p_T of the photon in the lepton+jets channel to check the impact of negative weights on signal ($t\bar{t}\gamma$ from production) and background (all other processes). The uncertainties are statistical.

All other values are transformed accordingly. This procedure makes sure that variables with larger values (e.g. given by the chosen unit) do not become more important than other variables with lower values. The training itself is then done by applying a K-fold cross-validation with 5 folds. The MC events are divided into the different folds based on their event number n , using $n \bmod 5$. Data events are divided into five sets in the same way to decide later which of the five trained networks will be applied to which data event.

In general, the results gave reasonable outputs nevertheless a few hyperparameters were checked to find the optimal setup. These were the following parameters:

- hidden layers
- neurons per layer
- dropout layers
- batch normalisation layers

Therefore a grid search was made. This means that the values for these parameters were chosen randomly in a given range. The number of hidden layers varied between one and six, while the number of nodes per layer was varied for each layer independently between 10 and 90. The random assignment of the dropout layer behind a hidden layer or the input layer results in zero to five of such layers. The same was done with batch normalisation layers. In total 100 different neural network architectures were tested. The other

5. Analysis Strategy

parameters listed above were not changed, either because the given choice is known to be rather optimal or because the problems (convergence for example) that one could mitigate with these variables have not shown up here. It was also tested to use fewer and more classes, but four classes showed again a good balance between complexity and necessary information reduction.

From the output results of these tests, it was then decided to use the original network architecture. The reason for this was that it showed a relatively good trade-off for the different classes in the confusion matrix and the ROC curves (see below). All the configurations tested to optimise the hyperparameters gave similar results or the improvement in one class came with a worse classification of other classes. Therefore, the results of the original network architecture were eventually chosen to be applied to define the signal and control regions for the final fit to extract the inclusive cross sections. The quality of the training procedure itself can be controlled by comparing the evolution of the training loss and the validation loss (see Figures 5.13 to 5.15), as well as the ROC curves (see Figures 5.16 and 5.17).

The evolution of the loss functions is shown for each fold. As can be expected the loss function of the training set decreases faster at the beginning of the training, while the validation loss needs around 10 to 20 epochs to decrease to the same or nearly the same value. For a few tens of epochs, both, the validation loss and the training loss, decrease in a similar way but with larger fluctuations for the validation loss. This is because the optimisation is done on the training set, while the validation set controls the optimisation to prevent overtraining. From the loss curves, it is seen that at some point the validation loss reaches a plateau, while the training loss still decreases. This is the point where overtraining starts. The overtraining is prevented because in the end the set of weights for the neural network that is chosen is the one giving the smallest validation loss. After 100 epochs without improving the validation loss, the training stops automatically (early stopping) and does not carry on till the full 5000 epochs are executed. The ROC curves do show almost no sign of overtraining. Only for the “other photon” classifier a slight overtraining is visible (see Figure 5.17). As the main focus of the classification is to separate $t\bar{t}\gamma$ from production from $t\bar{t}\gamma$ from decay and it is not too large, this was taken to be acceptable. ROC-curves do compare the dependence of the true positive rate as a function of the false positive rate (for the definitions see below). For this, they use the final output distributions for the training set and the testing set. For each of the four output nodes, there is one output distribution. The higher the output value for a certain event the more signal-like the event. In this particular case, the output values are between 0 and 1 and can be interpreted as a Bayesian probability of how signal-like the event is. For each output node, one of the four defined classes is considered signal. The rest is considered background. This form of multiclass classification is called one versus all.

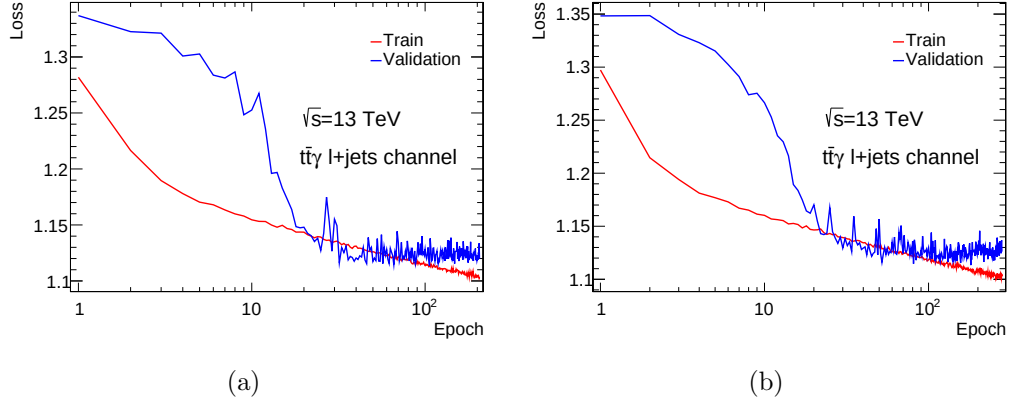


Figure 5.13.: The evolution of the loss function for the first (a) and second fold (b) for training and validation samples.

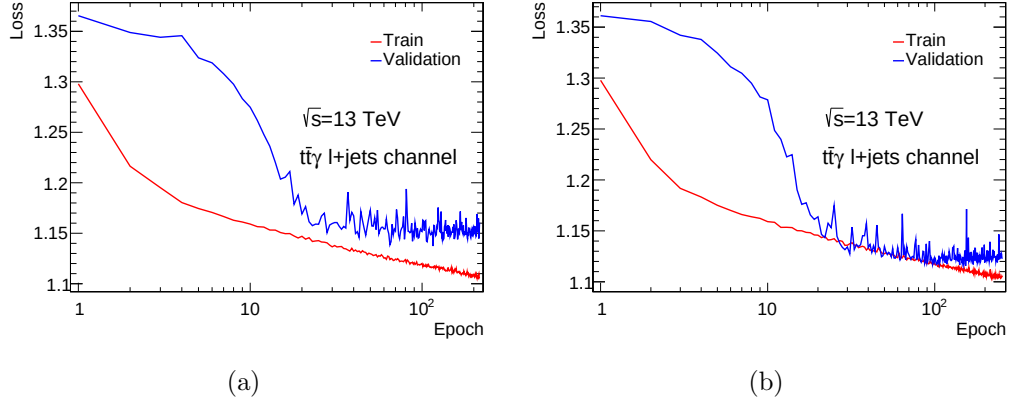


Figure 5.14.: The evolution of the loss function for the third (a) and fourth fold (b) for training and validation samples.

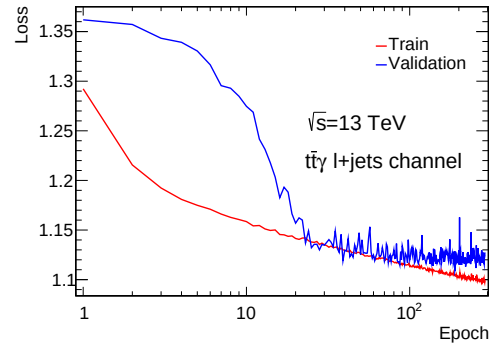


Figure 5.15.: The evolution of the loss function for the fifth fold for training and validation samples.

5. Analysis Strategy

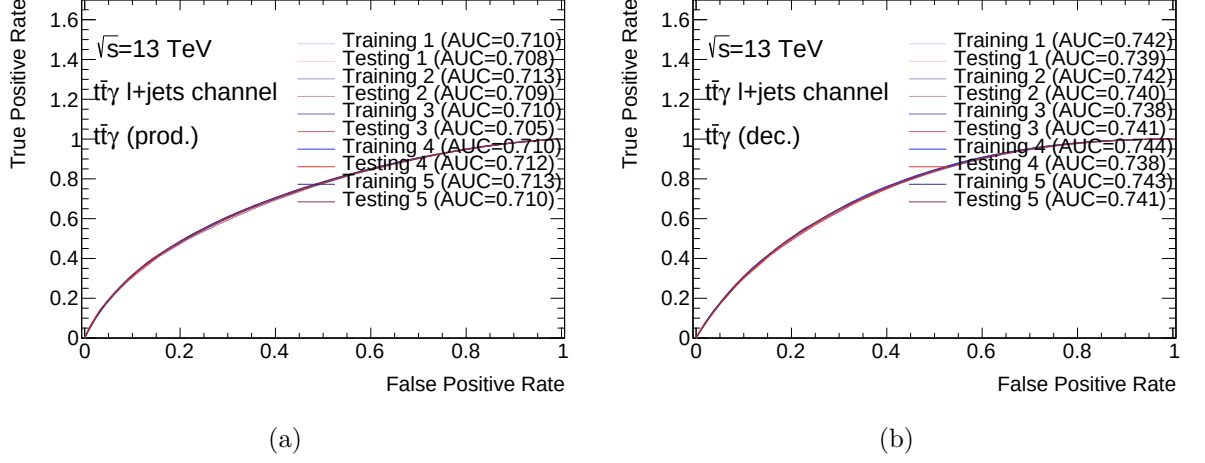


Figure 5.16.: The ROC curves (and AUC values) for all 5 folds for training and testing for (a) the $t\bar{t}\gamma$ (prod.) and (b) the $t\bar{t}\gamma$ (dec.) classifier.

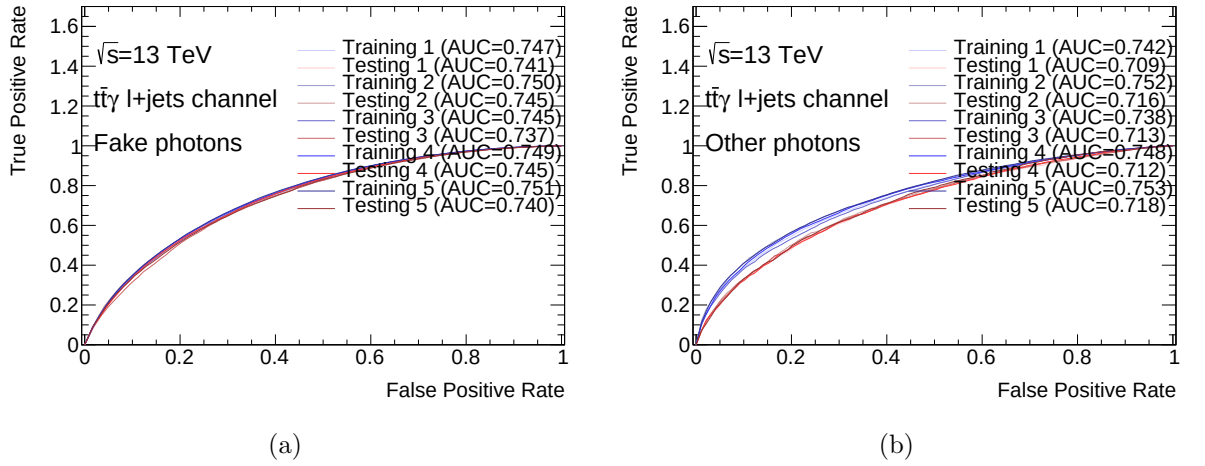


Figure 5.17.: The ROC curves (and AUC values) for all 5 folds for training and testing for (a) the “fakes” and (b) the “other photon” classifier.

The output distributions for the testing and the training sets for fold 1 for each output node are shown in Figures 5.18 and 5.19. The distributions for the other folds are in Appendix D.1. These distributions check the agreement of the neural network output by comparing training and testing sets. As described before, the neural network should not be sensitive to statistical fluctuations of the training set but should generalise well to the testing set. From the ratio panels, it can be seen, that mostly there is an agreement within statistical uncertainties. Only for the “other photon” classifier a slight mismatch is visible showing the small tendency of overtraining for this class. From these distributions, the ROC curves are calculated. For a given output value the true positive rate is calculated as the fraction of all signal events being above that given output value. The false positive rate is calculated similarly, but using the background events. The more background and signal are separated by the output distribution, the steeper the ROC curve rises for small values of the false positive rate. The area under the curve (AUC) of a ROC curve gives a measure of how well the classification works. If signal and background are fully separable it becomes 1. If the neural network cannot differentiate between signal and background it becomes 0.5. In this case, the neural network is not better than tossing a coin to tell if an event is either signal or background. If there is no overtraining, the ROC curves for the training set and the testing set should look similar and AUC values should be nearly the same, although it can be expected that the value is slightly higher for the training set, because it was used to optimise the weights. All ROC curves for all folds and all output nodes are shown in Figures 5.16 and 5.17. In general, there is no overtraining seen but slightly for the “other photon” classifier, reflecting the slight mismatch between training and testing curves described before. This was considered to be acceptable, as the main focus was to separate $t\bar{t}\gamma$ from production from $t\bar{t}\gamma$ from decay.

The importance of the variables was checked using the procedure of permutation importance. This technique shuffles the input values of one input variable among the different events. This breaks the expected correlations among the different variables and results in smaller AUC values. The larger the difference between the nominal AUC and the AUC obtained with permutation importance, the more important the respective variable is. As indicated before, especially for the “fakes” classifier the variables are mostly of similar importance (see Figure 5.20). The other plots showing the permutation importance are shown in Appendix D.1.

Due to the usage of K-fold cross-validation with five folds, five neural networks have been trained. They are combined such that each event is passed through the network for which it was in the testing set. In that way, the final output distributions are derived. For the actual data, the procedure is the same. The final output distributions are shown in Figures 5.21 and 5.22 and show clearly that the neural network can classify the events quite correctly. This can be shown too by the confusion matrix. To derive the confusion matrix, each event is classified according to the highest output value of the four output nodes. The confusion matrix is shown in Figure 5.23. The fraction of correctly classified events by this method varies between 45% to 60%.

5. Analysis Strategy

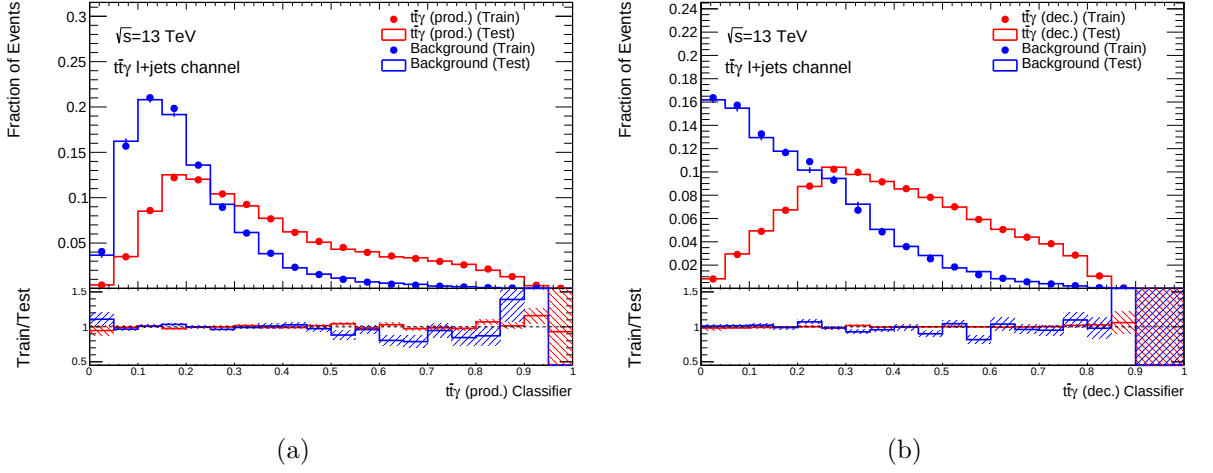


Figure 5.18.: Output distributions for signal ((a) $t\bar{t}\gamma$ (prod.) and (b) $t\bar{t}\gamma$ (dec.)) and background (everything not considered as signal) for (a) the $t\bar{t}\gamma$ (prod.) classifier and (b) the $t\bar{t}\gamma$ (dec.) classifier for fold 1. The uncertainties are statistical uncertainties.

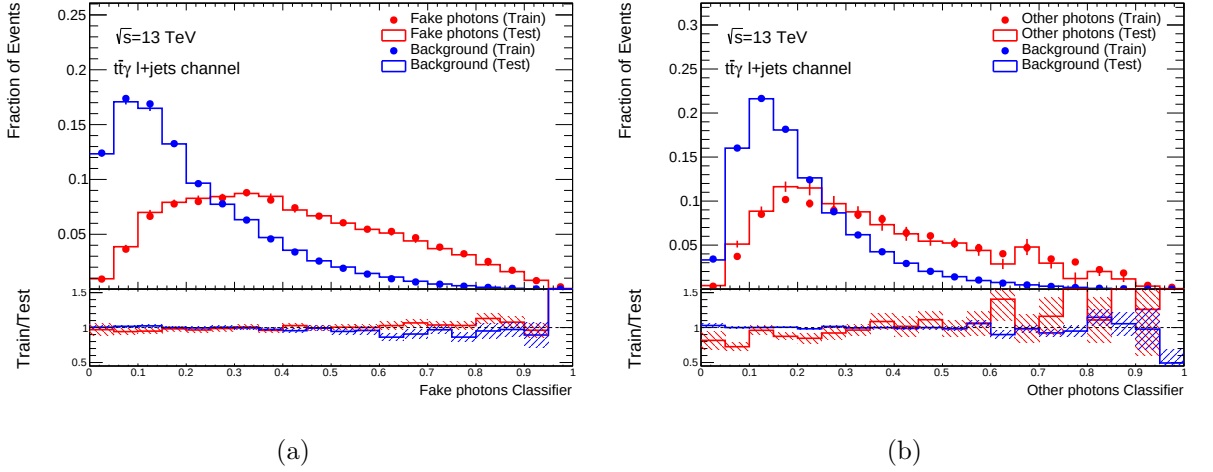


Figure 5.19.: Output distributions for signal ((a) fake photons and (b) other photons) and background (everything not considered as signal) for (a) the “fake photons” classifier and (b) the “other photon” classifier for fold 1. The uncertainties are statistical uncertainties.

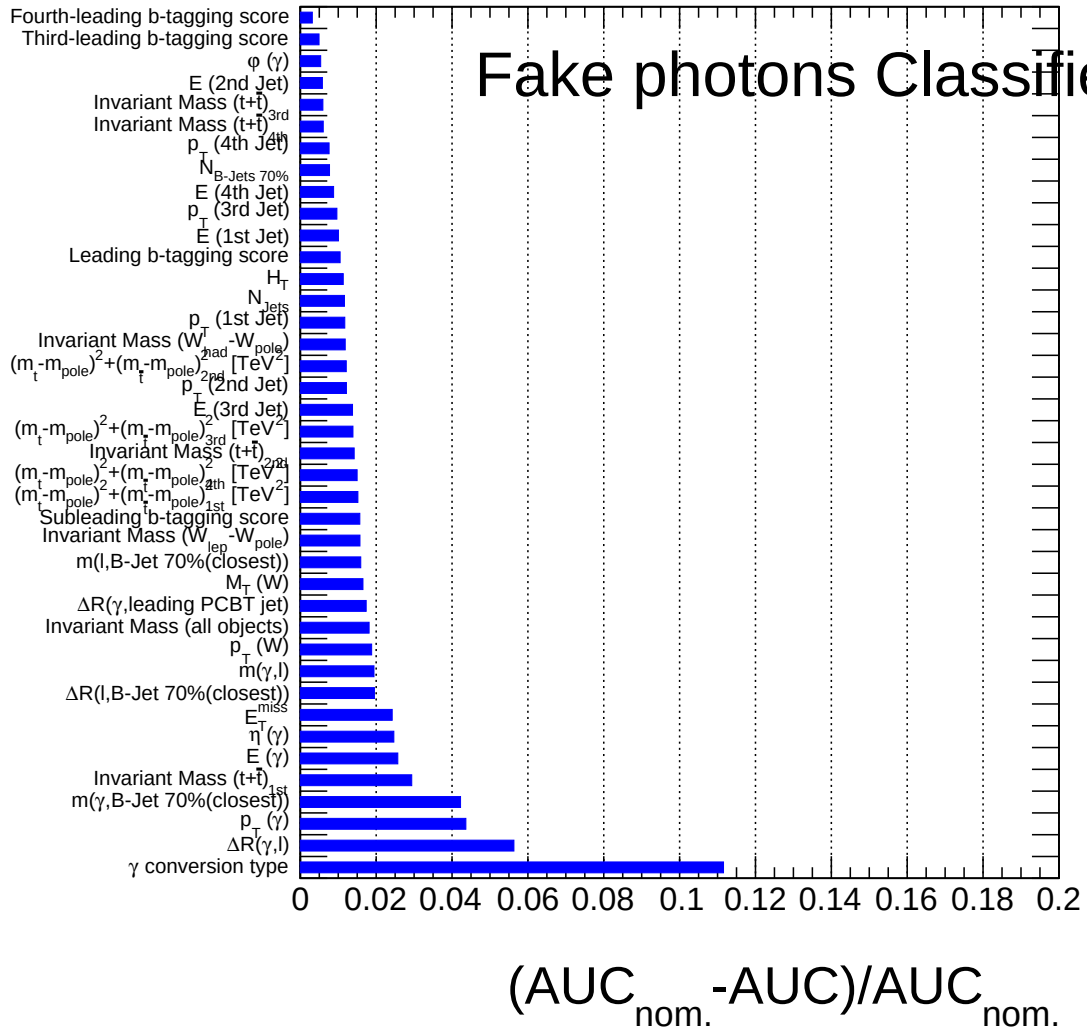


Figure 5.20.: Permutation importance for the “fake photons” classifier.

5. Analysis Strategy

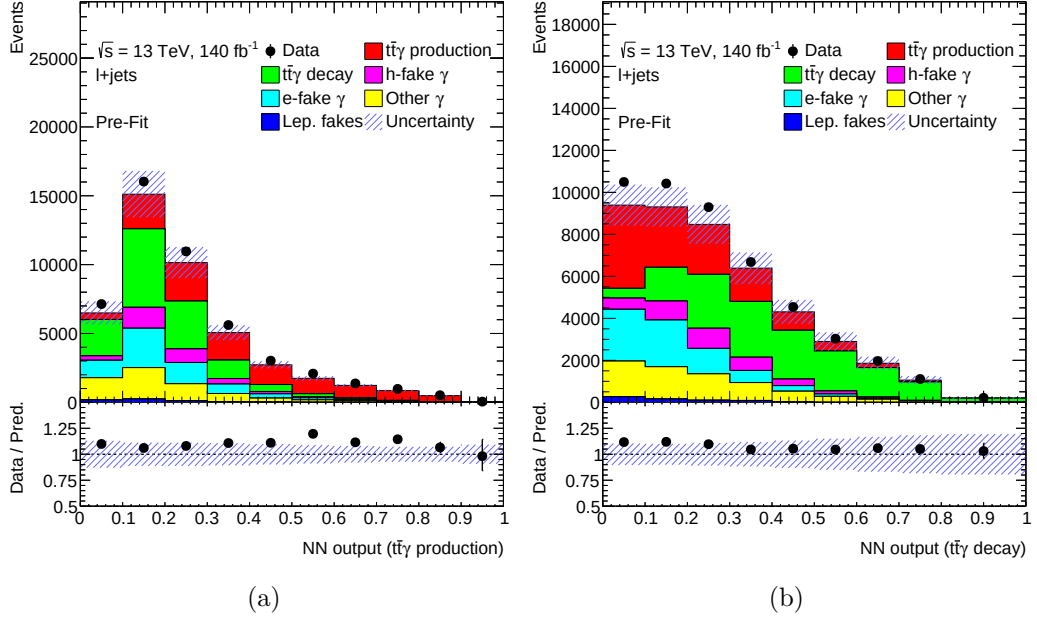


Figure 5.21.: The full output distributions of the (a) $t\bar{t}\gamma$ (prod.) classifier and (b) the $t\bar{t}\gamma$ (dec.) classifier. The uncertainty band corresponds to the total uncertainty, including stat. and syst. uncertainties.

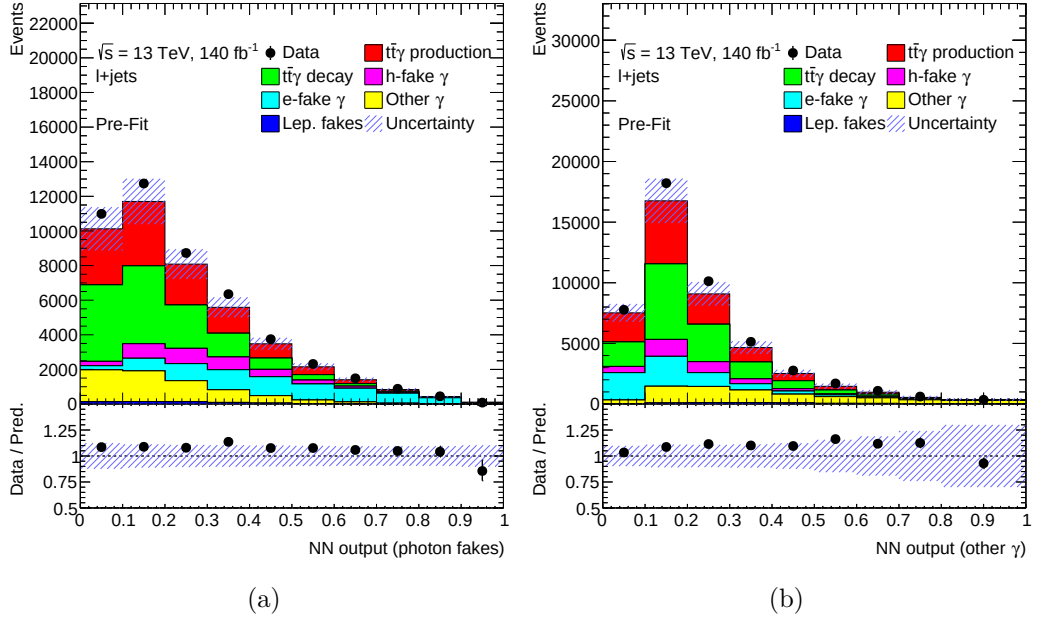


Figure 5.22.: The full output distributions of the (a) "fake photons" classifier and the (b) "other photon" classifier. The uncertainty band corresponds to the total uncertainty, including stat. and syst. uncertainties.

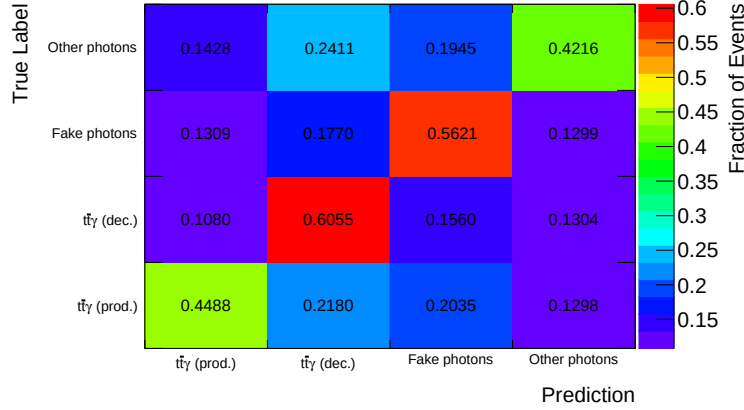


Figure 5.23.: The confusion matrix for all four classifiers.

Definitions of signal and control regions

For the final cross-section fits pure signal and control regions are needed. For the lepton+jets channel, this means that four regions are needed, where each region is enhanced in one of the four classes defined for the neural network training. The signal region is expected to be dominated by $t\bar{t}\gamma$ from production, while one control region should be dominated by $t\bar{t}\gamma$ from decay, one by photon fake events (e and $h \rightarrow \gamma$ fakes) and one by the prompt photon background. These regions are defined using the neural network output. As they form a four-dimensional space, it is not easy to visualise it. To get a better understanding of this complex space two-dimensional distributions were created, showing one classifier output as a function of another classifier output. For the four classes, this yields six different possibilities. To see the correlations of two classifiers for a certain category there is one distribution for each of the six output classifier combinations. The categories that were checked individually are: $t\bar{t}\gamma$ from production, $t\bar{t}\gamma$ from decay, $h \rightarrow \gamma$ fakes, $e \rightarrow \gamma$ fakes, $Wt\gamma$ and other prompt photons.

Hence in total 36 plots were checked to define the signal and control regions²². An example plot is shown in Figure 5.24 for the “fake photons” and “other photon” classifiers for $t\bar{t}\gamma$ from production events. One can see that the $t\bar{t}\gamma$ from production events accumulate in the lower left corner. This shows that $t\bar{t}\gamma$ from production events get small output values from the “fake photons” and the “other photon” classifier, as it should be. If one compares this with the two-dimensional distributions for the same output classifiers but for $h \rightarrow \gamma$ fakes and $e \rightarrow \gamma$ fakes (see Figure 5.25) events, one can see that most events have higher output values on the “fakes” classifier (as expected) but similar output values for the “other photon” classifier. One can also see that $e \rightarrow \gamma$ fakes are better classified than $h \rightarrow \gamma$ fakes. The final signal and control regions were defined by

²²These are the regions used in the fits (see Chapter 6). The signal region is dominated by the signal process. The other regions are called control regions, but they are all treated the same way in the fit.

5. Analysis Strategy

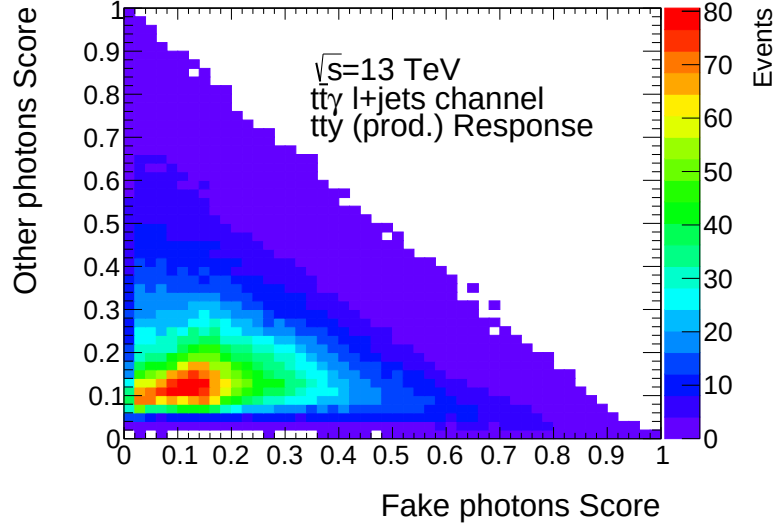


Figure 5.24.: The two dimensional distribution of MC events for $t\bar{t}\gamma$ from production with respect to the “fake photons” and the “other photon” classifiers.

using these two-dimensional distributions trying to separate the classes as much as possible and keeping enough statistics in each region²³. As there is a trade-off between the different classes and the focus of this analysis is on the $t\bar{t}\gamma$ (from production and decay) contributions, first, these two regions were defined. Especially the $t\bar{t}\gamma$ from production region is important as it defines the region for the EFT fits and the combination with the $t\bar{t}Z$ analysis. For the signal region, the following cuts were found:

- $t\bar{t}\gamma$ (dec.) classifier below 0.15, “fake photons” classifier below 0.2 and “other photon” classifier below 0.5

Next, the control region for $t\bar{t}\gamma$ from decay was defined to be:

- $t\bar{t}\gamma$ (dec.) classifier above 0.25 and “other photon” classifier below 0.4

From the two classes left, the two photon fake categories are the much larger background. Therefore this region was defined next:

- $t\bar{t}\gamma$ (dec.) classifier below 0.15, “fake photons” classifier above 0.2 and “other photon” classifier below 0.5

All remaining events form the control region for the prompt photon background. In general, due to the focus on getting pure regions for $t\bar{t}\gamma$ from production and decay the two regions for photon fakes and prompt photon background are less pure. Deriving purer regions for photon fakes and the prompt photon background would result in less pure regions for both $t\bar{t}\gamma$ contributions and less precise differential measurements. The

²³The corresponding two-dimensional distributions not shown here are found in Appendix D.1.

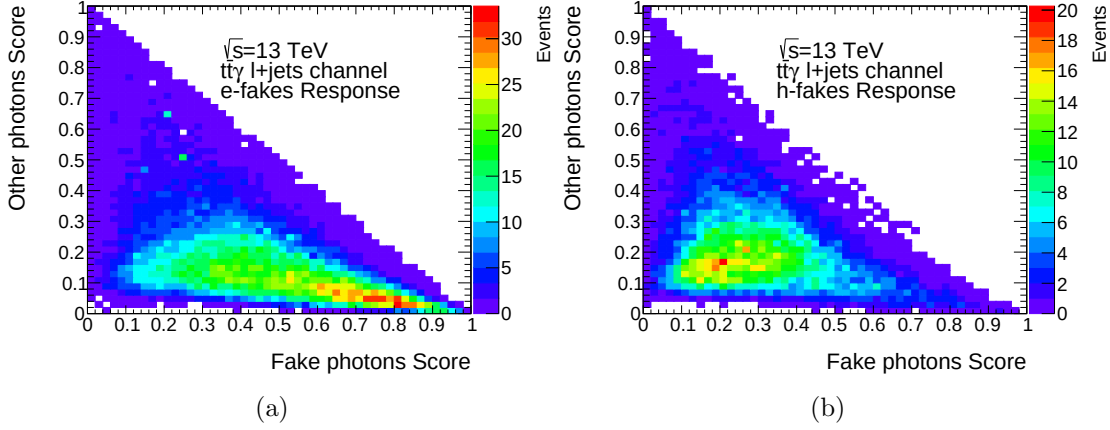


Figure 5.25.: The two dimensional distribution of MC events for (a) $e \rightarrow \gamma$ fakes and (b) $h \rightarrow \gamma$ fakes, both with respect to the “fake photons” and the “other photon” classifiers.

distributions chosen to be used in the fit in each of the four regions are the following ones:

- Signal region: Neural network output distribution for $t\bar{t}\gamma$ from production classifier
- Control region $t\bar{t}\gamma$ (dec.): Neural network output distribution for $t\bar{t}\gamma$ from decay classifier
- Control region for photon fakes: Subleading PCBT score
- Control region for prompt photon background: Subleading PCBT score

For the two $t\bar{t}\gamma$ enhanced regions the respective neural network output distributions were used as these provide the best sensitivity. The other two control regions are not as pure and here different distributions were chosen. As photon fakes and other prompt photon background show a different behaviour as a function of the subleading PCBT score²⁴, this variable was chosen. The four final distributions used in the inclusive cross-section fits are shown in Figure 5.26.

5.3.3. Event classification in the dilepton channel

Variable selection

The dilepton channel has a much lower background than the lepton+jets channel due to the second lepton which suppresses events contributing to photon fake events, especially $e \rightarrow \gamma$ fakes which become almost negligible. $Z\gamma$ events contribute to ee and $\mu\mu$ subchannels and $Wt\gamma$ to all combinations of leptons. In contrast to the lepton+jets

²⁴In most cases no second b -jet is expected.

5. Analysis Strategy

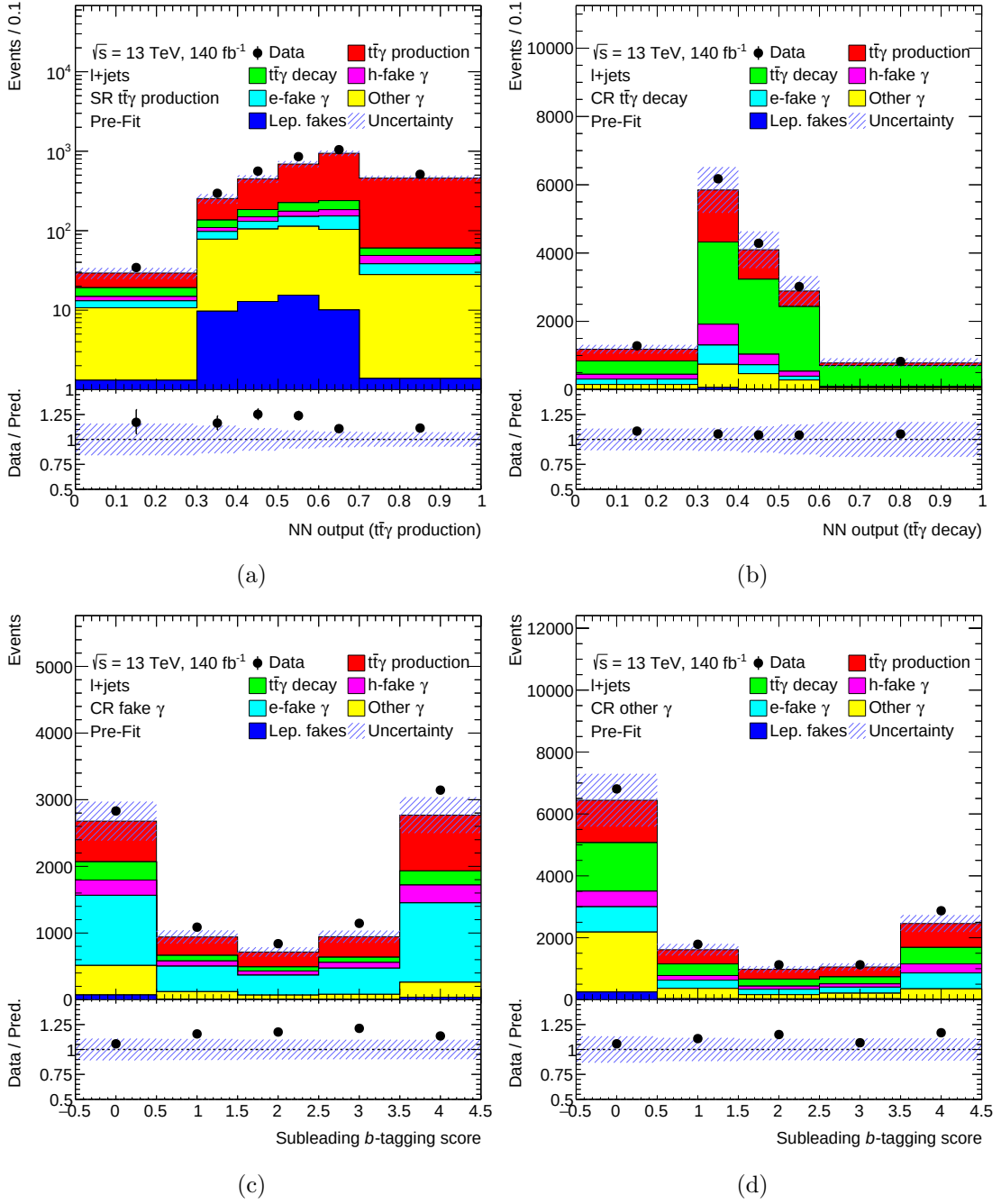


Figure 5.26.: Distributions of the $t\bar{t}\gamma$ (prod.) classifier for the signal region (SR) (a), the $t\bar{t}\gamma$ (dec.) classifier for the $t\bar{t}\gamma$ (dec.) control region (CR) (b), the subleading PCBT score for the photon fakes CR (c) and for the other prompt photon background CR (d). The uncertainty band corresponds to the total uncertainty, including statistical and systematic uncertainties.

channel, where about 60% $t\bar{t}\gamma$ events are expected (after pre-selection), in the dilepton channel almost 80% of the events are $t\bar{t}\gamma$ (production and decay counted together). This allows a binary classification instead of the multiclass classification and one can expect that fewer variables are sufficient. $t\bar{t}\gamma$ from production was defined as the signal, while everything else composes the background. The following variables were used²⁵:

- ΔR and invariant mass between photon and closest lepton
- Photon p_T
- Invariant mass between photon and closest b -jet
- Photon E
- Scalar sum of all jet p_T
- Jet p_T and E of the first two jets
- ΔR between photon and closest b -jet
- E_T^{miss}
- Number of jets
- Photon η
- Number of b -jets
- Photon ϕ

Those variables that are related to the photon kinematics are used for the separation between $t\bar{t}\gamma$ from production and $t\bar{t}\gamma$ from decay. The other variables shall help to separate the other backgrounds from $t\bar{t}\gamma$ from production. Again, as already seen in the lepton+jets channel, the MC simulations underestimate the number of events observed in data. The offset seems to be constant, so the shapes of the distributions are described well, which is the important point to use the variables in the neural network training, besides their separation power. Both, distributions and separation plots are shown in Appendix D.2. In the tails sometimes there seems to be some mismodelling and a slope in the ratio panel is seen. Two important examples are shown in Figure 5.27. This kind of mismodelling has been accepted due to the importance this variable has shown in the training. Unfortunately, the mismodelling got enhanced in one of the folds, so it shows a visible impact on how data behaves when it gets processed by the neural network (see next Section). It could be proven that this has no impact on the final results (see next but one Section).

²⁵Again, the variables are ordered with respect to their separation power between both $t\bar{t}\gamma$ contributions.

5. Analysis Strategy

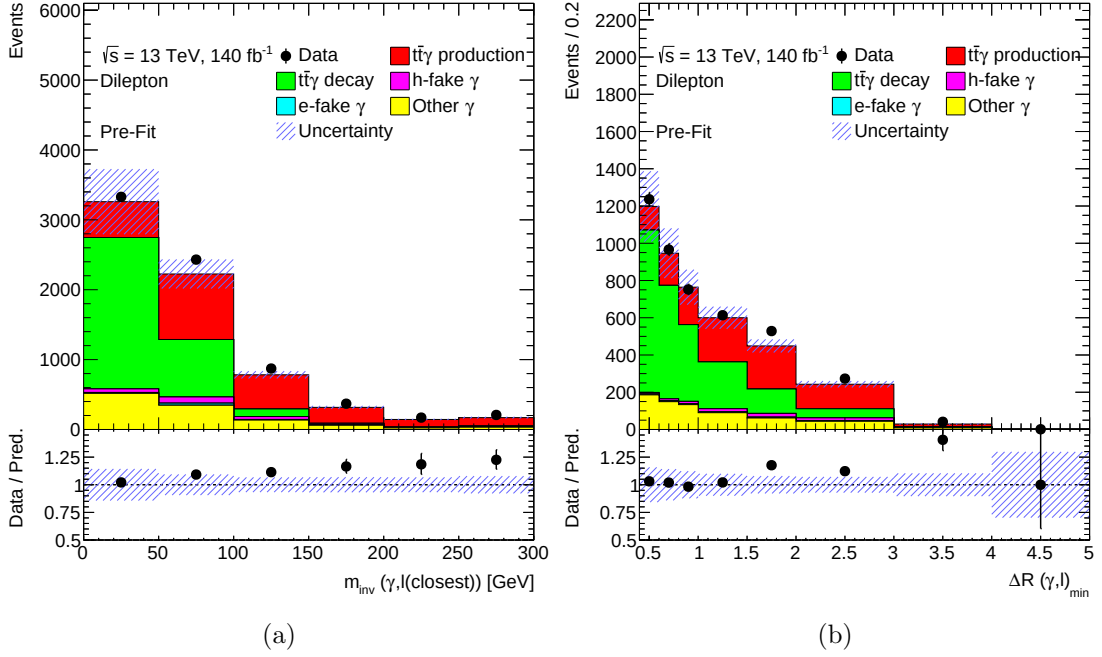


Figure 5.27.: The invariant mass (a) and ΔR (b) distributions between the photon and the closest lepton. The uncertainty band corresponds to the total uncertainty, including stat. and syst. uncertainties.

Training, optimisation and results

For the initial training, the same architecture and hyperparameters were chosen as in the lepton+jets channel:

- four hidden layers with 50 neurons each
- one dropout layer following the first hidden layer
- no batch normalisation layers
- learning rate: 0.01
- activation function: rectifier linear unit
- patience: 100 epochs
- epochs: 5000
- batch size: 100000
- validation size: 5% of the fold used for training
- dropout probability: 0.2

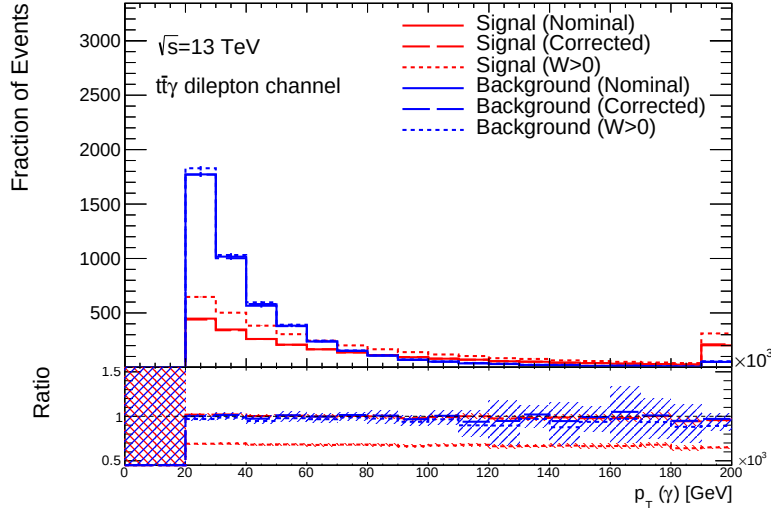


Figure 5.28.: Comparison of distributions for the p_T of the photon in the dilepton channel to check the impact of negative weights on signal ($t\bar{t}\gamma$ from production) and background (all other processes). The uncertainties are statistical.

- output activation: softmax
- metrics: accuracy
- loss: binary cross-entropy

Again, the input variables are transformed to the interval $[0,1]$, a 5-fold cross-validation is applied and negative weights are not taken into account, but the positive weights are scaled so that the original normalisation is kept the same. As before the input distributions are only varied slightly by this procedure and the overall shape is kept. This can be seen again in the constant ratios between nominal distributions and modified ones. An example plot presenting the photon p_T is shown in Figure 5.28. For all other variables, the corresponding plots are in Appendix D.2. For the hyperparameter optimisation, the following parameters were varied:

- hidden layers
- neurons per layer
- dropout layers
- batch normalisation layers

A grid search was performed, where the number of hidden layers was varied between one and six, while the number of nodes per layer was varied for each layer independently between 10 and 90. The numbers were chosen randomly. To assign a dropout layer

5. Analysis Strategy

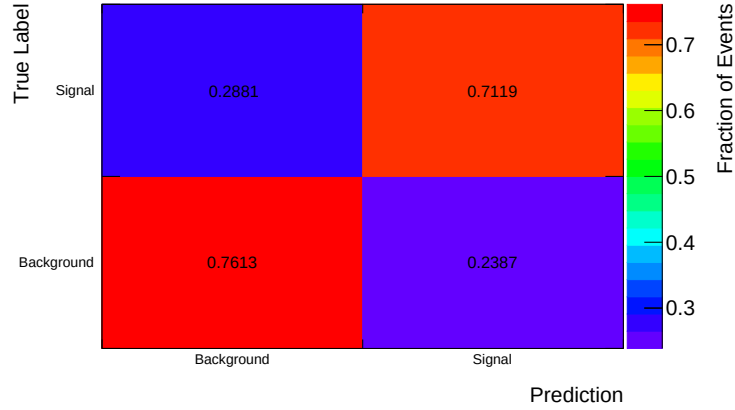


Figure 5.29.: The confusion matrix for the binary classifier in the dilepton channel.

behind a hidden layer or the input layer was again decided randomly, which resulted in zero to five of such layers. The same was done with batch normalisation layers. In total 100 different neural network architectures were tested. In the end, the initial architecture proved to be the one to use, based on the confusion matrix (see Figure 5.29). Looking at the comparison of the testing and the validation loss for all five folds (see Figures 5.30 to 5.32) shows a coherent convergence behaviour. The agreement of training and testing sets for signal and background is given within statistical uncertainties. For fold one, three and five a slight slope for the background is seen, which is not significant. As an example, the distributions for fold one are shown in Figure 5.33a, while the distributions for all other folds are in Appendix D.2. No further sign of overtraining is observable. The corresponding ROC curves are shown in Figure 5.33b and there is no indite of overtraining. The five neural networks are applied the same way, as it was done for the lepton+jets channel. The final output distribution is shown in Figure 5.34. A good separation between $t\bar{t}\gamma$ from production and decay is visible, which can also be seen in the confusion matrix in Figure 5.29. The data point in bin 8 was further investigated to understand the deviation between data and prediction and it could be proven that it does not cause any problems or that any mistake was made previously. The results are presented in the next section.

Cross checks for deviating data point

In Figure 5.34 it can be seen that for bin number eight the difference between data and MC prediction gets large. Investigating this, showed that it is a consequence of the procedure to assign the events to the folds and that it has no impact on the measurement. The first check that was done, was to look for a mismodelling in one of the three subchannels ee , $\mu\mu$ and $e\mu$. No relevant mismodelling could be found. The next step was to redo the neural network training by sorting out the variables with the smallest χ^2 probability of data and MC to agree. First, only those variables having a χ^2 probability

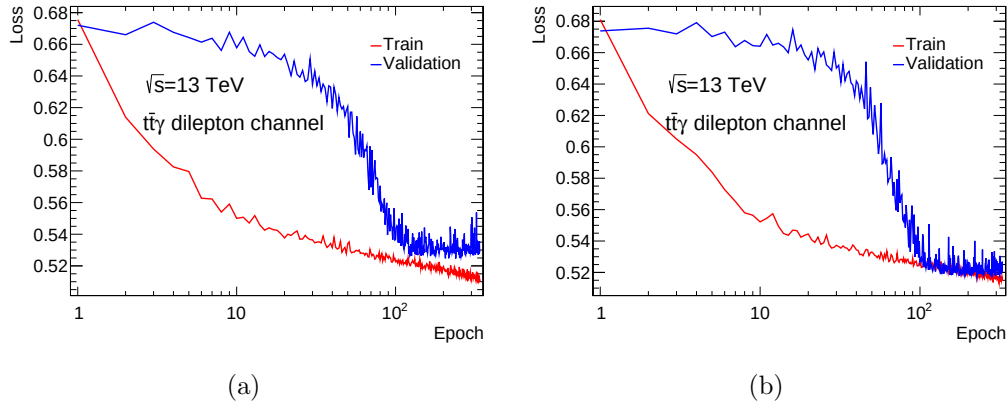


Figure 5.30.: The evolution of the loss function for the first (a) and second fold (b) for training and validation samples.

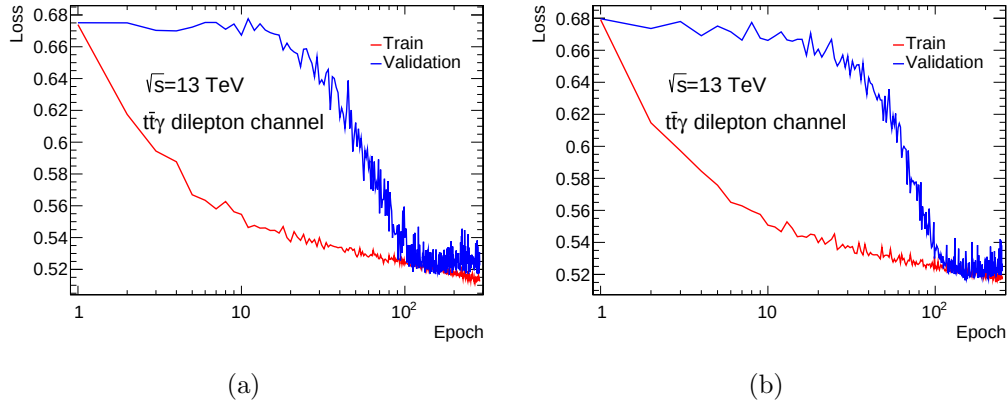


Figure 5.31.: The evolution of the loss function for the third (a) and fourth fold (b) for training and validation samples.

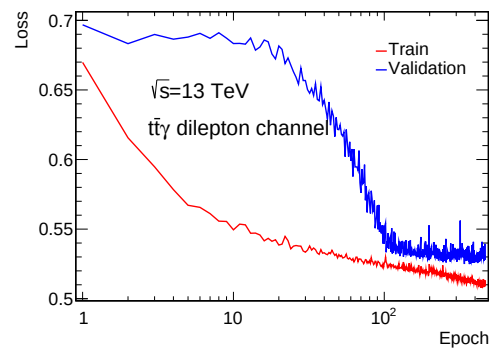
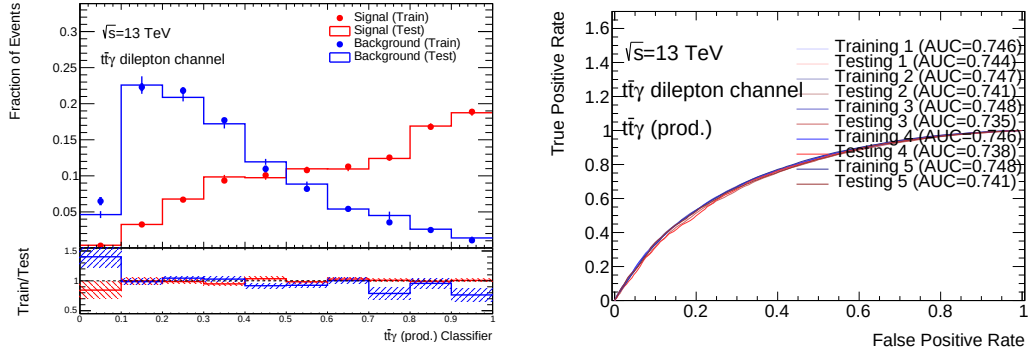


Figure 5.32.: The evolution of the loss function for the fifth fold for training and validation samples.

5. Analysis Strategy



(a) Output distributions for signal and back- (b) The ROC curves (and AUC values) for all
ground for fold 1. Only statistical uncer- 5 folds for training and testing.
tainties are shown.

Figure 5.33.

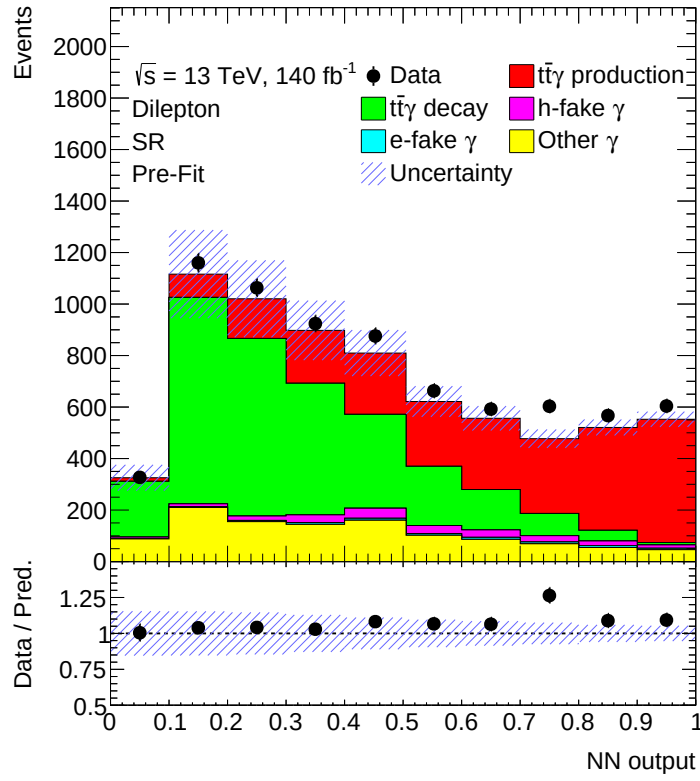


Figure 5.34.: The full output distribution for the binary classification in the dilepton channel. The uncertainty band corresponds to the total uncertainty, including stat. and syst. uncertainties.

below 0.2 were removed. These were the following ones:

- E (1st jet)
- $\Delta R(l, \gamma)$
- N_{jets}
- H_T

Afterwards, all variables with a χ^2 probability below 0.3 were removed. The following additional variables were removed from the training:

- $\eta(\gamma)$
- $m_{\text{inv}}(l, \gamma)$

While the bump was still visible in the first check, it was gone in the second check. So as a closure test, the neural network training was repeated by removing the variables with a χ^2 probability between 0.2 and 0.3. Again, the bump was still visible. This meant, that not only one variable is the reason for the bump but several of them. A look at the permutation importance (see Figure 5.35) of the original neural network training showed that the two most important variables were $\Delta R(l, \gamma)$ and $m_{\text{inv}}(l, \gamma)$. Removing only these two variables made the bump disappear. Looking at the ROC curves (see Figure 5.36), it can be seen that the separation is much worse than in the original training (see Figure 5.33b). By reducing the complexity of the network in three different ways by either halving the number of layers, halving the number of nodes per layer or halving both, it could be shown that the neural network architecture has no impact on the bump. Scaling the weights of the signal events such that signal and background have the same sum of weights only smeared out the bump. From these results, it was concluded that the bump is a feature of the simulated events used for training. Checking the output distributions for data and MC for the testing set for each fold individually, showed a small bump in each of the folds, especially in fold 4 (see Figure 5.37). So this showed that the main reason behind the bump is due to the distribution of data and simulation into the five folds. In the tests of the fit setup it could be shown that the bin with the bump has no significant impact on the final measurement (see Section 6.2), and in consequence this neural network setup and its final output distribution, as shown in Figure 5.34, can be used for the fit to extract the cross-section.

5. Analysis Strategy

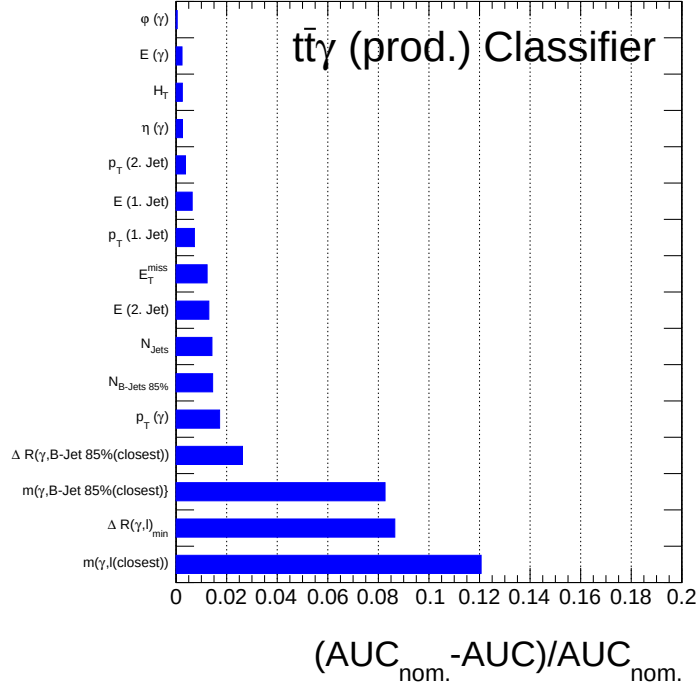


Figure 5.35.: Permutation importance for the binary classification in the dilepton channel.

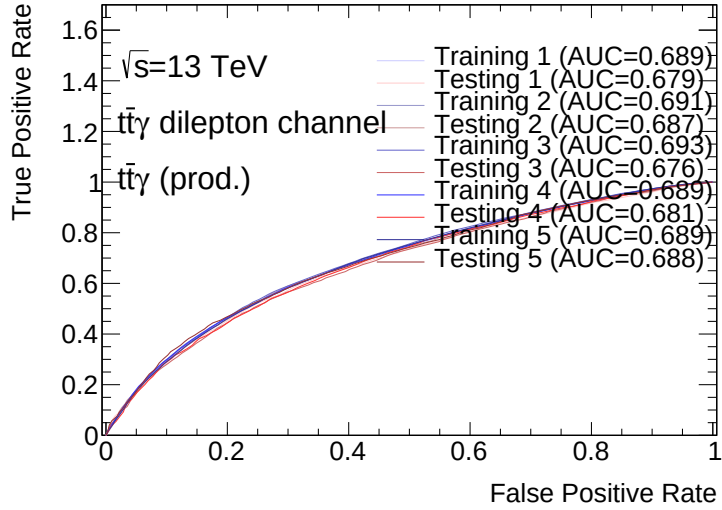


Figure 5.36.: The ROC curves (and AUC values) for all 5 folds for training and testing for the neural network training removing $\Delta R(l, \gamma)$ and $m_{\text{inv}}(l, \gamma)$ as input variables.

5.3. Neural network

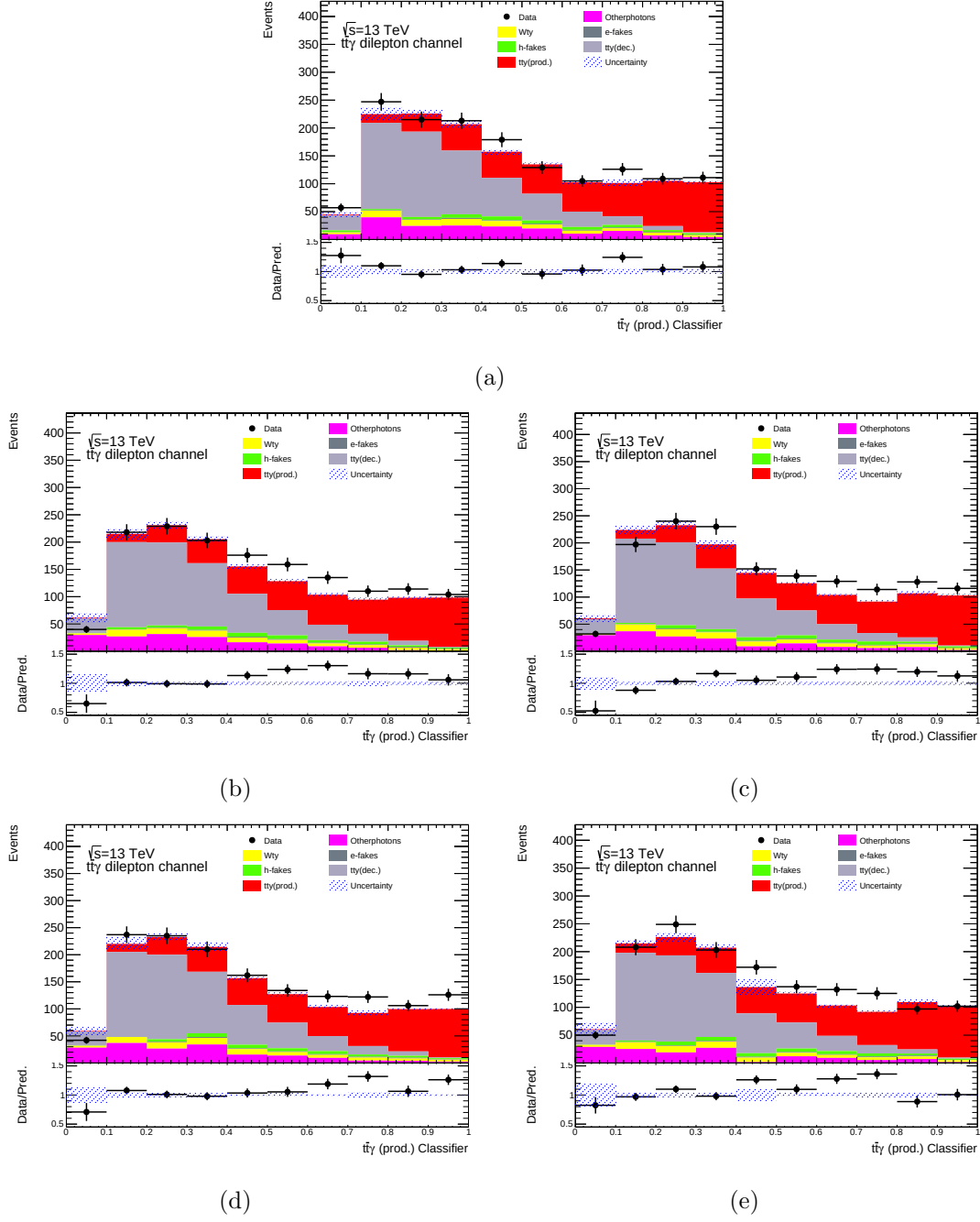


Figure 5.37.: The distributions of the neural network output for each fold for the original training. (a) Fold 1, (b) Fold 2, (c) Fold 3, (d) Fold 4 and (e) Fold 5.

5.4. Systematic uncertainties²⁶

Several systematic uncertainties need to be taken into account. On the one hand, there are theoretical uncertainties affecting the simulations and therefore the SM prediction. These uncertainties are described in Section 5.4.1. On the other hand, the detector itself and the event reconstruction are sources of systematic uncertainties, which are generally called experimental uncertainties. These are described in Section 5.4.2. The individual uncertainties are taken as uncorrelated, while each uncertainty is taken to be correlated among the different regions used in the final fits (see Chapter 6). Before describing the actual systematic uncertainties (see Sections 5.4.1 and 5.4.2) three procedures (symmetrisation, pruning and smoothing) on how systematic uncertainties are treated are explained in the following.

Symmetrisation of systematic uncertainties

Uncertainties are either two-sided or one-sided. Two-sided means that there is an up and a down variation of a given uncertainty (e.g. scale uncertainties), where the up variation is connected to a larger value of a certain parameter and the down variation to a smaller value. These uncertainties are symmetrised to correct for asymmetric up and down variations or cases where both variations are in the same direction. The symmetrisation formula giving the symmetrised uncertainty σ_{new} is the following:

$$\sigma_{\text{new}} = \frac{|\sigma_{\text{up}}| + |\sigma_{\text{down}}|}{2}, \quad (5.27)$$

where σ_{up} is the up variation and σ_{down} is the down variation. In case of uncertainties where only one variation is available (e.g. parton shower uncertainty), the difference to the nominal prediction is mirrored with respect to the nominal prediction.

Pruning

The overall impact on the final measurement of some of the uncertainties that are taken into account as nuisance parameters²⁷ in the fit is negligibly small. Therefore a “pruning technique” is used. This reduces the number parameters used in the fit and helps to make the fit numerically more stable and faster, without affecting the precision of the measurement. A threshold to drop uncertainties was set at 0.1% on the relative impact of an uncertainty on the total input for both effects that an uncertainty can induce, shape and normalisation effects (see Section 6.1). As a closure test to prove that the precision of the measurement is not artificially improved, this threshold was compared to a threshold of 0.05%. The respective Asimov fits showed only an impact on the rounding digit, which is the third digit.

²⁶The information in this chapter are based on the ATLAS-internal documentation of this analysis.

²⁷How uncertainties are incorporated as nuisance parameters in fit is described in Section 6.1.

Smoothing

Uncertainties that are evaluated either by using an alternative sample, as is done for some of the theoretical uncertainties, or those experimental uncertainties not evaluated with a reweighting method, are smoothed. Smoothing means that the histograms of the systematic uncertainties are adjusted following an algorithm (here the algorithm described in Reference [281] was used) to avoid statistical fluctuations, which are physically not meaningful. The underlying assumption is that the behaviour of uncertainties is smooth and that large fluctuations in their estimation are due to statistical noise.

5.4.1. Theoretical uncertainties

For the main processes considered in this analysis, which are $t\bar{t}\gamma$ from production and $t\bar{t}\gamma$ from decay, $Wt\gamma$ and $t\bar{t}$, different theoretical uncertainties are taken into account. These uncertainties are estimated with different generator setups described in Section 4.2, where the difference between the nominal prediction and the alternative prediction is taken to be the uncertainty. The alternative predictions are all normalised to the same fiducial cross section of the respective process.

$t\bar{t}\gamma$ (production and decay)

For both $t\bar{t}\gamma$ contributions, the same theoretical uncertainties are taken into account.

Scale uncertainties μ_R and μ_F :

Renormalisation and factorisation are mathematical procedures that allow calculations in QFT with physically meaningful results. The choice of the corresponding scales (for more details see Sections 2.4 and 4.2.1) should not affect the final results. For MC event generators has to be chosen. To estimate the effect of the scale choices on the final result one varies the scales by a factor of 2 up and down. This results in different event weights, which change the input distributions. The difference between the nominal and the varied distributions serves as a estimate of the uncertainty.

Parton shower uncertainty:

As QCD becomes non perturbative for energies approximately around and below 1 GeV, the whole process of hadronisation, low energy gluon emissions and jet formation, cannot be calculated from first principles. Hence parametric models are used and tuned to data. The differences between the models is quite significant. To account for the resulting uncertainty one compares different parton shower algorithms interfaced with the same event generator. In this case the event generator is MadGraph and the two shower algorithms are Pythia 8 and Herwig 7 (see Section 4.2 for the exact version numbers). The difference in the distributions is taken as a one-sided uncertainty and symmetrised accordingly.

5. Analysis Strategy

ISR radiation:

The frequency and the hardness of gluon emissions in the initial state is modelled by the parton shower algorithm, which depends on the value of the strong coupling constant α_s . Varying the value within its experimental uncertainties results in different tunes for the shower algorithm. The so called *var3c* parameter of the Pythia A14 tune, describing higher and lower ISR emission rates, is used to reweight the events and again the differences between the resulting distributions and the nominal are taken as up and down variations to account for this uncertainty.

PDF uncertainties:

As already described in Section 2.4, the inner structure of the colliding protons is described by the parton density functions (PDF). These functions cannot be derived from first principles. They are derived at a certain energy scale and afterwards extrapolated to higher energies through the DGLAP equations. The uncertainties on the PDFs also need to be taken into account. For this the so called PDF4LHC15 prescription is used [282]. This prescription provides event weights for 30 independent systematic variations and one baseline prediction. The difference of the baseline prediction with respect to each systematic uncertainty is taken as a one sided, independent uncertainty, which is symmetrised accordingly.

$Wt\gamma$

For the $Wt\gamma$ samples, the parton shower uncertainty is taken into account as well as the scale uncertainties. The same procedure as described for the $t\bar{t}\gamma$ samples is used. The resulting uncertainty is taken to be uncorrelated for $Wt\gamma$ and inclusive Wt samples.

$t\bar{t}$

The uncertainties described here are only relevant for the $t\bar{t}$ events with prompt photons surviving the overlap removal. Besides the aforementioned uncertainties on the renormalisation and factorisation scales μ_R and μ_F , the parton shower uncertainty and the ISR uncertainty (all treated in the same way as described for the $t\bar{t}\gamma$ samples), two further uncertainties are taken into account.

Uncertainty on the *hdamp* parameter:

The so called *hdamp* parameter controls the energy scale of the first hard emission in the development of the showering. The nominal value given in Section 4.2 is doubled. The difference of the distributions of the nominal sample and the alternative sample are taken as a one-sided uncertainty and symmetrisation is applied.

Matrix element matching to parton shower:

As quarks and gluons are generated at NLO with the matrix element generator, their matching with the parton shower algorithm is a non-trivial task, needed to avoid double

counting of phase space regions. The different matching schemes²⁸ (see for example [285]) lead to different results, which form a further source of uncertainty on the prediction. This uncertainty is taken into account by comparison with an alternative sample, where the matching within Pythia is changed. Therefore, the so called pT_{hard}-parameter is changed to the alternative model described in Reference [286]. The difference between this sample and the nominal sample is again taken as a one-sided uncertainty and symmetrised.

5.4.2. Experimental uncertainties

Experimental uncertainties are grouped by the reconstructed object they affect. These uncertainties are related to the object reconstruction, identification, resolution, calibration and possible isolation criteria. All these uncertainties, but two concerning E_T^{miss} , are two-sided. Further experimental uncertainties are related to the luminosity and pile-up.

Photons

The photon identification data-to-MC scale factors are derived using three different methods: radiative Z -decays, an electron extrapolation method with $Z \rightarrow ee$ events using the similarities of photons and electrons in the ECAL, and the matrix method with samples with isolated high- p_T photons. The results from each method are combined. The influence of isolation and identification requirements on the scale factors is studied to derive the uncertainties [252]. Further uncertainties are related to the combined energy calibration of photons and electrons [253].

Electrons

Besides the uncertainties related to the combined photon and electron energy calibration [253], there are several uncertainties on the electron data-to-MC scale factors, related to the identification, reconstruction, isolation and trigger efficiencies. The scale factors for electrons are derived with the tag-and-probe method using $Z \rightarrow ee$ events in slices of E_T^{miss} , η and p_T [254, 255].

Muons

To derive data-to-MC scale factors for muons, as well as their momentum calibration, the tag-and-probe method is employed. In this case $Z \rightarrow \mu\mu$ and $J/\Psi \rightarrow \mu\mu$ events are used. Uncertainties are estimated by performing the identification, isolation, reconstruction and trigger efficiency measurements under slightly modified conditions [255, 256].

Jets

For jet-related uncertainties, the so called reduction scheme is used [259]. This results in 30 independent uncertainties concerning the jet energy scale. It takes reconstruction

²⁸Pythia uses the CKKW matching [283] for dipole cascades [284, 285].

5. Analysis Strategy

and calibration effects into account related to pile-up, jet flavour composition, single-particle response, punch-through effects (jets passing through the calorimeter into the muon spectrometer), modelling and statistical uncertainties on the extrapolation of the jet calibration from the central region, high- p_T jet behaviour and b -jet energy scale uncertainties. For each uncertainty an up and a down variation is provided.

The jet energy resolution contributes with 12 uncertainties derived by an eigenvalue decomposition of the corresponding covariance matrix. One further uncertainty is added to account for data-MC (dis-)agreement in samples where the detector simulation is done with the fast simulation setup [224]. The jet energy resolution itself is measured by a comparison of the energy measurements of photons and Z bosons with those of jets. The impact of the jet-vertex cut is estimated by varying it [260].

b -tagging

B -tagging is done using a multiclass neural network, which tries to tag b -jets, c -jets and light jets [287, 288]. The systematic uncertainties assigned to its calibration are quite numerous and hence a decomposition of the eigenvalues of the covariance matrix is done. In this analysis the pseudo-continuous b -tagging (PCBT) is used, whose uncertainty decomposition gives 45 independent uncertainties for b -jets and 20 independent uncertainties for c -jets and light jets each. The PCBT is only calibrated for jets below a certain energy. For b -jets this limit is 400 GeV, for c -jets 250 GeV and for light jets 300 GeV [287, 288]. As there is a fraction (around 1%) of jets above these thresholds, a 50% normalisation uncertainty for events with jets above these thresholds is applied, independently for uncalibrated b -jets, c -jets and light jets, resulting in three additional uncertainties.

E_T^{miss}

As outlined in Section 4.3.5, E_T^{miss} is derived from the measurements of all other reconstructed objects plus the soft term [263, 264]. While the uncertainties on the energy measurements of the reconstructed objects are already taken into account, the uncertainties on the soft term are new. These uncertainties are added as well. They are derived from the parallel and perpendicular projections of the soft term onto the hard term in $Z \rightarrow \mu\mu$ events with one or zero jets. The root mean square of the deviations from the expectation (momentum imbalance) is taken as the uncertainty [263, 264].

Luminosity

The luminosity uncertainty of 0.83% is taken from Reference [210].

Pile-up

For the pile-up reweighting procedure an uncertainty is derived based on the pile-up dependence of two different algorithms that are used for instantaneous luminosity measurements, as pile-up and the instantaneous luminosity are closely related [289, 290].

5.4.3. Background normalisation uncertainties

The different background contributions are the subject of normalisation uncertainties. As the $t\bar{t}$ normalisation is known up to a precision of about 5% [291], this was the uncertainty assigned. In the lepton+jets channel the smaller other prompt photon backgrounds all receive a conservative uncertainty of 50%²⁹. The reason to have a large normalisation uncertainty for the $W\gamma$ process too, is that a significant data-MC mismatch was observed in the $W\gamma$ control region (see Section 5.2.1). With the large uncertainty the fit will be able to vary and constrain it accordingly. In the dilepton channel again all minor contributions to the other prompt photon category receive an uncertainty of 50%. Because $Z\gamma$ is well modelled (see Section 5.2.1), an uncertainty of 20% seems to be sufficient. For the $t\bar{t}\gamma$ from decay process in both channels a normalisation uncertainty of 20% was assigned (to account for the k-factor used; see Section 4.2.1) in cases where only one normalisation factor for the whole $t\bar{t}\gamma$ process is used in the fit (see Chapter 6). In cases where both normalisations, $t\bar{t}\gamma$ from decay and $t\bar{t}\gamma$ production, are left free floating, no uncertainty is assigned.

The photon and lepton fake contributions receive the uncertainties, which were derived for the respective scale factors (see Sections 5.2.2 to 5.2.4). While photon fakes are important in both channels, lepton fakes are only relevant for the lepton+jets channel, while being negligible in the dilepton channel.

²⁹In general this exceeds the precision of theory calculations.

6.1. Profile likelihood fits and inclusive cross-section measurements

To extract the cross-sections for $t\bar{t}\gamma$ from production and for $t\bar{t}\gamma$ from production and decay together, as described in Section 5.1, using the signal and control regions defined in Section 5.3, the method of maximum likelihood estimator is applied in different setups. The idea of the likelihood function is to describe the joint probability distribution of N measured data points, which depend on several parameters. Maximising the likelihood function L gives the most likely set of values for these parameters. For a fit to extract the signal strength μ with predictions for signal and background and a given number of nuisance parameters, the likelihood function L is defined in the following way:

$$L(\mathbf{n}, \boldsymbol{\theta}_0 | \mu, \boldsymbol{\theta}) = \prod_i P(n_i | \mu \cdot S_i(\boldsymbol{\theta}) + B_i(\boldsymbol{\theta})) \times \prod_j G(\theta_{j,0} | \theta_j, \Delta\theta_j) \times \prod_i P(n_i | \gamma_i \nu_i). \quad (6.1)$$

The likelihood depends on the data $\mathbf{n}$ with n_i entries in the i -th bin, the normalisation factor μ and the nuisance parameters $\boldsymbol{\theta}$, where $\boldsymbol{\theta}_0$ is the nominal value for all nuisance parameters. Each nuisance parameter represents one systematic uncertainty. S_i and B_i are the signal and background predictions for bin i respectively. Their distribution is taken to be a Poisson distribution P . The nuisance parameters are constrained by a Gaussian distribution G for each nuisance parameter, with $\theta_{j,0}$ being the nominal central value of the j -th nuisance parameter and $\Delta\theta_j$ the corresponding $1\text{-}\sigma$ uncertainty. Here, the Gaussian distributions are chosen to be standard Gaussians with $\theta_{j,0} = 0$ and $\Delta\theta_j = \pm 1$. Additional nuisance parameters are included for the uncertainty introduced by the statistics of the MC samples in each bin. One nuisance parameter per bin is assigned. They are described by Poisson distributions, where γ_i is a parameter and ν_i is the MC prediction for the respective bin i [292]. This likelihood is used in the case where

6. Results

both $t\bar{t}\gamma$ contributions are measured together as signal. In the case of the measurement where one normalisation factor for each $t\bar{t}\gamma$ contribution is used, the likelihood function changes to

$$L(\mathbf{n}, \boldsymbol{\theta}_0 | \mu_1, \mu_2, \boldsymbol{\theta}) = \prod_i P(n_i | \mu_1 \cdot S_{i,1}(\boldsymbol{\theta}) + \mu_2 \cdot S_{i,2}(\boldsymbol{\theta}) + B_i(\boldsymbol{\theta})) \times \prod_j G(\theta_{j,0} | \theta_j, \Delta\theta_j) \times \prod_i P(n_i | \gamma_i \nu_i). \quad (6.2)$$

The fit is done by minimising the negative log-likelihood $-\ln(L)$ simultaneously for μ and $\boldsymbol{\theta}$, which is equivalent to maximising L but numerically easier and more stable. The results are called $\hat{\mu}$ and $\hat{\boldsymbol{\theta}}$. For this the ATLAS internal framework TREx-Fitter¹ is used. It uses the CERN-software package ROOT [293], with the HISTFACTORY [294] package for using ROOFIT [295] and ROOSTATS [296].

The uncertainties on the nuisance parameters $\boldsymbol{\theta}$ and the signal strength μ are estimated using the profile likelihood ratio R_{PLR} . This ratio is defined as

$$R_{\text{PLR}} = \frac{L(\mathbf{n}, \bar{\mu}, \bar{\boldsymbol{\theta}})}{L(\mathbf{n}, \hat{\mu}, \hat{\boldsymbol{\theta}})}, \quad (6.3)$$

where $\hat{\mu}$ and $\hat{\boldsymbol{\theta}}$ are the values obtained from maximising the likelihood. $\bar{\boldsymbol{\theta}}$ and $\bar{\mu}$ are the set of optimised nuisance parameters and signal strength for one parameter η being varied, which can be either the signal strength or a nuisance parameter. By varying this one parameter and optimising all others, the respective likelihood is profiled. From the resulting profile, the uncertainty for each parameter is estimated. Because $-2\ln R_{\text{PLR}}$ follows a χ^2 distribution, the uncertainties are given for those values of η , where $-2\ln R_{\text{PLR}} = 1/2$. This evaluation is implemented using MINUIT and the MINOS and MIGRAD algorithms therein [297].

The number of expected events after the fit (post-fit values) depends on the nuisance parameters. In general, there are three ways in which nuisance parameters influence the number of expected events in a bin. Firstly, this is the influence of a finite number of MC events. The corresponding nuisance parameter is called γ_i , where the index i denotes the bin. For each bin, an independent nuisance parameter is assigned. Furthermore, each systematic uncertainty that is incorporated as a nuisance parameter into the fit has two effects on the post-fit values. One component is its effect on the shape of the fitted distribution, called σ_i , and the second component is the effect on the normalisation of the fitted distribution, called δ . If the overall normalisation is left free floating in the fit, as a signal strength μ_p , this is included as well. Hence the expected yields $\phi_{p,i}$ in a bin (of a given process p) is given as

$$\phi_{p,i} = \gamma_i \mu_p \nu_{p,i} \prod_j \sigma_{i,j} \delta_j, \quad (6.4)$$

where j runs over all nuisance parameters described by a Gaussian function. Before the fit all these factors are set to unity and $\phi_{p,i} = \nu_{p,i}$. After the fit, the yields are adjusted

¹More information on this tool can be found here: <https://trexfitter-docs.web.cern.ch/trexfitter-docs/>. The website is not publicly accessible.

according to the extracted values for the nuisance parameters. While the results for γ_i and μ_p can be taken directly from the fit, the results for the other nuisance parameters must be interpolated. For the shape component, a linear interpolation is used:

$$\sigma_{i,j} = 1 + I_{i,j} \text{ with } I_{i,j} = \begin{cases} \theta(I_{i,\text{up}} - I_{i,\text{nom.}}) & \theta \geq 0 \\ \theta(I_{i,\text{down}} - I_{i,\text{nom.}}) & \theta < 0 \end{cases}, \quad (6.5)$$

where θ is obtained from the fit and $I_{i,\text{up}}$, $I_{i,\text{down}}$ and $I_{i,\text{nom.}}$ are the predicted yields for the up variation, down variation and the nominal prediction respectively in the corresponding bin i for the respective nuisance parameter j . This interpolation is done for each nuisance parameter taken into account and not pruned away. For the normalisation component, an exponential interpolation is used as follows:

$$\delta_j = 1 + I_j \text{ with } I_j = \begin{cases} (I_{\text{up}}/I_{\text{nom.}})^\theta & \theta \geq 0 \\ (I_{\text{down}}/I_{\text{nom.}})^{-\theta} & \theta < 0 \end{cases}, \quad (6.6)$$

where θ is obtained from the fit and I_{up} , I_{down} and $I_{\text{nom.}}$ are the predicted yields for the up variation, down variation and the nominal prediction respectively for the total number of events for the respective nuisance parameter j . Again this is only applied to nuisance parameters not pruned away.

From the extracted value for μ and its uncertainty one can determine the measured fiducial inclusive cross-section with

$$\sigma_{\text{measured}} = \mu \cdot \sigma_{\text{pred.}}, \quad (6.7)$$

where $\sigma_{\text{pred.}}$ is the predicted fiducial cross-section. The prediction is calculated as

$$\sigma_{\text{pred.}} = \frac{N}{L}, \quad (6.8)$$

where N is the number of weighted MC-events in the fiducial phase space and L is the luminosity. Possible influences of the reconstruction efficiency and possible migration effects are described by the correction factor $C = N_{\text{reco}}/N$, where N_{reco} is the weighted number of reconstructed MC events. This factor is taken into account when normalising the MC events to the same fiducial cross-section and is therefore indirectly taken into account. No uncertainty on $\sigma_{\text{pred.}}$ is assigned as all theoretical uncertainties are taken into account for the extraction of μ and therefore are quantified in the uncertainty on μ .

6.2. Tests of the fit setup

Before performing the fit to data and extracting the values for the signal strength, different tests of the fit setups were done to check for any problems that might occur. The first step was an Asimov fit, where the MC prediction was also used as “pseudo-data” in the fit. This helps to check the general consistency and convergence of the fit.

6. Results

The next step was to unblind the fit but keep the signal strength blinded. This was done to investigate the pulls and constraints of some nuisance parameters. The pull is defined as $(\hat{\Theta} - \Theta_0)/\Delta\Theta$, where $\hat{\Theta}$ is the best-fit value of the corresponding nuisance parameter and Θ_0 the expectation (usually 0 or 1) and $\Delta\Theta$ being the pre-fit uncertainty on the nuisance parameter (usually 1). Constraints mean by how much the uncertainties on the nuisance parameters are reduced by the fit.

6.2.1. Asimov tests

Lepton+jets channel

In the lepton+jets channel, two fit setups were tested. The first setup uses two normalisation factors (NF) for both $t\bar{t}\gamma$ processes and no normalisation uncertainty for $t\bar{t}\gamma$ from decay. The second setup uses only one NF for a joint measurement of $t\bar{t}\gamma$ from production and decay. For this fit the normalisation uncertainty for $t\bar{t}\gamma$ from decay was used. In both cases, the resulting NFs are 1, as expected. The uncertainties on these Asimov results show the expected uncertainties for the final fit. These were obtained to be:

$$\Delta\mu_{t\bar{t}\gamma} = {}^{+0.072}_{-0.066}, \quad (6.9)$$

$$\Delta\mu_{t\bar{t}\gamma(\text{prod.})} = {}^{+0.073}_{-0.067}, \quad (6.10)$$

for the fit setup with 1 and 2 (only the uncertainty on the NF for $t\bar{t}\gamma$ from production is shown) NF respectively. As expected no pulls are observed. The main observations concerning the nuisance parameters are the respective constraints. While most nuisance parameters are unconstrained a few show relevant constraints. For the normalisation uncertainty on the $W\gamma$ contribution a constraint to +0.48 and -0.40 was observed in both fits. The large constraint is expected, due to the large prior uncertainty assigned on purpose. Further nuisance parameters with relevant constraints between 0.75 and 0.9 are:

- Parton shower uncertainties for $t\bar{t}\gamma$ from production and $t\bar{t}\gamma$ from decay,
- B-tagging uncertainty on light jets, first eigenvalue of principle component analysis (PCA) decomposition
- B-tagging uncertainty on b -jets, first and second eigenvalue of PCA decomposition (B0, B1)
- h- and e-fakes scale factor uncertainties,
- Uncertainty on uncalibrated light jets in PCBT,
- Jet pile-up density (called: rho topology),
- Lepton fake scale factor uncertainty.

In the case of the fit with one NF the normalisation uncertainty on $t\bar{t}\gamma$ from decay is also largely constrained. As in the case of the $W\gamma$ normalisation uncertainty, this is not a surprise, as the 20% are quite large. The constraints show that the uncertainties might be overestimated, but still, they are not large enough to hint at an undetected problem in the fit.

To analyse the importance of each nuisance parameter for the total uncertainty on the signal strength the following four fits are done:

- The fit is done with the respective nuisance parameter being fixed to its pre-fit value² and its up variation of the pre-fit uncertainty.
- The fit is done with the respective nuisance parameter being fixed to its pre-fit value and its down variation of the pre-fit uncertainty.
- The fit is done with the respective nuisance parameter being fixed to its post-fit value³ and its up variation of the post-fit uncertainty.
- The fit is done with the respective nuisance parameter being fixed to its post-fit value and its down variation of the post-fit uncertainty.

For each of these four fits the difference $\Delta\mu$ concerning the signal strength of the nominal fit is calculated. The larger $\Delta\mu$, the more important that nuisance parameter is. The ranking plots for both fit setups is shown in `Finumpages = 36`, figure 6.1. While most nuisance parameters have an expectation of 0, γ factors (white dots in the ranking plots) and additional normalisation factors (red dots in the ranking plots) have an expectation of 1. Hence their pulls are not centered around 0 but around 1 in the ranking plots⁴. In both cases, the ranking is quite similar, with the impact of the second NF for the fit with two NFs in one case and the normalisation uncertainty on $t\bar{t}\gamma$ from decay in the other case. The most relevant difference is that in the fit where the signal strength only describes $t\bar{t}\gamma$ from production the showering uncertainty “ $t\bar{t}\gamma$ Herwig 7” is much less important. This makes sense as this nuisance parameter is not directly relevant for $t\bar{t}\gamma$ from production. In general, the most important uncertainties are those on the $t\bar{t}\gamma$ modelling. But also jet, b -tagging and photon-related uncertainties play an important role. MC-related statistical uncertainties are important, too. The correlations between the nuisance parameters are mostly as expected because they occur between nuisance parameters describing similar systematic uncertainties. For completeness, the matrices are shown in Appendix E.1. Those uncertainties that are ranked quite high do also have large correlations with the signal strength, which is expected. The main observation is that the impact of the normalisation uncertainty in the fit with one NF has a negative correlation with the signal strength and in the fit with two NF the second normalisation factor has a positive correlation with the signal strength. The reason for this are the differences in the signal definition. In the end, for the lepton+jets channel consistent

²The nuisance parameter is taken as given by the respective distribution and not varied.

³The nuisance parameter’s optimal value determined by the fit is taken.

⁴This holds for all ranking plots shown in this thesis.

6. Results

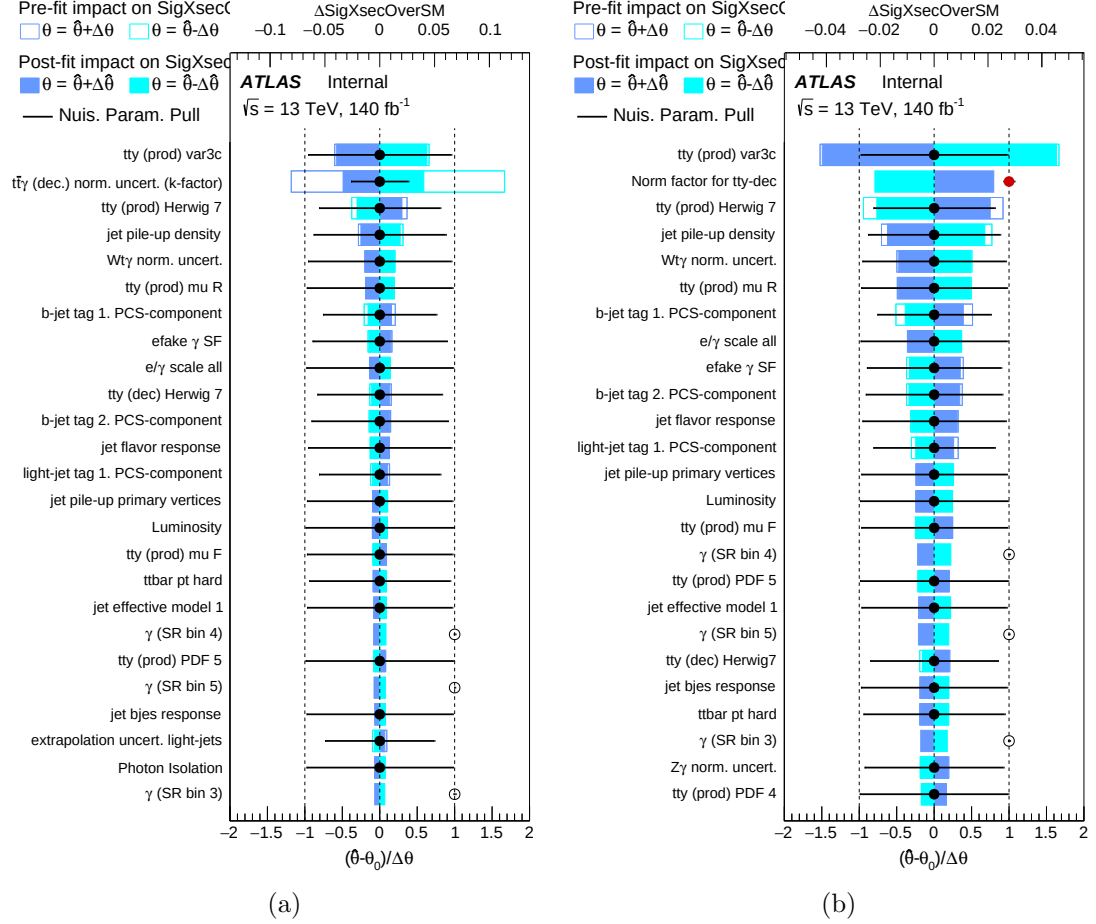


Figure 6.1.: Rankings of the 25 nuisance parameters with the largest impact on the signal strength for the Asimov fits in the lepton+jets channel. (a) shows the results for the fit setup with one NF, while (b) shows it for the fit setup with two NFs. The blue and turquoise bars show the pre-fit and post-fit importance expressed by $\Delta\mu$, while the black bars and dots show the post-fit values for the nuisance parameters and their uncertainties. More details about the different nuisance parameters are given in Section 5.4 and the corresponding references.

results are observed and the general fit setup shows no problems.

For completeness and for later comparisons with the unblinded results the impact of several groups of uncertainties are estimated. These estimations are done by applying the following procedure: first, the systematic uncertainties are grouped into different groups. For each group, the fit is done without those uncertainties. This results in a smaller uncertainty σ' for the signal strength. Doing a quadratic subtraction with the nominal uncertainty gives the impact σ_{group} to be

$$\sigma_{\text{group}} = \sqrt{\sigma^2 - \sigma'^2}. \quad (6.11)$$

In the tables 6.1 and 6.2 the groups and their impact is shown. The largest impact have the $t\bar{t}\gamma$ from production modelling uncertainties, jet-related uncertainties and b -tagging uncertainties. More details on these uncertainties are given in Section 5.4. This is in agreement with the results from the nuisance parameter ranking. Large differences between the fits are seen where expected. So the impact of $t\bar{t}\gamma$ from decay is less important for the signal strength in the fit with two NFs, as here the signal strength describes $t\bar{t}\gamma$ from production. In the case of the fit where both $t\bar{t}\gamma$ contributions are described with one NF, their impact gets large, while that of the $t\bar{t}\gamma$ from production modelling uncertainties reduces slightly. The statistical uncertainty also is larger for the fit with two NFs, as the signal region for $t\bar{t}\gamma$ from production has less statistics than the control region for $t\bar{t}\gamma$ from decay. In the fit with one NF, both regions become important and so the statistical uncertainty gets smaller. The uncertainty on the second NF (in the fit with two NF) has less impact on the signal strength than the normalisation uncertainty on $t\bar{t}\gamma$ from decay (in the fit with one NF). In the end, this results in a similar total uncertainty for the signal strength.

Dilepton channel

In the dilepton channel, the same two fit setups were used. One fit was used to measure $t\bar{t}\gamma$ from production as the signal strength and $t\bar{t}\gamma$ from production as a second normalisation factor. The second setup measures $t\bar{t}\gamma$ from production and decay together as the signal strength. In this case, where only one normalisation factor is used in the fit, $t\bar{t}\gamma$ from decay has a normalisation uncertainty of 20% assigned. Again the Asimov setup leads to the expected values of 1 for the normalisation factors. The uncertainties are estimated to be

$$\sigma_{1\text{NF}} = {}^{+0.074}_{-0.069}, \quad (6.12)$$

$$\sigma_{2\text{NF}} = {}^{+0.076}_{-0.070}. \quad (6.13)$$

In comparison to the lepton+jets channel, these uncertainties are a bit larger. This is mainly due to the statistical component.

No pulls are observed and in both cases, only one nuisance parameter gets constrained by the fits. This is the parton shower uncertainty for the $t\bar{t}\gamma$ sample. The ranking of the most important nuisance parameters shows a similar ranking between the two setups

6. Results

(see Figure 6.2). As seen before in the lepton+jets channel, the rankings in the dilepton channel look similar and therefore consistent between the two setups. Comparing the rankings plots in Figure 6.1 (lepton+jets channel) with those in Figure 6.2 (dilepton channel), one can observe that the same uncertainties are the most important ones. Most uncertainties appear in all four rankings and also at a similar position within the ranking. One main difference between the two channels can be seen for the $h \rightarrow \gamma$ fakes scale factor uncertainty, which is among the most important nuisance parameters in the dilepton channel, while not relevant in the lepton+jets channel, where the $e \rightarrow \gamma$ fakes scale factor uncertainty is important. The reason for this becomes clear when looking at the distributions for the fits (see Figures 5.26 and 5.34). In the lepton+jets channel the $e \rightarrow \gamma$ fake contribution has a much larger impact than in the dilepton channel. In contrast to that, the $h \rightarrow \gamma$ fake contribution, although not negligible, is less important in the lepton+jets channel, while being relatively important in the dilepton channel. Another difference concerns the $var3c$ uncertainty for $t\bar{t}\gamma$ from production. In the lepton+jets channel, it is the most important nuisance parameter, while in the dilepton channel, it is much less important. Nevertheless, the similar rankings demonstrate the consistency not just between the two fit setups but also between the channels. The correlations between the nuisance parameters are smaller in comparison to the lepton+jets channel, which mainly concerns the jet and b -tagging uncertainties, while modelling, normalisation and photon-related uncertainties are similar. Due to its importance the uncertainty on the $h \rightarrow \gamma$ fake scale factor shows larger correlations. The correlation matrices are shown in Appendix E.1.

The impact of different groups of uncertainties was studied for the dilepton channel in the same way as it is done for the lepton+jets channel. The corresponding tables with the concrete values are shown in Tables 6.1 and 6.2. While $t\bar{t}\gamma$ from production modelling uncertainties have become smaller, other uncertainties have become larger, especially the statistical uncertainties. Their impact has doubled, which can be expected as the whole number of events in the dilepton channel is roughly four times smaller than in the lepton+jets channel. Still one can say that the behaviour seen in both channels is consistent.

Combination

Once the robustness of the fit is ensured, at least at Asimov level, the combination of both channels can be done. Combining different decay channels into one single fit is the logical step, as the decay channel has no or only a small impact on the photon radiation. Hence, it is expected that both $t\bar{t}\gamma$ contributions behave the same in both channels. The previously seen consistency between the fits in the dilepton channel and the lepton+jets channel supports this prediction. Doing the combined fits can validate this.

The Asimov fit results on the signal strength and a comparison with the results from the individual fits are shown for both fit setups in Figure 6.3. The central values are 1, as they should be. The uncertainties are reduced, which is also expected. As the fits are not totally uncorrelated, the uncertainty reduction is a bit less than a factor of $1/\sqrt{2}$. Again for both fit setups a ranking was investigated (see Figure 6.4). The ranking follows more

6.2. Tests of the fit setup

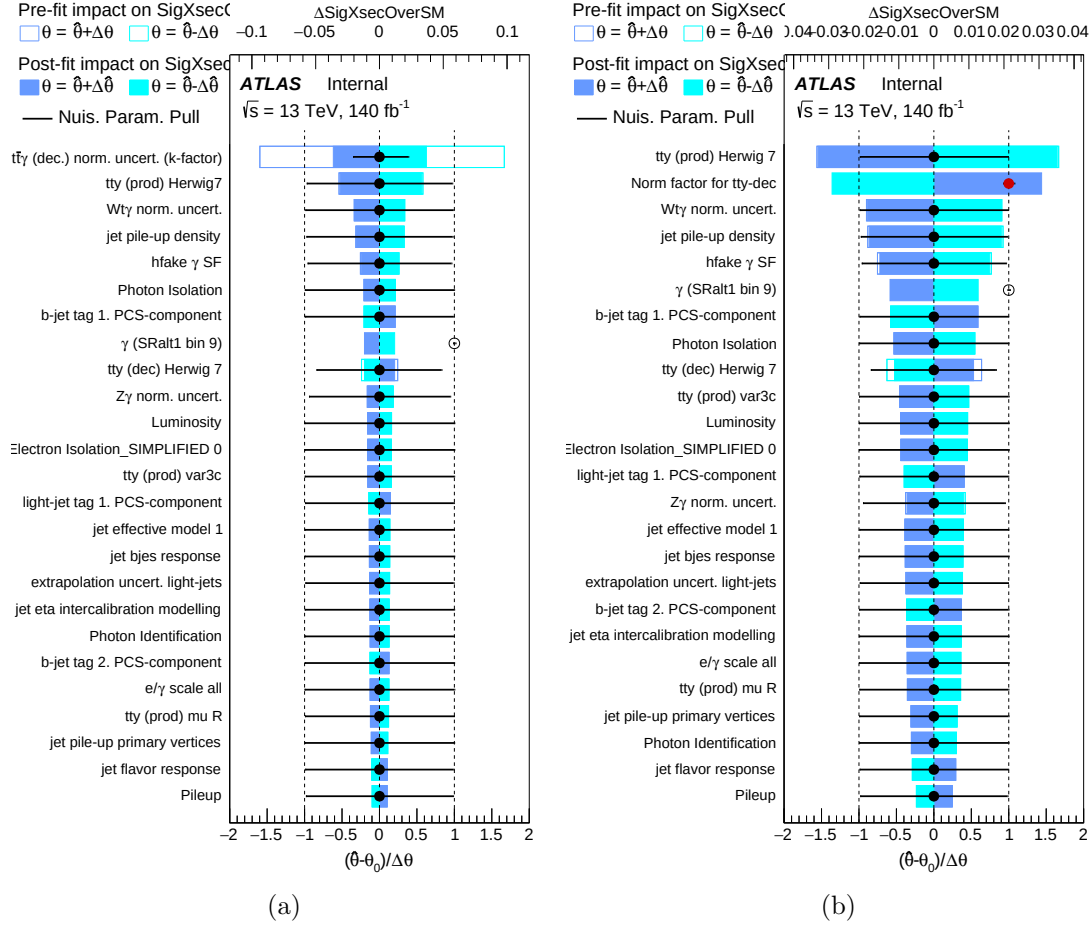


Figure 6.2.: Rankings of the 25 nuisance parameters with the largest impact on the signal strength for the Asimov fits in the dilepton channel. (a) shows the results for the fit setup with one NF, while (b) shows it for the fit setup with two NFs. The blue and turquoise bars show the pre-fit and post-fit importance expressed by $\Delta\mu$, while the black bars and dots show the post-fit values for the nuisance parameters and their uncertainties. More details about the different nuisance parameters are given in Section 5.4 and the corresponding references.

6. Results

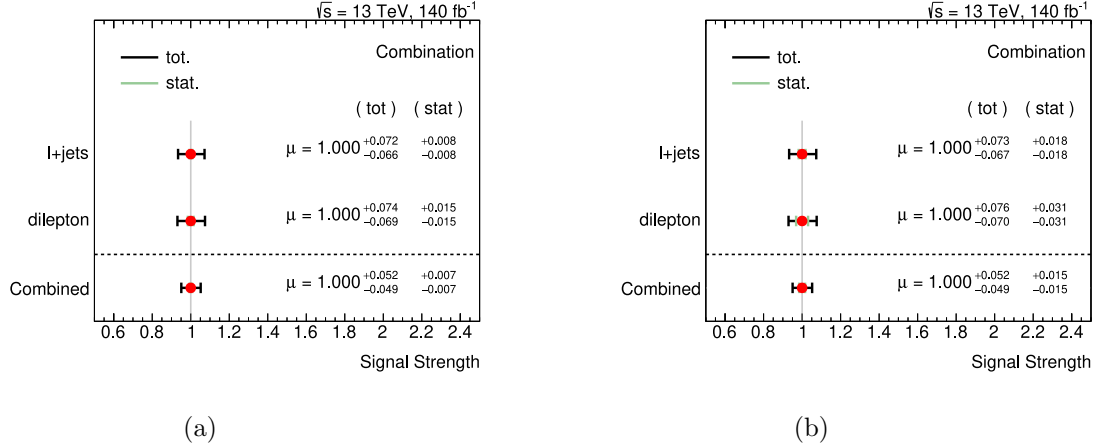


Figure 6.3.: A comparison of the best-fit value with one NF (a) and with two NF (b). The central value is, as expected, 1, while the uncertainties are reduced in the combinations with respect to the fits in the individual channels.

the ranking seen in the lepton+jets channel, which seems to dominate the fit. Also, the constraints are similar to the lepton+jets channel. The combined correlations do not reveal anything unexpected (see Appendix E.1). The final step was to extract the grouped uncertainties and compare them to the uncertainties from the individual fits. Table 6.1 and 6.2 provide the results. The main reduction comes from the $t\bar{t}\gamma$ modelling uncertainties, the jet-related uncertainties and the b -tagging uncertainties. In the case of the fit with one NF also the normalisation uncertainty on $t\bar{t}\gamma$ from decay has a strongly reduced impact.

Table 6.1.: Summary of the impact of the systematic uncertainties on the signal strength for the Asimov fits with 1 NF grouped into different categories. The quoted uncertainties are obtained by repeating the fit with certain sets of nuisance parameters removed from the fit and subtracting in quadrature the resulting uncertainty from the uncertainty from the full fit. For the total uncertainty average means the sum of the up and down uncertainty divided by two.

Group of uncertainties	Lepton+jets	Dilepton	Combination
Total uncertainty (average)	0.069	0.072	0.051
Statistical uncertainty	0.008	0.015	0.007
MC statistical uncertainty	0.012	0.015	0.009
Normalisation uncertainty for $t\bar{t}\gamma$ (dec.)	0.37	0.036	0.021
Modelling uncertainties			
$t\bar{t}\gamma$ (prod.) modelling	0.049	0.035	0.029
$t\bar{t}\gamma$ (dec.) modelling	0.009	0.012	0.009
Background modelling	0.007	0.003	0.006
Prompt background normalisation	0.015	0.023	0.022
Experimental uncertainties			
Jet uncertainties	0.028	0.026	0.016
B-tagging uncertainties	0.021	0.017	0.010
Extrapolation uncertainties	0.006	0.008	0.004
Fake lepton background estimate	0.004	-	0.001
Fake photon background estimate	0.010	0.015	0.016
Photon	0.005	0.014	0.008
Missing transverse energy	0.003	0.003	0.004
Other experimental	0.014	0.017	0.014

6. Results

Table 6.2.: Summary of the impact of the systematic uncertainties on the signal strength for the Asimov fits with 2 NF grouped into different categories. The quoted uncertainties are obtained by repeating the fit with certain sets of nuisance parameters removed from the fit and subtracting in quadrature the resulting uncertainty from the uncertainty from the full fit.

Group of uncertainties	Lepton+jets	Dilepton	Combination
Total uncertainty (average)	0.070	0.073	0.051
Statistical uncertainty	0.018	0.031	0.015
MC statistical uncertainty	0.012	0.014	0.009
Normalisation factor for $t\bar{t}\gamma$ (dec.)	0.022	0.030	0.021
Modelling uncertainties			
$t\bar{t}\gamma$ (prod.) modelling	0.052	0.037	0.029
$t\bar{t}\gamma$ (dec.) modelling	0.005	0.012	0.008
Background modelling	0.007	0.005	0.005
Prompt background normalisation	0.015	0.022	0.022
Experimental uncertainties			
Jet uncertainties	0.028	0.026	0.015
B-tagging uncertainties	0.021	0.017	0.010
Extrapolation uncertainties	0.005	0.009	0.004
Fake lepton background estimate	0.004	-	0.001
Fake photon background estimate	0.009	0.017	0.016
Photon	0.004	0.014	0.007
Missing transverse energy	0.002	0.004	0.003
Other experimental	0.014	0.017	0.014

6.2. Tests of the fit setup

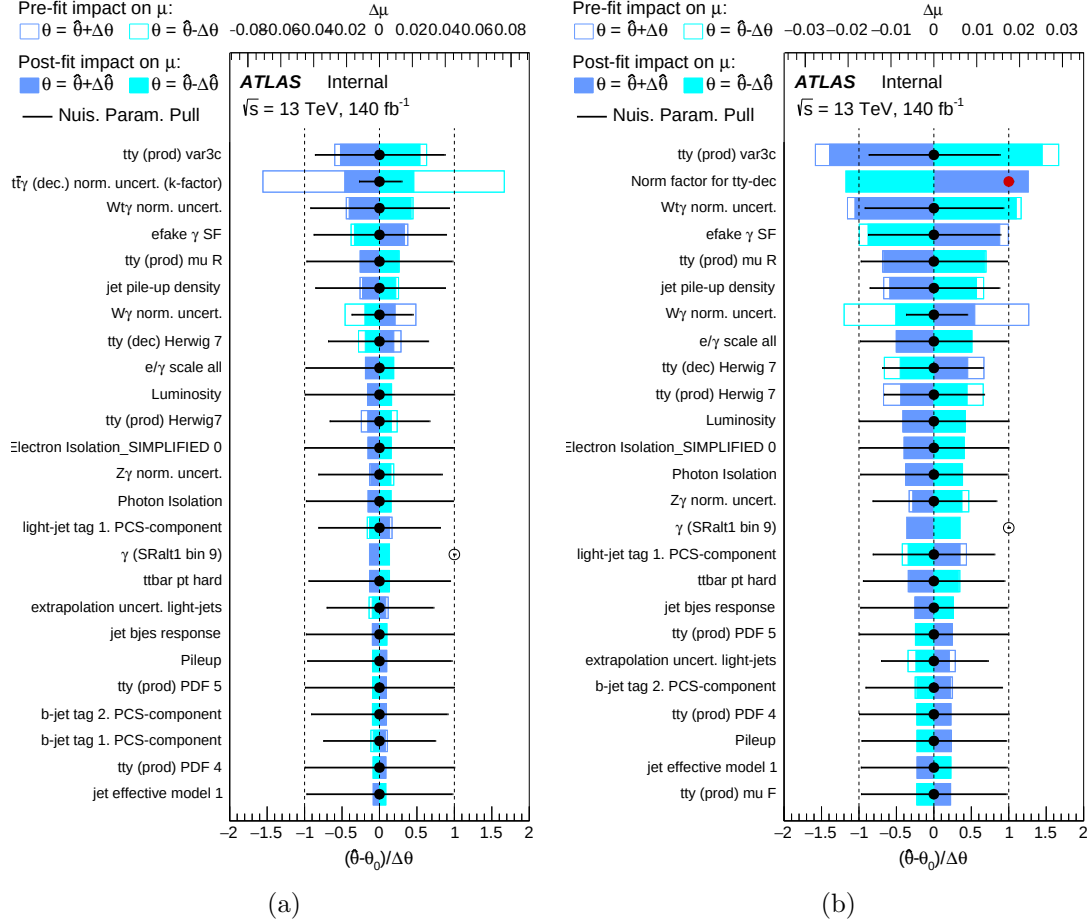


Figure 6.4.: Rankings of the 25 nuisance parameters with the largest impact on the signal strength for the combined Asimov fits in the lepton+jets and dilepton channel. (a) shows the results for the fit setup with one NF, while (b) shows it for the fit setup with two NFs. The blue and turquoise bars show the pre-fit and post-fit importance expressed by $\Delta\mu$, while the black bars and dots show the post-fit values for the nuisance parameters and their uncertainties. More details about the different nuisance parameters are given in Section 5.4 and the corresponding references.

6. Results

6.2.2. Tests for fits to data

From the previous section, it is known that the fit setup seems to be working fine. The Asimov results look consistent between the two fit-setups and between the two channels. The combination gives also sensible results. Hence, the next step was to do the fit to data but keep the final value of the signal strength blinded. This allows one to see how the fit behaves when real data is used and one can investigate possible issues concerning the different nuisance parameters.

Lepton+jets channel

Unblinding the fit revealed a consistent behaviour of the nuisance parameters between the two fit setups. In comparison to the Asimov fits, the constraints are quite similar and only a few pulls are observed (see Appendix E.2). From the pulls, the most important ones to investigate further are the b -tagging uncertainties B0 and B1, as well as the extrapolation uncertainty on light jets. From the modelling uncertainties, the scale and parton shower uncertainties on the $t\bar{t}\gamma$ from production sample are investigated further, as well as the parton shower uncertainty on the $t\bar{t}\gamma$ from decay sample. For the fit with one NF the uncertainty on the $t\bar{t}\gamma$ from decay normalisation is checked too. These uncertainties were chosen because they also appear in the ranking to be among the most important uncertainties. Other uncertainties with moderate pulls were therefore not checked more deeply. The rankings (see Figure 6.5) look similar to the rankings from the Asimov fits (see Figure 6.1). Except for the pulls of the extrapolation uncertainty on light jets and the shower uncertainty for $t\bar{t}\gamma$ from production, all other investigated pulls are moderate and less than half a σ . The changes seen in the correlation matrices are also moderate with respect to the Asimov fit setup (see Appendix E.2).

To investigate the uncertainties mentioned above, two studies were done. One study was to decorrelate each uncertainty concerning its shape and normalisation effect. The second check was to decorrelate them between the four regions. For the two b -tagging uncertainties under investigation, the decorrelation between shape and normalisation effect shows that the pull comes from the shape effect. Decorrelating between the regions shows that the pulls are strongest in the two control regions that use the PCBT distributions as input, which is reasonable. In both cases, for the renormalisation and factorisation scale uncertainties, the pull is driven by the impact of the uncertainties on the shape of the distributions. Because both uncertainties act on the $t\bar{t}\gamma$ from production sample, their impact is largest in the signal region. Again this is a reasonable result. The pull on the parton shower uncertainty for $t\bar{t}\gamma$ from decay comes from both effects, shape and normalisation. Decorrelating between the regions shows that the pull is driven by the $t\bar{t}\gamma$ from decay control region, but also by the other photon control region, while the fakes control region and the signal region show a pull in the opposite direction. From the corresponding plots, showing the impact of the uncertainty (see Figures 6.6 and 6.7), one can see that the negative pull preferred by the fit corresponds to the up-variation, which is obtained by interfacing the nominal simulation with Herwig 7 (while the down-variation is the symmetrised up-variation). This indicates that the Herwig 7

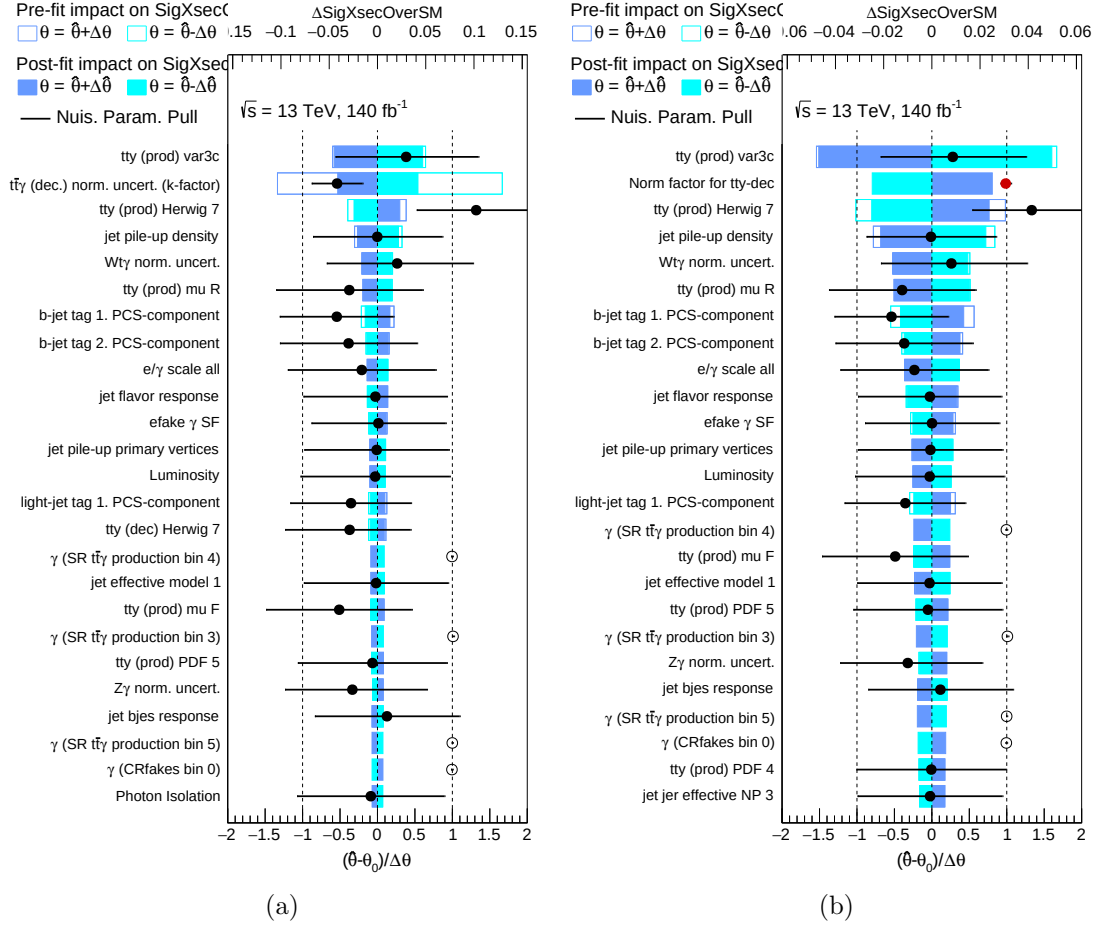


Figure 6.5.: Rankings of the 25 nuisance parameters with the largest impact on the signal strength for the fits to data in the lepton+jets channel. (a) shows the results for the fit setup with one NF, while (b) shows it for the fit setup with two NFs. The blue and turquoise bars show the pre-fit and post-fit importance expressed by $\Delta\mu$, while the black bars and dots show the post-fit values for the nuisance parameters and their uncertainties. More details about the different nuisance parameters are given in Section 5.4 and the corresponding references.

6. Results

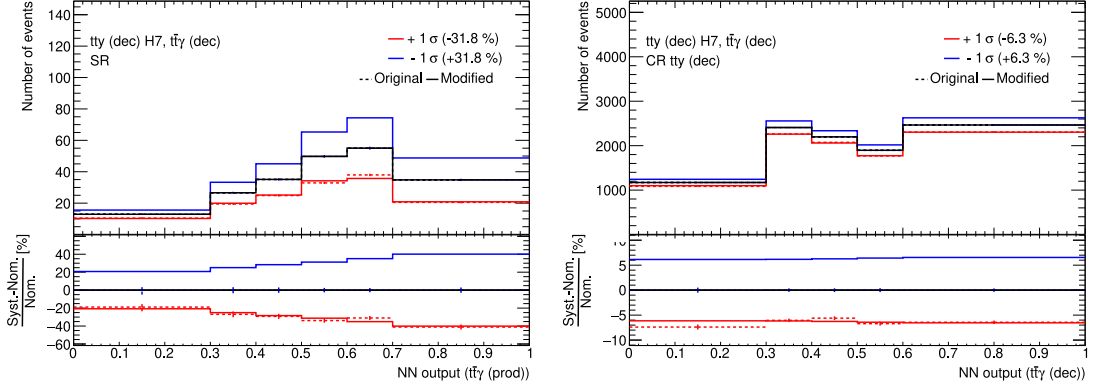


Figure 6.6.: The distributions show the size of the parton shower uncertainty in the lepton+jets channel for the $t\bar{t}\gamma$ (dec.) contribution in the signal region (left) and the $t\bar{t}\gamma$ decay control region (right). The black line shows the nominal expectation, while the red and blue lines show the up and down variations.

prediction is closer to data than the nominal prediction interfaced with Pythia. As one can expect, the pull and constraint on the normalisation uncertainty on the $t\bar{t}\gamma$ from decay sample is driven by the $t\bar{t}\gamma$ from decay control region. No decorrelation into shape and normalisation component is made, because the uncertainty is defined as an overall flat uncertainty of 20%.

Two uncertainties show large pulls of about 1σ (extrapolation uncertainty on light jets) and 1.3σ (parton shower uncertainty for $t\bar{t}\gamma$ from production). The pull for the extrapolation uncertainty on light jets is driven by its shape component and is strongest in the signal region. This behaviour is supported by the plots showing the impact of this uncertainty for the four regions (see Figures 6.8 and 6.9). While the impact is almost negligible in the control region for $t\bar{t}\gamma$ from decay, for the two other control regions a slight slope is seen, which is even much stronger in the signal region. This slope in the signal region drives the pull. The largest pull is observed for the parton shower uncertainty on the $t\bar{t}\gamma$ from production sample. The decorrelation into regions shows that the main impact comes from the signal region, although all regions show a pull in the same direction. That this pull is a shape effect shows the decorrelation into shape and normalisation components. Hence, the large pull is mainly explained by the strong slope seen in the SR, while the other regions also show clear slopes (see Figures 6.10 and 6.11) but are less important for the $t\bar{t}\gamma$ from production contribution, because the control regions are not dominated by it.

With these results, the sources of the relevant pulls have become clear and are understandable. As only two nuisance parameters show pulls around or above 1σ in a rather complicated fit setup no need to change anything on the nuisance parameter model used for the final fit has become visible.

6.2. Tests of the fit setup

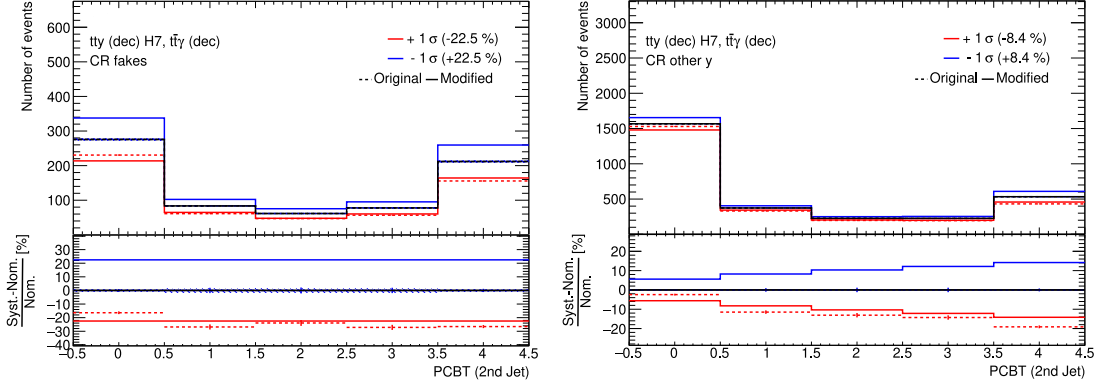


Figure 6.7.: The distributions show the size of the parton shower uncertainty in the lepton+jets channel for the $t\bar{t}\gamma$ (dec.) contribution in the control region for fakes (left) and the other γ control region (right). The black line shows the nominal expectation, while the red and blue lines show the up and down variations.

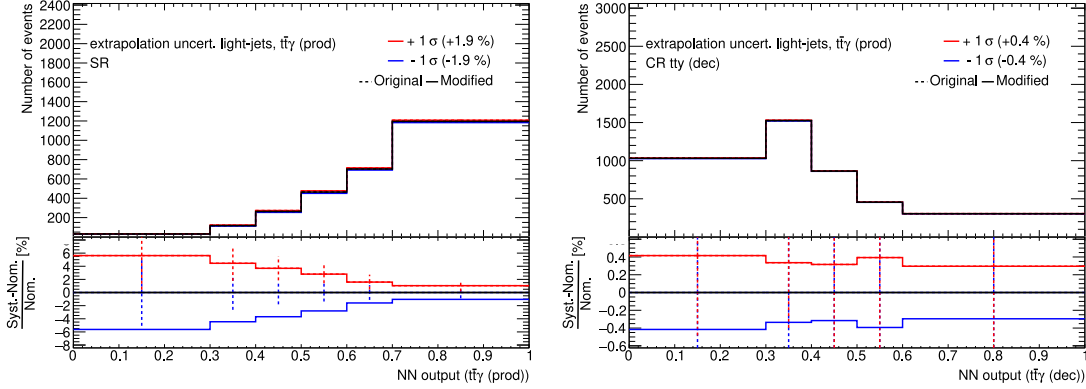


Figure 6.8.: The distributions show the size of the extrapolation uncertainty on light jets in the lepton+jets channel for the $t\bar{t}\gamma$ from production contribution in the signal region (left) and the $t\bar{t}\gamma$ decay control region (right). The black line shows the nominal expectation, while the red and blue lines show the up and down variations.

6. Results

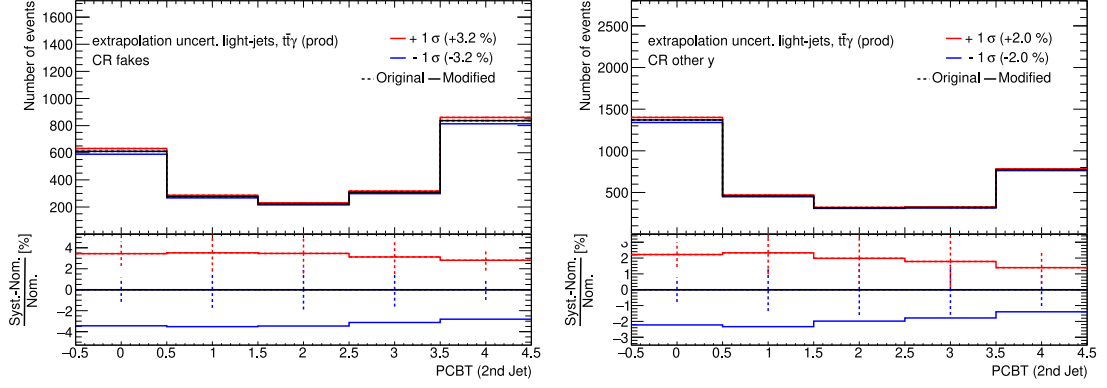


Figure 6.9.: The distributions show the size of the extrapolation uncertainty on light jets in the lepton+jets channel for the $t\bar{t}\gamma$ from production contribution in the control region for fakes (left) and the other γ control region (right). The black line shows the nominal expectation, while the red and blue lines show the up and down variations.

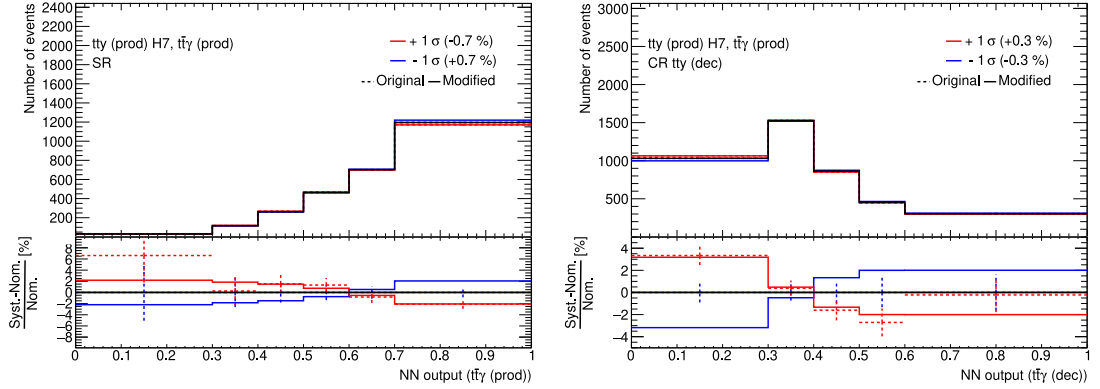


Figure 6.10.: The distributions show the size of the parton shower uncertainty in the lepton+jets channel for the $t\bar{t}\gamma$ from production contribution in the signal region (left) and the $t\bar{t}\gamma$ decay control region (right). The black line shows the nominal expectation, while the red and blue lines show the up and down variations.

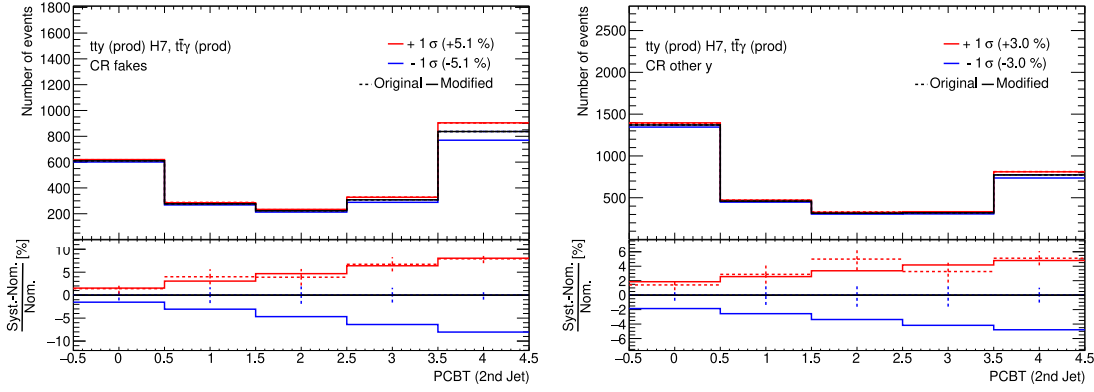


Figure 6.11.: The distributions show the size of the parton shower uncertainty in the lepton+jets channel for the $t\bar{t}\gamma$ from production contribution in the control region for fakes (left) and the other γ control region (right). The black line shows the nominal expectation, while the red and blue lines show the up and down variations.

Dilepton channel

In the dilepton channel, the fits to data reveal good results. Almost no pulls are seen. The visible pulls come from the following uncertainties: uncertainty on $h \rightarrow \gamma$ fake scale factor, uncertainty on pile-up reweighting scale factor, extrapolation uncertainty for light jets and the parton shower uncertainties on both $t\bar{t}\gamma$ samples. None of these pulls goes beyond 0.5σ and the constraints look similar to the Asimov fits (see Appendix E.2). The ranking of the nuisance parameters is again very similar to the Asimov fits (compare Figure 6.2 with Figure 6.12) and correlations look as expected from Asimov fits. For completeness they are shown in Appendix E.2. So in general the results for the dilepton channel look good. No indication of a problematic behaviour of any of the nuisance parameters could be detected. Therefore no further investigations are done.

In Section 5.3.3 it is checked that the bump in the data/MC ratio in the neural network output distribution is a statistical fluctuation caused by two input variables and the way how MC events are distributed into the five folds used for cross-validation. Despite having studied that the neural network behaves correctly, still, the impact of the bump on the final result needs to be checked. Therefore the following procedure was used in the fit with two normalisation factors (no check was done for the fit with one NF), one for each of the two $t\bar{t}\gamma$ samples: the blinded output values for the normalisation factors were stored in hexadecimal values (to keep them hidden but usable for calculations), whereas the uncertainties on the normalisation factors were unblinded. With such a setup several fits were made, where different bins of the fit distribution (see Figure 5.34) were removed from the fit. The results (hexadecimal values) were used to calculate the difference of the normalisation factors $\mu_{\text{nom.}} - \mu_{\text{alt.}}$ concerning the default fit setup with all bins included. $\mu_{\text{nom.}}$ is the signal strength of the nominal fit setup and $\mu_{\text{alt.}}$ the signal

6. Results

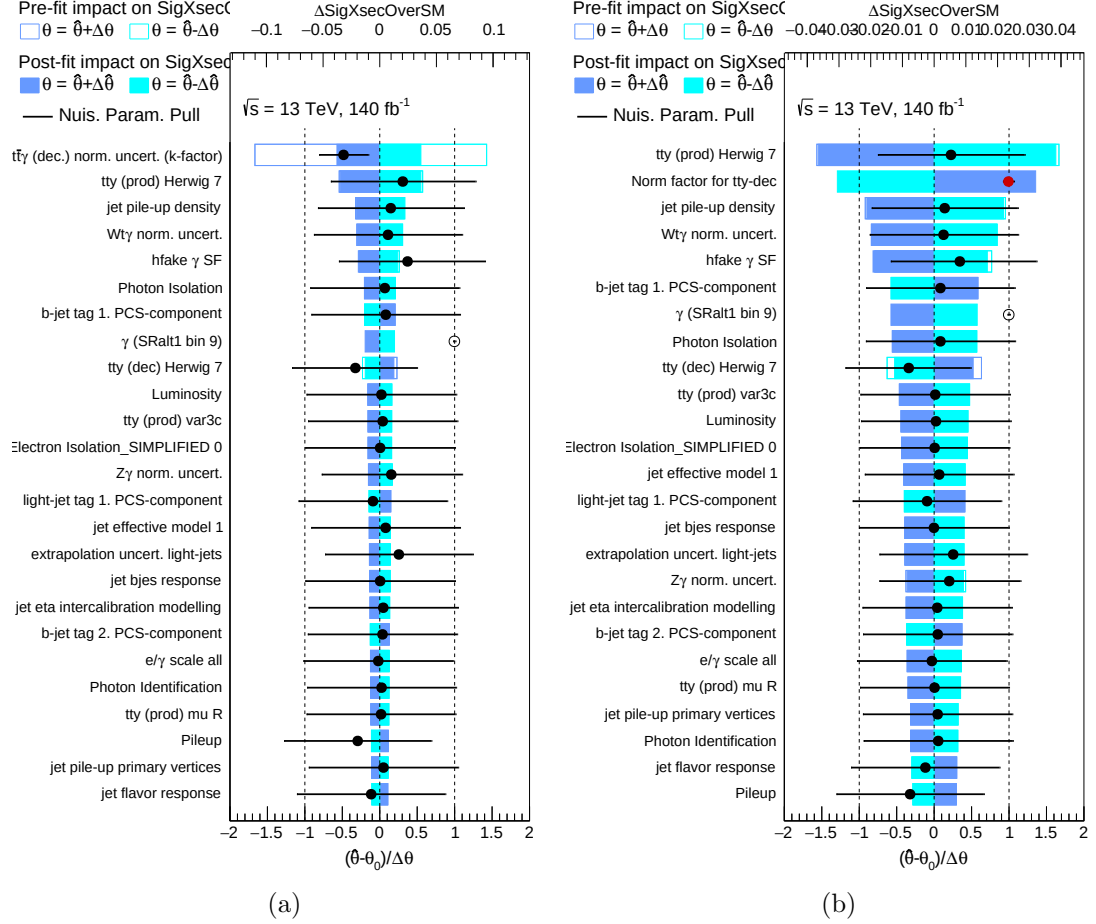


Figure 6.12.: Rankings of the 25 nuisance parameters with the largest impact on the signal strength for the fits to data in the dilepton channel. (a) shows the results for the fit setup with one NF, while (b) shows it for the fit setup with two NFs. The blue and turquoise bars show the pre-fit and post-fit importance expressed by $\Delta\mu$, while the black bars and dots show the post-fit values for the nuisance parameters and their uncertainties. More details about the different nuisance parameters are given in Section 5.4 and the corresponding references.

6.2. Tests of the fit setup

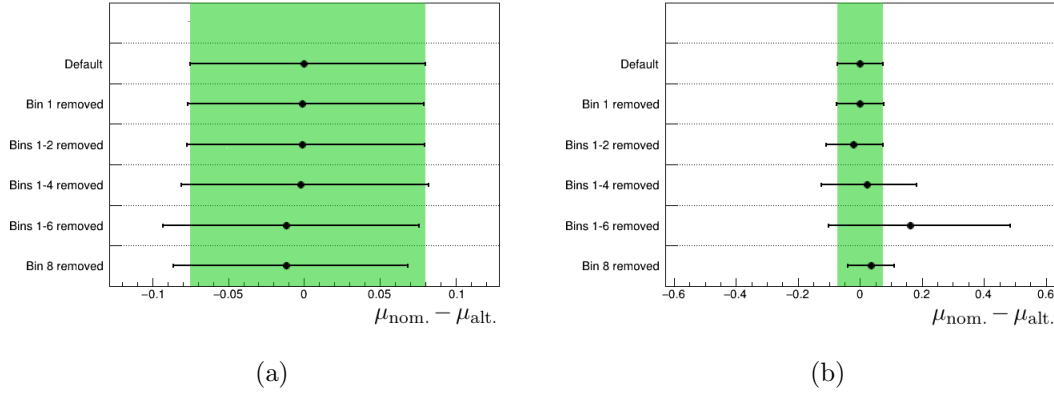


Figure 6.13.: Both plots show the difference of the normalisation factors $\mu_{\text{nom.}} - \mu_{\text{alt.}}$ between the nominal fit setup and fit setups with certain bins being removed from the fit, as well as the corresponding uncertainties. The left one shows the results for the normalisation factor for the $t\bar{t}\gamma$ from production sample and the right plot shows the results for the normalisation factor for the $t\bar{t}\gamma$ from decay sample.

strength in the alternative fit setups. The following configurations were used: default (with all bins), bin 1 is removed, bins 1 and 2 are removed, bins 1 to 4 are removed, bins 1 to 6 are removed and bin 8 (the bin with the bump) is removed. The idea of removing bins from the left is, to see if that makes the bin with the bump (bin 8) more important and changes the final results, which would be visible in a larger or smaller $\Delta\mu$. Bin 8 was removed to see if that bin in general has an impact on the measurement. The results for both normalisation factors are shown in Figure 6.13. The pulls and constraints between the different setups do not vary much and therefore show that the results on the normalisation factors are stable. It can be seen that none of the different fit setups has a significant impact on the normalisation factor for $t\bar{t}\gamma$ from production. That this is the case for the fits where the bins dominated by $t\bar{t}\gamma$ from decay are removed can be expected, but it also shows that the bin has no special impact on the measurement. Removing that bin itself also has no significant impact on the central value for the normalisation factor for $t\bar{t}\gamma$ from production, but also on the uncertainties no impact is observed. This proves that the bump has no special impact on the measurement. The central values for the normalisation factor for $t\bar{t}\gamma$ from decay do not differ much for removing up to the first four bins but the uncertainties get much larger, which can be expected due to the reduced statistics. Removing six bins makes the result on the normalisation factor for $t\bar{t}\gamma$ from decay useless, as the uncertainty grows so much that the shift in the central value does not become significant. The most important fit setup where the bin with the bump is removed shows again consistent results concerning the default fit for both, central value and uncertainties. Hence, the conclusion is that the default fit procedure is reliable and the bump does not bias the final result.

6. Results

Combination

Unblinding the combination of dilepton and lepton+jets channels in both fit setups behaves mostly as expected. The pulls and constraints are mostly similar to the pulls and constraints seen in the individual fits in the lepton+jets channel. The rankings, which are consistent with the Asimov results, are shown in Figure 6.14. Nevertheless, two observations are made: first, the pull on the $var3c$ uncertainty for the $t\bar{t}\gamma$ from production sample is pulled much stronger than in the lepton+jets fit. In the dilepton fit it was not pulled at all. Second, the parton shower uncertainty is less pulled than in the lepton+jets fit. Both uncertainties are highly correlated⁵ and so their interplay was investigated for the fit setup with two NF. To understand the larger pull in the combination, the $var3c$ uncertainty was decorrelated into the different regions. This showed that the main impact on the pull comes from the signal region in the lepton+jets channels. Decorrelating shape and normalisation effect of the nuisance parameter showed that the pull is a normalisation effect, which becomes clear when checking the impact plots of the systematic uncertainty for the different regions (see Figures 6.15 and 6.16). The shape of the uncertainty is rather flat and therefore the pull is related to the impact this uncertainty has on the normalisation.

These decorrelation studies help to understand the origin of the pull but still, the interplay of it with the parton shower uncertainty on the $t\bar{t}\gamma$ from production sample must be investigated. In a first step, this was done by removing one of the two uncertainties from the fits completely. The result was that in case the $var3c$ uncertainty is removed the parton shower uncertainty is pulled a little less in the combination, while in the other case, besides a pull reduction of the $var3c$ uncertainty, larger pulls for the first two uncertainties on the b -tagging are observed. This can be explained by the larger correlations these important uncertainties have concerning the parton shower uncertainty. To further study the impact of the combination the parton shower uncertainty was decorrelated between the two channels. This leads to pulls similar to those seen in the fits for the individual channels but larger uncertainties on the NF for $t\bar{t}\gamma$ from production, which rise from approximately 5.2% to 6%. The correlations also show that they are dominated strongly by the high correlations showing up between the $var3c$ uncertainty and the parton shower uncertainty in the lepton+jets channel, which was not present in the single channel fits. Doing so-called $2\text{-}\mu$ fits (each channel has its normalisation factor defined) in both cases (decorrelating the parton shower uncertainty between the two channels or not) gives compatible results with the fits where a combined normalisation factor is defined. Therefore, a χ^2 test between the default and the $2\text{-}\mu$ fits is done, showing a better compatibility of almost 77% in the case of the decorrelation. In case of not decorrelating the parton shower uncertainty between the channels, the test gives a compatibility of 35%. Also, no significant modelling differences in the basic objects are seen between the two channels.

The results of the studies led to the decision to keep the setup unchanged. Although decorrelating reduces the pulls slightly it does not release any constraints as there were

⁵Correlation matrices are shown for completeness in Appendix E.2.

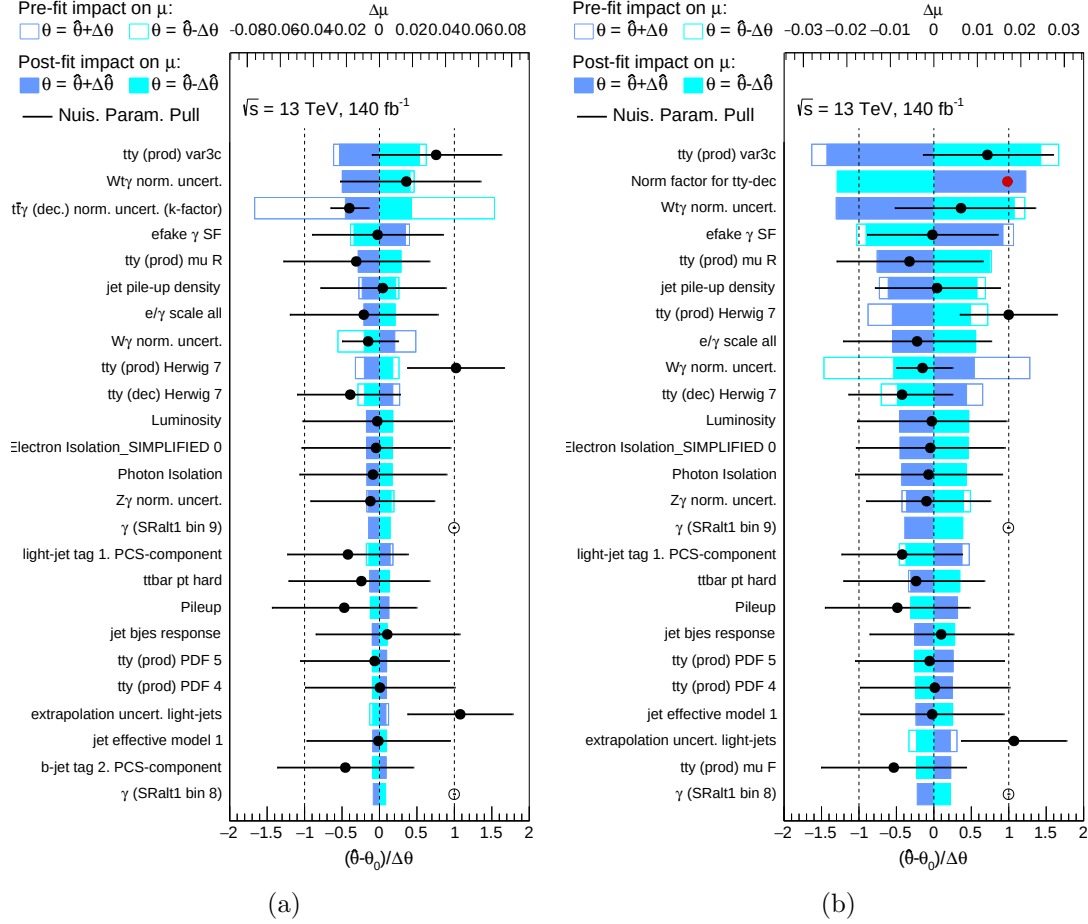


Figure 6.14.: Rankings of the 25 nuisance parameters with the largest impact on the signal strength for the fits to data in the combination of lepton+jets and dilepton channel. (a) shows the results for the fit setup with one NF, while (b) shows it for the fit setup with two NFs. The blue and turquoise bars show the pre-fit and post-fit importance expressed by $\Delta\mu$, while the black bars and dots show the post-fit values for the nuisance parameters and their uncertainties. More details about the different nuisance parameters are given in Section 5.4 and the corresponding references.

6. Results

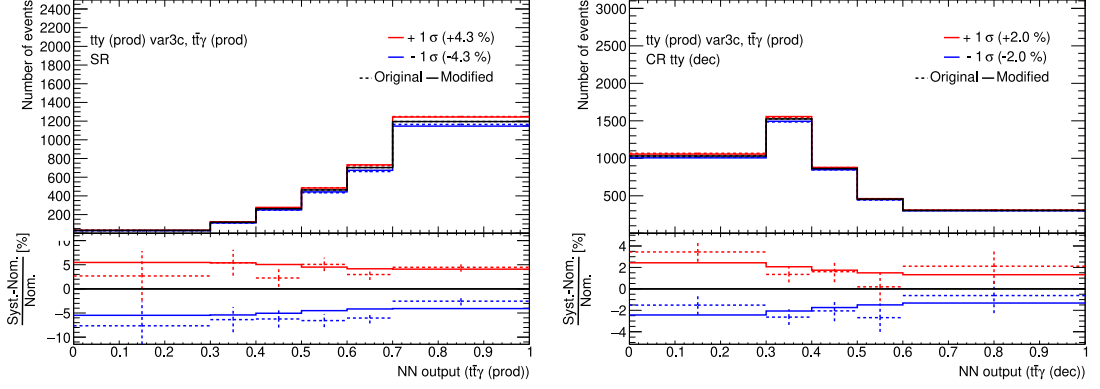


Figure 6.15.: The impact of the $var3c$ uncertainty in the lepton+jets channel for the $t\bar{t}\gamma$ from production contribution in the signal region (left) and the $t\bar{t}\gamma$ decay control region (right).

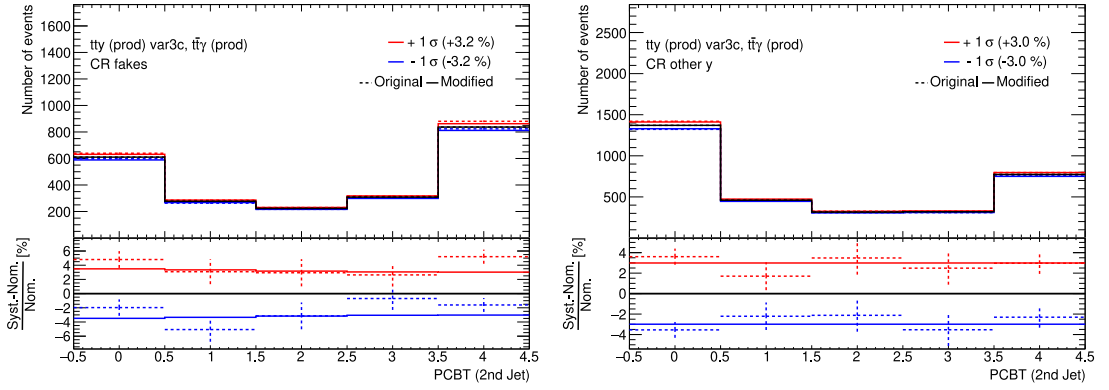


Figure 6.16.: The impact of the $var3c$ uncertainty in the lepton+jets channel for the $t\bar{t}\gamma$ from production contribution in the control region for fakes (left) and the other γ control region (right).

practically none. The correlations seen are not unphysical as both uncertainties do affect the parton shower. In one case the hadronisation model is different and in the other case, the rate of additional radiation is changed. The fact that the $2\text{-}\mu$ fits show better agreement for the decorrelated setup can be explained by the larger uncertainties on the normalisation factors observed. And in addition to these reasons, there is no physics argument in favour of the decorrelation, as both channels are not too different in their modelling.

6.3. Final fit and inclusive cross-section extraction

After unblinding the fits, while keeping the normalisation factors blinded, different pulls of the nuisance parameters were investigated more closely. In the end, no changes were made to the final fit setup. Hence, the rankings, pull plots and correlation matrices from the previous section are also valid for the final fit setup, where the results are presented in this section. The final values for the signal strength (the normalisation factor for $t\bar{t}\gamma$ from production) are shown in Figure 6.17 for the fit with two normalisation factors. The results on the second normalisation factor for $t\bar{t}\gamma$ from decay are shown in Figure 6.18. The results for the fit with one normalisation factor (combined measurement of both $t\bar{t}\gamma$ contributions) are shown in Figure 6.19. In both cases the combined result is lower than the two results for the individual channels. The reason for this are correlations among the different nuisance parameters, which leads to different pulls and constraints, which eventually make up the difference. The corresponding post-fit plots for fits with one or two NF in the individual channels for the regions used in the fits are shown in Figures 6.20 and 6.21. They all show that the fit can describe the data well. Further post-fit plots are shown in Appendix E.3. The grouped uncertainties⁶ are shown in Table 6.3 for the signal strength for the fit with two normalisation factors. The grouped uncertainties for the fit setup with one normalisation factor are shown in Table 6.4. The dominant uncertainties are those related to the modelling of $t\bar{t}\gamma$ from production. In the case of the fit with two normalisation factors, the second factor and its uncertainty have also quite an impact on the total uncertainty, while in the fit with one normalisation factor, this impact comes from the normalisation uncertainty for $t\bar{t}\gamma$ from decay. Further groups that contribute largely are jet and b -tagging uncertainties. In the combination, the uncertainty reduction comes mostly from reductions on the uncertainties of exactly the aforementioned groups.

From the measured values of the normalisation factors the cross-sections are derived. For that Equation 6.7 is used. The uncertainty on the calculated values is given as

$$\Delta_{\sigma} = \Delta_{\mu}\sigma_{\text{pred.}} , \quad (6.14)$$

⁶For the final results the grouping was slightly changed with respect to the tests shown in the previous section. While the extrapolation uncertainties are added to the b -tagging group, the other experimental group was split into pile-up, luminosity and lepton-related uncertainties. The uncertainty on the JVT scale factor was moved to jet-related uncertainties.

6. Results

where Δ_μ is the uncertainty on the corresponding normalisation factor. These are the final results for the lepton+jets channel:

$$\sigma_{t\bar{t}\gamma}(\text{prod.}) = 288_{-19}^{+21} \text{ fb} = 288_{-19}^{+20}(\text{syst.}) \pm 5(\text{stat.}) \text{ fb}, \quad (6.15)$$

$$\sigma_{t\bar{t}\gamma}(\text{prod.}+\text{dec.}) = 704_{-46}^{+49} \text{ fb} = 704_{-46}^{+49}(\text{syst.}) \pm 5(\text{stat.}) \text{ fb}. \quad (6.16)$$

For the dilepton channel, the following values are obtained:

$$\sigma_{t\bar{t}\gamma}(\text{prod.}) = 45.7_{-3.1}^{+3.3} \text{ fb} = 45.7_{-2.8}^{+3.0}(\text{syst.})_{-1.3}^{+1.4}(\text{stat.}) \text{ fb}, \quad (6.17)$$

$$\sigma_{t\bar{t}\gamma}(\text{prod.}+\text{dec.}) = 116.1_{-7.7}^{+8.2} \text{ fb} = 116.1_{-7.6}^{+8.0}(\text{syst.}) \pm 1.7(\text{stat.}) \text{ fb}. \quad (6.18)$$

And finally the results for the combination of the two channels:

$$\sigma_{t\bar{t}\gamma}(\text{prod.}) = 319 \pm 15 \text{ fb} = 319_{-14}^{+15}(\text{syst.}) \pm 4(\text{stat.}) \text{ fb}, \quad (6.19)$$

$$\sigma_{t\bar{t}\gamma}(\text{prod.}+\text{dec.}) = 788_{-37}^{+38} \text{ fb} = 788_{-37}^{+38}(\text{syst.}) \pm 5(\text{stat.}) \text{ fb}. \quad (6.20)$$

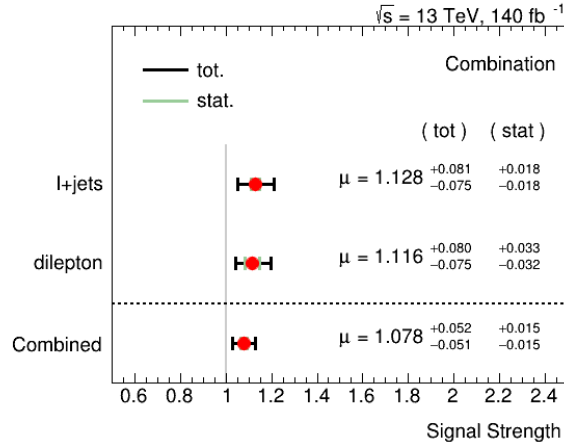


Figure 6.17.: The results for the signal strength in the combination of lepton+jets and dilepton channel in comparison to the individual fits in the setup with two NF, showing total and statistical uncertainties.

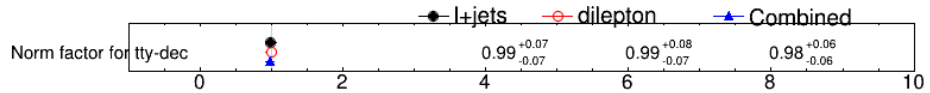


Figure 6.18.: The results for the best-fit value for the $t\bar{t}\gamma$ from decay normalisation factor in the combination of lepton+jets and dilepton channel in comparison to the individual fits for the setup with two NF.

6.3. Final fit and inclusive cross-section extraction

Table 6.3.: Summary of the impact of the systematic uncertainties on the signal strength for the unblinded fits with two NFs grouped into different categories. The quoted uncertainties are obtained by repeating the fit with certain sets of nuisance parameters removed from the fit and subtracting in quadrature the resulting uncertainty from the uncertainty from the full fit.

	lepton+jets	dilepton	combination
Total uncertainty (average)	0.078	0.077	0.052
Statistical uncertainty	0.018	0.033	0.015
Total syst. uncertainty	0.076	0.071	0.050
MC statistical uncertainty	0.015	0.015	0.010
Normalisation factor for $t\bar{t}\gamma$ (dec.)	0.024	0.031	0.021
Modelling uncertainties			
$t\bar{t}\gamma$ (prod.) modelling	0.051	0.016	0.030
$t\bar{t}\gamma$ (prod.) showering	0.024	0.037	0.009
$t\bar{t}\gamma$ (dec.) modelling	0.003	0.013	0.008
Background modelling	0.007	0.002	0.005
Prompt background normalisation	0.015	0.020	0.020
Experimental uncertainties			
Jet uncertainties	0.035	0.030	0.017
B-tagging uncertainties	0.026	0.021	0.010
Fake lepton background estimate	0.004	-	0.001
Fake photon background estimate	0.008	0.015	0.016
Photon	0.005	0.015	0.008
Missing transverse energy	0.003	0.004	0.004
Leptons	0.013	0.014	0.013
Pile-up	0.003	0.007	0.005
Luminosity	0.008	0.010	0.008

6. Results

Table 6.4.: Summary of the impact of the systematic uncertainties on the signal strength for the unblinded fits with one NF grouped into different categories. The quoted uncertainties are obtained by repeating the fit with certain sets of nuisance parameters removed from the fit and subtracting in quadrature the resulting uncertainty from the uncertainty from the full fit.

	lepton+jets	dilepton	combination
Total uncertainty (average)	0.076	0.076	0.052
Statistical uncertainty	0.008	0.016	0.007
Total syst. uncertainty	0.075	0.074	0.051
MC statistical uncertainty	0.015	0.015	0.010
Modelling uncertainties			
$t\bar{t}\gamma$ (prod.) modelling	0.049	0.015	0.029
$t\bar{t}\gamma$ (prod.) showering	0.023	0.035	0.009
$t\bar{t}\gamma$ (dec.) modelling	0.006	0.013	0.008
$t\bar{t}\gamma$ (dec.) normalisation	0.041	0.037	0.020
Background modelling	0.007	0.002	0.005
Prompt background normalisation	0.015	0.021	0.020
Experimental uncertainties			
Jet uncertainties	0.034	0.029	0.017
B-tagging uncertainties	0.026	0.021	0.011
Fake lepton background estimate	0.004	-	0.001
Fake photon background estimate	0.008	0.014	0.016
Photon	0.005	0.016	0.008
Missing transverse energy	0.003	0.004	0.004
Leptons	0.012	0.014	0.012
Pile-up	0.004	0.007	0.005
Luminosity	0.008	0.010	0.008

6.3. Final fit and inclusive cross-section extraction

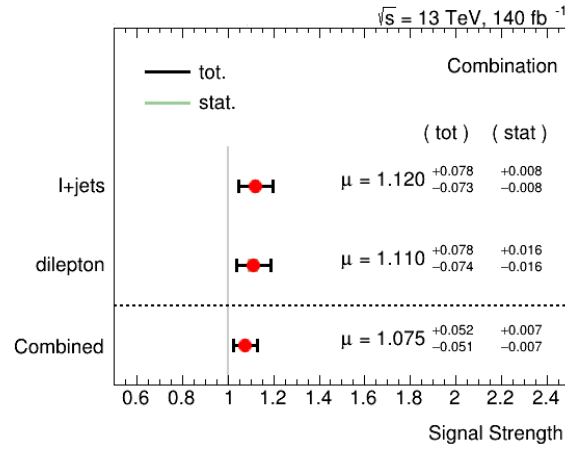


Figure 6.19.: The results for the signal strength in the combination of lepton+jets and dilepton channel in comparison to the individual fits in the setup with one NF, showing total and statistical uncertainties.

6. Results

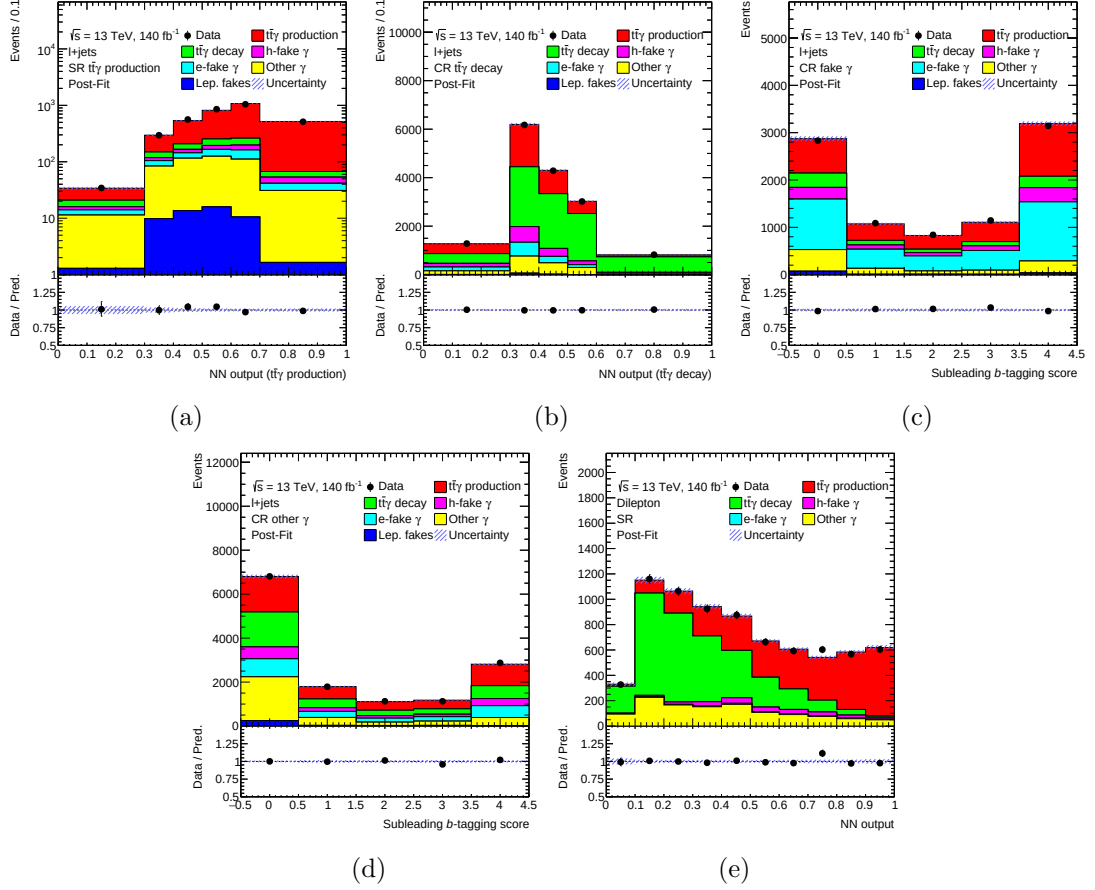


Figure 6.20.: Post fit distributions for the fit with two NF for the fit-regions in the lepton+jets channel ((a)-(d)) and the dilepton channel (e). The fits were made separately in the channels.

6.3. Final fit and inclusive cross-section extraction

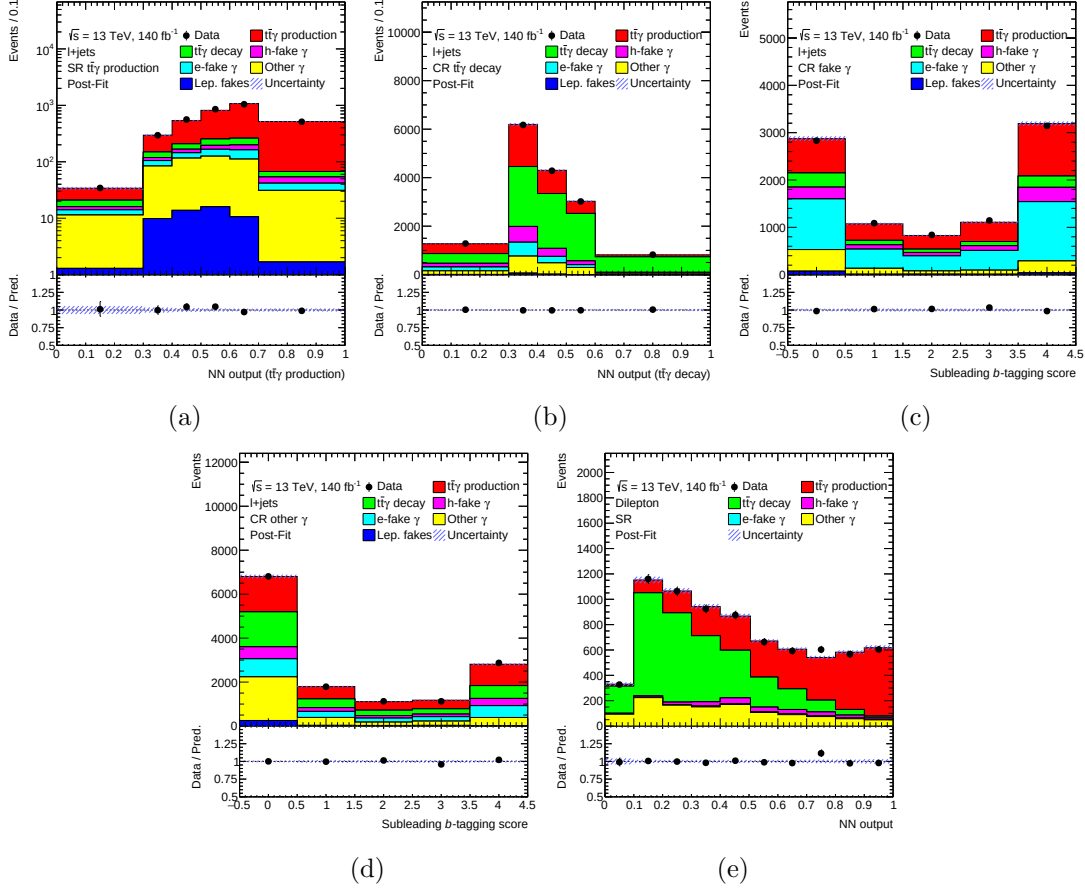


Figure 6.21.: Post fit distributions for the fit with one NF for the fit-regions in the lepton+jets channel ((a)-(d)) and the dilepton channel (e). The fits were made separately in the channels.

6. Results

6.4. Further results

The analysis of the $t\bar{t}\gamma$ process with the full Run 2 data set was not limited to only measuring inclusive cross-sections as presented in the previous section. In addition, differential cross-sections are measured and an EFT interpretation is made. Because the $t\bar{t}\gamma$ and $t\bar{t}Z$ process are both sensitive to electroweak couplings, the same EFT operators are relevant. To make use of this, a combined EFT interpretation is done too. These three measurements, also published in Reference [3], are summarised here to give a full picture of the publication. Further details can be found there.

6.4.1. Differential cross-sections

The top-photon coupling, as presented in Section 2.6.2, does not only predict the frequency with which $t\bar{t}$ pairs appear in association with a photon but also determines the kinematic properties of the photon. If the top-photon coupling is modified this might not change the overall cross-section but only the shape of certain distributions. To be sensitive to such modifications, differential cross-sections are measured. Their measurement is done by applying the method of profile likelihood unfolding [251]. Here the unfolding was done to particle level.

The method of profile likelihood unfolding turns the unfolding into a maximum likelihood fit. In principle, the fit is based on the same likelihood as the inclusive cross-section measurement with the difference that each bin gets an individual normalisation factor for the signal strength. The predictions are given by the respective distributions at the particle level and the application of the response matrix. The response matrix describes how distributions on the particle level get transformed into distributions on the reconstruction level, taking migration, efficiency and acceptance effects into account. As the normalisation stays the same, fitting the folded distributions gives the normalisations at the particle level. Uncertainties are taken into account via dedicated response matrices. For the lepton+jets channel (on the figures in this section it is called single lepton) the same regions are used that are used in the inclusive cross-section measurements. In the dilepton channel, two regions (one control region and one signal region) are created, based on the neural network output. The cut is placed at a neural network output of 0.6. The cut is chosen to minimise the expected uncertainties.

Further optimisation is done for the binning. Two criteria are applied. First, the bins should be wide enough to keep the statistical uncertainties below 10% and fairly similar in each bin. The second criterion is related to the migration effects. Due to the resolution of the detector events might migrate from one bin to another bin when the response matrix is applied to the particle level distribution. To have only small migrations between the bins, they should have at least a width of two times the respective resolution. The bin optimisation concerning the statistical uncertainty already assures small migrations.

A closure test is done to verify that the unfolding procedure reproduces the underlying true distribution. For this, the input shapes of the distributions which are unfolded are altered and the resulting true distributions are compared. The necessary closure is

achieved. Furthermore, stress tests are made. They verify that the unfolding itself is not biased to any specific distribution. Therefore the shape of the distributions is linearly changed and it is checked if the procedure recovers the changed distributions, which it does. Hence, the whole setup for the differential measurements is working correctly. Additionally, this was confirmed by Asimov fits.

Differential distributions are measured for $t\bar{t}\gamma$ from production only and for $t\bar{t}\gamma$ from production and decay. Hence, the fit setups from the inclusive cross-section measurements are repeated. The combination of both channels is only done for $t\bar{t}\gamma$ from production and there only for the photon η and p_T distributions. In general, the following variables are chosen to be measured:

- Both channels: $p_T(\gamma)$, $|\eta|(\gamma)$, $\Delta R_{min}(\gamma, l)$, $\Delta R_{min}(\gamma, b)$, $\Delta R_{min}(l, j)$, p_T (jet with highest p_T),
- Only dilepton channel: $\Delta R(\gamma, l_1)$, $\Delta R(\gamma, l_2)$, $|\Delta\eta(l, l)|$, $\Delta\phi(l, l)$, $p_T(l, l)$.

All of these variables have a physics motivation which is outlined in the following. While the p_T of the photon as a kinematic property is sensitive to the strength of top-photon coupling, hence the charge and the electromagnetic coupling, the angular variables are especially sensitive to the structure of the coupling. The jet p_T is not directly sensitive to the top-photon coupling but is sensitive to the energy scale of the process. In the dilepton channel, additional variables are tested. While the angular distances between the photon and the leptons are again sensitive to the coupling structure, the variables concerning both leptons are sensitive to the spin correlations of the top quarks, which are also altered by the photon emission and the changed fraction of the different initial states in comparison to plain $t\bar{t}$ events.

The final results for $t\bar{t}\gamma$ from production are shown for the absolute and normalised differential cross-sections in Figure 6.22 for the photon p_T . Further results for $t\bar{t}\gamma$ from production are in Appendix F. The complete set of results, including all differential distributions for $t\bar{t}\gamma$ from production and decay, are found in Reference [3]. The uncertainties for the absolute cross-sections range from slightly below 10% up to 30% in some bins of the tails of some distributions for the dilepton channel. In the case of the normalised cross-sections, the uncertainties drop in most cases below 10%. This is due to the cancellation of the normalisation uncertainties. The agreement between data and prediction is quantified as a χ^2 per degree of freedom and the corresponding p -values. Summarising them, one can say that the agreement between data and prediction is modest. One general observation is that for the measurements for $t\bar{t}\gamma$ from production the agreement is on average better than for the full $t\bar{t}\gamma$ measurements. This is explained by the fact the $t\bar{t}\gamma$ from decay sample is an LO sample and hence gives a poorer description than the NLO sample for $t\bar{t}\gamma$ from production. The exact values are given in Appendix F.

6. Results

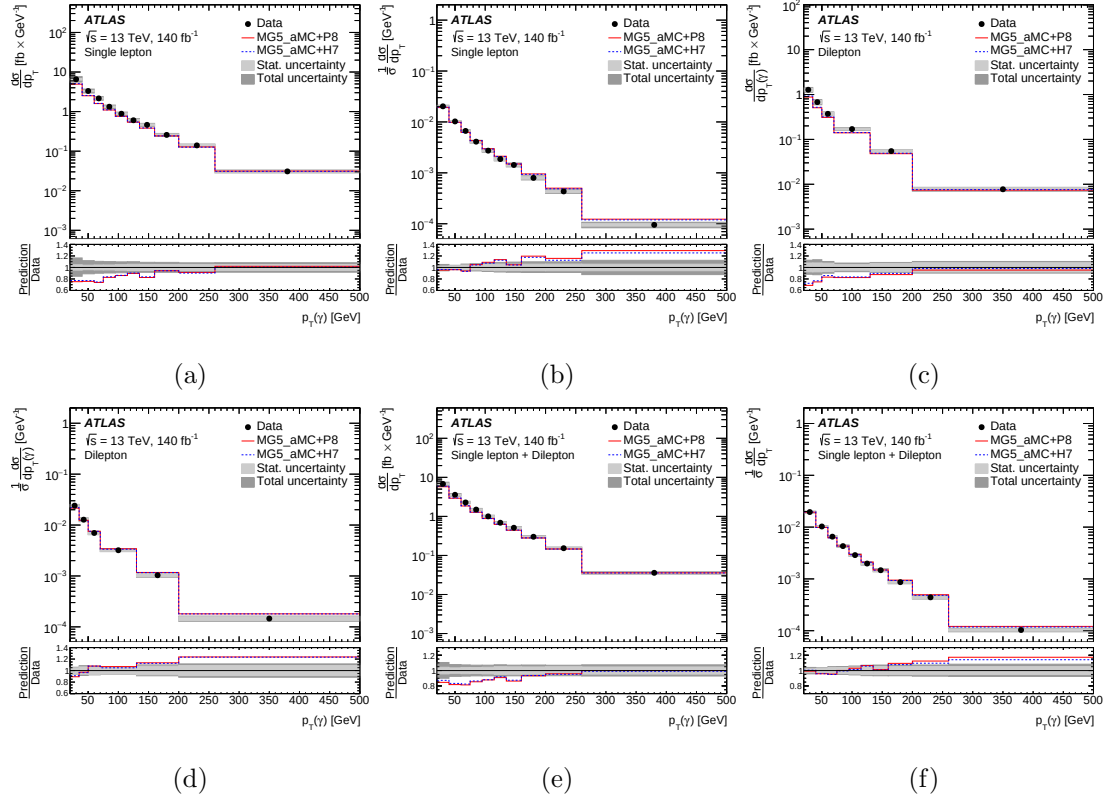


Figure 6.22.: The particle-level unfolded distributions of the photon p_T in the lepton+jets channel ((a) and (b)), dilepton channel ((c) and (d)) and their combination ((e) and (f)) in absolute ((a),(c),(e)) and normalised ((b),(d),(f)) form. The error bars represent both statistical and systematic uncertainties. Overflow events are included in the last bin of each distribution. Plots were taken from [3].

6.4.2. EFT interpretation

Differential distributions unfolded to particle level are a good method to search for possible deviations between the SM predictions and measured data to find hints for new physics. But they only allow to quantify the possible (dis-)agreement. One way to use possible disagreements is to compare them to very specific models. This allows to set relatively precise limits for the parameters of such models. However, the drawback is that these limits are not generalisable. To derive more general, model-independent (and therefore less precise) limits on possible BSM physics the EFT method is a good choice (see Section 2.6.2 and especially Equation 2.52).

Following Reference [172] two operators are relevant to probe the top-photon coupling. These two operators are called O_{tB} and O_{tW} . They modify the weak hypercharge coupling and the weak isospin coupling. Both couplings are the underlying couplings forming

the electromagnetic coupling. The cut-off scale Λ (see Equation 2.52) was set to 1 TeV. The most sensitive distribution to these two operators is the photon p_T distribution. Therefore the EFT interpretation is done based on the differential measurement obtained for the photon p_T for the combined phase space of lepton+jets and dilepton channel. The necessary MC samples are simulated at LO and the nominal EFT sample (SM-like) is scaled to the NLO cross-section prediction for the different bins of the photon p_T distribution for the $t\bar{t}\gamma$ from production sample. The obtained scale factors are then applied to all other EFT samples, which provide predictions for different Wilson coefficients for the named operators. To test the fit setup, injection tests are done. The limits on the operators are obtained with different fit-setups, using a Bayesian statistical approach with EFTfitter [298]. Two sets of fits are done. One fit is done with both operators (real and imaginary parts are taken as independent parameters) simultaneously to obtain the best-fit value for the Wilson coefficients (called global mode). The two-dimensional contours are shown in Figure 6.23. The limits are then derived from the marginalised distributions and can be found in Appendix G. A second set of fits is done where each parameter is fitted independently with a quadratic parametrisation and the corresponding limits are derived. They are in the Appendix G too.

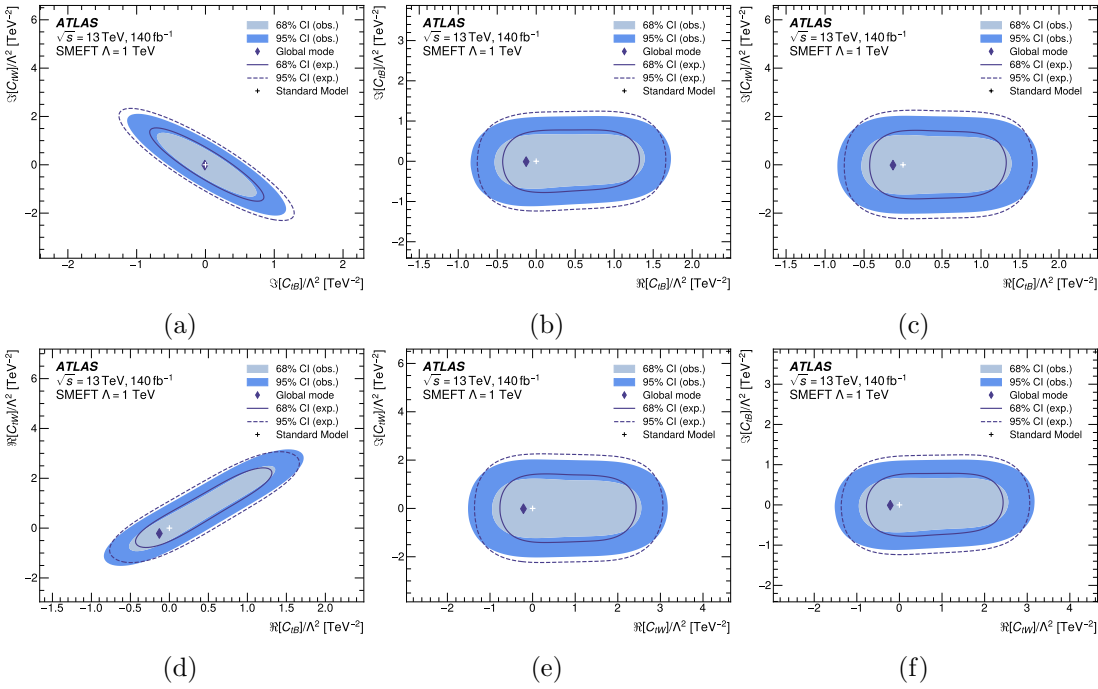


Figure 6.23.: The two-dimensional contours for the real and imaginary part of c_{tB} and c_{tW} Wilson-coefficients derived from the photon p_T distribution from $t\bar{t}\gamma$ from production from the combined measurement in the lepton+jets and dilepton channels. Plots were taken from [3].

6. Results

6.4.3. $t\bar{t}Z$ - $t\bar{t}\gamma$ combination

The connection between the photon and the Z -boson described in the electroweak unification does lead to the situation that both couplings, the top-photon and the top- Z coupling depend on the same underlying couplings of hypercharge and weak isospin. As a consequence, the same EFT operators are probed. In both analyses, this was done with similar fit setups, based on the minimisation of a negative log-likelihood function. Hence, a combination is somehow the logical consequence to exploit the complementarity of both processes, $t\bar{t}\gamma$ and $t\bar{t}Z$.

For a combined EFT-interpretation the photon and Z -boson p_T distributions from both analyses are taken. While for the photon p_T the differential distribution from both channels is used, for the Z -boson p_T distribution the combination of the 3-lepton and 4-lepton channels is used [251]. The treatment of uncertainties was harmonised in advance. So all experimental uncertainties are correlated between the channels, while modelling and background normalisation uncertainties are taken to be uncorrelated. The cuts on the fiducial phase space were set to avoid any overlap and assure statistical independence. Again the same fitting procedure as outlined in the previous Section is applied. The limits from the global fit using marginalised distributions and the limits from the independent fits are in Appendix G. To illustrate the importance of combinations the two dimensional contour plots are shown in Figure 6.24. There the Wilson coefficients c_{tY} and c_{tZ} for the operators $\mathcal{O}_{tY}$ and $\mathcal{O}_{tZ}$ are shown. These operators are obtained as rotations of $\mathcal{O}_{tB}$ and $\mathcal{O}_{tW}$. While for $t\bar{t}Z$ a degenerate structure is seen (as expected) the degeneracy is resolved by adding the results from the $t\bar{t}\gamma$ measurement, which in addition leads to more constrained contours as well.

6.4. Further results

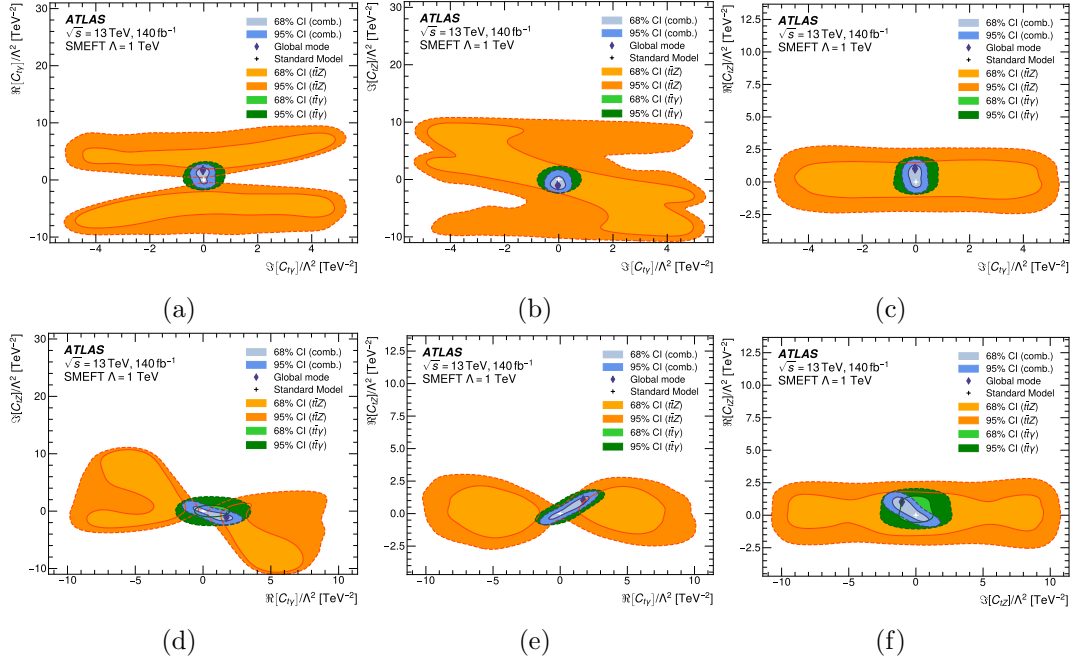


Figure 6.24.: The two-dimensional contours for the c_{tY} and c_{tZ} Wilson-coefficients. The individual contributions from the $t\bar{t}Z$ (orange) and $t\bar{t}\gamma$ (green) measurements are displayed in addition to the combined results (blue). The best-fit value (global mode) and the SM prediction are shown as well. Plots were taken from [3].

This analysis aimed at the investigation of the top-photon coupling and the charge of the top quark. While previous analyses had the same aim, this analysis added new insights due to its design. First, there is the separation between photons coming from the production part of the $t\bar{t}\gamma$ process and those coming from the decay. This approach, based on the narrow-width approximation, was used for the first time in a $t\bar{t}\gamma$ cross-section measurement. The separation was only achievable with sufficient precision due to the usage of feed-forward neural networks and multiclass classification, which allowed to exploit the predicted shape differences of both $t\bar{t}\gamma$ contributions and their differences concerning the different background processes.

In addition to the results of this thesis the full analysis measured a large number of differential distributions and for the first time, an EFT interpretation of the top-photon coupling was performed within a $t\bar{t}\gamma$ ATLAS analysis. Another measurement performed for the first time was the combination of $t\bar{t}\gamma$ and $t\bar{t}Z$ [251], leading to improved constraints on EFT operators compared to the individual measurements.

Although this novel treatment of the $t\bar{t}\gamma$ process leads to new uncertainties, related to the separation, the results are compatible with or even better than similar analyses. In the following, the inclusive cross-section measurements done by the author will be analysed and possible improvements for possible Run 3 analyses will be suggested. Some of the suggestions have a somewhat general character and are relevant for other analyses as well.

Possible improvements for future analyses

As already mentioned, the newly introduced treatment of $t\bar{t}\gamma$ events divided into a production part and a decay part was done to enhance the sensitivity to the top-photon coupling. The feed-forward neural network used for this yielded a good separation between $t\bar{t}\gamma$ from production and $t\bar{t}\gamma$ from decay. This is shown by the high purity of

7. Conclusions

the defined signal region and $t\bar{t}\gamma$ from decay control region in the lepton+jets channel. The drawbacks were the fewer pure fakes and other prompt photon control regions. To further improve this, one might search for variables specific to the different background processes. Such new variables would improve the binary classification in the dilepton channel too. For example, the $h \rightarrow \gamma$ fake events are very similar to real $t\bar{t}\gamma$ events. This comes from the fact that these events mainly come from $t\bar{t}$ events. Therefore the overall kinematic structure of these events is very similar to $t\bar{t}\gamma$ events, especially to $t\bar{t}\gamma$ from decay events. The most useful approach would be to better identify these photon fakes. In the 36 fb^{-1} ATLAS analysis [1] a tool called Prompt Photon Tagger was used for this. It is based on feed-forward neural networks using so-called shower-shape variables¹. Unfortunately, large shape uncertainties had to be assigned to its usage, which diluted the promising results. As most reconstruction and identification algorithms are refined for analyses concerning data from LHC Run 3, the usage of such a tool might be useful already at that stage. Hence, the rejection of photon fakes would improve. This would be a benefit for all analyses involving photons. Furthermore, a better background rejection could be archived with improved b -tagging and the possibilities a larger data set offers. With the data that is foreseen to be collected in Run 3, it might be useful to apply tighter efficiency requirements and/or require more b -tags and still have enough statistics with minimal statistical uncertainties.

The next point concerns the definition of signal and background. This analysis measured fiducial cross-sections where the predictions came from the simulated samples themselves. Better results can be expected from a more inclusive approach, as employed in the $t\bar{t}\gamma$ analysis for $e\mu$ final states [2]. In the full theory calculation, including non-doubly-resonant diagrams, $Wt\gamma$ events were included as signal, allowing for a more consistent signal treatment. Such calculations are much more difficult for lepton+jets events and hence were not available at the time of this thesis. The reason is that much more diagrams need to be taken into account due to the high number of involved quarks. In addition, a coherent treatment in the simulations of the interferences between $t\bar{t}\gamma$ and $Wt\gamma$ would be beneficial. Otherwise, the generally used diagram removal and diagram subtraction scheme would introduce further uncertainties, if one would simply go from LO (as used here) to NLO simulations for the $Wt\gamma$ process. So a simulation similar to the progress known from the $bb4l$ -generator [299] seems necessary. In consequence, $Wt\gamma$ was treated as background in this analysis although the production part of it contains the same information about the top-photon coupling as the $t\bar{t}\gamma$ from production events. Hence, including $Wt\gamma$ as signal should result in more precise measurements if the above-mentioned requirements are fulfilled.

One of the main systematic uncertainties concerns the normalisation of the $t\bar{t}\gamma$ from decay sample, which is a LO sample. A simulation at NLO would further improve the situation.

Furthermore, it is unclear where the approximately 10% difference between data and MC comes from, which results in the measured signal strengths for $t\bar{t}\gamma$ and $t\bar{t}\gamma$ from production. This might be due to even higher order corrections.

¹These variables describe the shower evolution in the calorimeters.

A further possibility Run 3 offers is related to the increased centre-of-mass energy from 13 to 13.6 TeV. This allows for a comparison between Run 2 and Run 3 results and allows for a measurement of the ratio of $q\bar{q}$ initiated events and those events initiated by gluon-gluon fusion. No ISR-photons are expected from events stemming from gluon-gluon fusion and in consequence cross-section ratios might be more sensitive. Such a measurement might also serve as input to constrain parton density functions.

Comparison to other $t\bar{t}\gamma$ measurements at $\sqrt{s} = 13$ TeV

Because previous analyses did not make the splitting into $t\bar{t}\gamma$ from production and $t\bar{t}\gamma$ from decay, only the obtained results for their combined measurements can be compared to previous analyses. In the end, this analysis achieved a precision in the individual channels of approximately 7.5%. This is a similar precision compared to the partial Run 2 analysis, although the statistical uncertainty decreased. The reasons are mainly the uncertainties introduced by the splitting of the $t\bar{t}\gamma$ process into production and decay. The combination of lepton+jets and dilepton channels was able to achieve similar uncertainties than the $e\mu$ analysis [2]. In comparison to the CMS measurements for inclusive $t\bar{t}\gamma$ cross-sections [29, 30], which have relative uncertainties of 6.1% in the lepton+jets channel and 3.9% in the dilepton channel, it looks like a less precise measurement. The problem is that the ATLAS and CMS measurements cannot be directly compared to each other. There are important differences in the way how uncertainties are taken into account and the definitions of the fiducial phase spaces. What can be compared is the measured agreement with the corresponding predictions. While this analysis measured signal strengths of about 10% above the predictions, the CMS dilepton analysis has similar results. In both cases, the theory uncertainty is larger than 10%, where the CMS predictions [30] have much larger uncertainties in comparison to the predicted uncertainties of this analysis (see Chapter 6), as LO-samples were used and this analysis used NLO samples, at least for the $t\bar{t}\gamma$ from production part. Hence both measurements are well in agreement with the SM prediction. In the case of the lepton+jets channel CMS measures lower signal strength in their regions with 4 or more jets than this analysis does. But again the final results lie well within the theory uncertainties, which again are larger for the CMS analysis due to the same reasons as given above (see Reference [29] and Chapter 6). Additionally, all the uncertainties coming from the splitting into $t\bar{t}\gamma$ production and decay part were not relevant for the CMS analyses and they make up quite a large fraction in this analysis. All in all the presented results in Section 6.3 are as precise as possible, taking the newly developed approach into account.

Interpretation of the results

Given the results of this novel analysis, what has been learned about the top-photon coupling and the top quark charge with this analysis?

From the decay of the top quark it is already known that an exotic model with a top quark charge of $-4/3$ [187, 188] is excluded [189–191] up to 8σ [192]. So is there room for a modified top quark charge? In principle, one can take the top-photon coupling

7. Conclusions

and add a continuous parameter c to it. Such a parameter would not spoil the charge conservation in the top quark decay but it would modify the overall rate of $t\gamma$ -events. At LO the cross-section of the $t\bar{t}\gamma$ from production process has a quadratic dependence on the top quark charge. So under the assumption of the narrow width approximation one can calculate the cross-section for $t\bar{t}\gamma$ from production as a $2 \rightarrow 3$ process, either as $q\bar{q} \rightarrow t\bar{t}\gamma$ or $gg \rightarrow t\bar{t}\gamma$. From the ISR in $q\bar{q}$ -annihilation, a constant term is part of the cross-section, so it can be parametrised as

$$\sigma_{t\bar{t}\gamma} = \frac{4c^2}{9}A + B, \quad (7.1)$$

where A is the part that gets changed by the parameter c and B is the part related to ISR and stays constant. Such a model can be implemented in MadGraph quite easily [280]. Doing the calculation with different top quark charges allows one to quantify A and B from a system of linear equations. With the signal strength μ , measured with different setups in Section 6.3, one gets the relation

$$\mu\sigma_{t\bar{t}\gamma, Q=2/3} = \frac{4c^2}{9}A + B, \quad (7.2)$$

assuming that $\sigma_{t\bar{t}\gamma, Q=2/3}$ is the cross-section expected from SM. From this follows that the modified charge in the coupling would be²

$$Q = \frac{2c}{3} = \sqrt{\frac{\mu\sigma_{t\bar{t}\gamma, Q=2/3} - B}{A}}. \quad (7.3)$$

With the result from the combination ($\mu = 1.078_{-0.051}^{+0.052}$) this would yield

$$Q = 0.70 \pm 0.02, \quad (7.4)$$

assuming the uncertainty on the signal strength to be dominant concerning the uncertainty on A and B coming from the integration uncertainty by MadGraph. The deviation of around 1.5σ concerning the SM prediction would presumably vanish when taking theory uncertainties (e.g. scale and PDF uncertainties) into account, but they were not quantified for this simple calculation.

The caveat concerning this calculation is the following: from a QED perspective the introduction of the parameter c might be ad-hoc but unproblematic concerning features such as gauge-invariance and renormalisability. The problem comes when looking at the whole SM. There the photon emerges from a rotation in the space of the gauge field for hyper-charge and the third gauge field for the weak force. The question becomes now, how the parameter c can be implemented into this rotation without any problems concerning the couplings of the particles to the Z -boson? Unfortunately, no definite answer to this question can be given here.

²The relative sign of the square-root has no importance here, as it enters squared any cross-section calculation. It might be relevant in constructing a model consistent with electroweak unification.

Inclusive and differential cross-sections can only reveal possible deviations. For a quantitative interpretation of any deviations either a well-defined BSM model is needed or model-independent EFT operators. Because limits on the Wilson-coefficients of EFT operators can be interpreted in a more general way than limits on concrete models, this analysis made an EFT interpretation. The limits (see Section 6.4.2) show that the measured top-photon coupling is consistent with the SM prediction. Consequently, any significant deviation that might be measurable in the future will be rather small. So the main message of this analysis is that, within current uncertainties, the top quark couples to the photon as predicted by the SM.

Bibliography

- [1] ATLAS Collaboration, *Measurements of inclusive and differential fiducial cross-sections of $t\bar{t}\gamma$ production in leptonic final states at $\sqrt{s} = 13$ TeV in ATLAS*, Eur. Phys. J. C **79** (2019) 382.
- [2] ATLAS Collaboration, *Measurements of inclusive and differential cross-sections of combined $t\bar{t}\gamma$ and $Wt\gamma$ production in the $e\mu$ channel at 13 TeV with the ATLAS detector*, JHEP **9** (2020) 49.
- [3] ATLAS Collaboration, *Measurements of inclusive and differential cross-sections of $t\bar{t}\gamma$ production in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, JHEP **10** (2024) 191.
- [4] S. Weinberg, *The Making of the Standard Model*, Eur. Phys. J. C **34** (2003) 5.
- [5] Particle Data Group Collaboration, R. L. Workman et al., *Review of Particle Physics*, PTEP **2022** (2022) 083C01.
- [6] ATLAS Collaboration, *Standard Model Summary Plots February 2022*, Tech. Rep. ATL-PHYS-PUB-2022-009, CERN, Geneva, 2022.
- [7] G. 't Hooft, *Renormalizable Lagrangians for massive Yang-Mills fields*, Nuc. Phys. B **35** (1971) 167.
- [8] G. 't Hooft and M. Veltman, *Regularization and renormalization of gauge fields*, Nuc. Phys. B **44** (1972) 189.
- [9] G. 't Hooft and M. Veltman, *Combinatorics of gauge fields*, Nuc. Phys. B **50** (1972) 318.
- [10] D. J. Gross and F. Wilczek, *Ultraviolet Behavior of Non-Abelian Gauge Theories*, Phys. Rev. Lett. **30** (1973) 1343.

BIBLIOGRAPHY

- [11] H. D. Politzer, *Reliable Perturbative Results for Strong Interactions?*, Phys. Rev. Lett. **30** (1973) 1346.
- [12] H. D. Politzer, *Asymptotic freedom: An approach to strong interactions*, Phys. Rep. **14** (1974) 129.
- [13] M. Kobayashi and T. Maskawa, *CP-Violation in the Renormalizable Theory of Weak Interaction*, Prog. Theor. Phys. **49** (1973) 652.
- [14] ATLAS Collaboration, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Phys. Lett. B **716** (2012) 1.
- [15] CMS Collaboration, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, Phys. Lett. B **716** (2012) 30.
- [16] S. Perlmutter, et al., *Measurements of Ω and Λ from 42 High-Redshift Supernovae*, Astrophys. J. **517** (1999) 565.
- [17] A. G. Riess, et al., *Observational Evidence from Supernovae for an Accelerating Universe and a Cosmological Constant*, Astron. J. **116** (1998) 1009.
- [18] F. Zwicky, *Die Rotverschiebung von extragalaktischen Nebeln*, Helv. Phys. Acta **6** (1933) 110.
- [19] V. C. Rubin, N. Thonnard, and W. K. Ford, Jr., *Rotational properties of 21 SC galaxies with a large range of luminosities and radii, from NGC 4605 / $R = 4\text{kpc}$ / to UGC 2885 / $R = 122\text{kpc}$ /*, Astrophys. J. **238** (1980) 471.
- [20] A. D. Sakharov, *Violation of CP invariance, C asymmetry, and baryon asymmetry of the universe*, Sov. Phys. Uspekhi **34** (1991) 392.
- [21] S. Dimopoulos and H. Georgi, *Softly broken supersymmetry and SU(5)*, Nuc. Phys. B **193** (1981) 150.
- [22] J. Wess and B. Zumino, *Supergauge transformations in four dimensions*, Nuc. Phys. B **70** (1974) 39.
- [23] L. Randall and R. Sundrum, *Large Mass Hierarchy from a Small Extra Dimension*, Phys. Rev. Lett. **83** (1999) 3370.
- [24] L. Randall and R. Sundrum, *An Alternative to Compactification*, Phys. Rev. Lett. **83** (1999) 4690.
- [25] N. Arkani-Hamed, S. Dimopoulos, and G. Dvali, *The hierarchy problem and new dimensions at a millimeter*, Physics Letters B **429** (1998) 263.

- [26] ATLAS Collaboration, *Observation of top-quark pair production in association with a photon and measurement of the $t\bar{t}\gamma$ production cross section in pp collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector*, Phys. Rev. D **91** (2015) 072007.
- [27] ATLAS Collaboration, *Measurement of the $t\bar{t}\gamma$ production cross section in proton-proton collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector*, JHEP **11** (2017) 86.
- [28] CMS Collaboration, *Measurement of the semileptonic $t\bar{t} + \gamma$ production cross section in pp collisions at $\sqrt{s} = 8$ TeV*, JHEP **10** (2017) 6.
- [29] CMS Collaboration, *Measurement of the inclusive and differential $t\bar{t}\gamma$ cross sections in the single-lepton channel and EFT interpretation at $\sqrt{s} = 13$ TeV*, JHEP **12** (2021) 180.
- [30] CMS Collaboration, *Measurement of the inclusive and differential $t\bar{t}\gamma$ cross sections in the dilepton channel and effective field theory interpretation in proton-proton collisions at $\sqrt{s} = 13$ TeV*, JHEP **5** (2022) 91.
- [31] J. J. Thomson, *Cathode Rays*, Philos. Mag. **44** (1897) 293.
- [32] M. Planck, *Ueber das Gesetz der Energieverteilung im Normalspectrum*, Ann. Phys. **309** (1901) 553.
- [33] H. Geiger and E. Marsden, *On the Scattering of α -Particles by Matter*, Proc. Roy. Soc. Lond. A **82** (1909) 495.
- [34] E. Rutherford, *The Scattering of α and β Particles by Matter and the Structure of the Atom*, Philos. Mag. **21** (1911) 669.
- [35] A. Einstein, *Über einen die Erzeugung und Verwandlung des Lichtes betreffenden heuristischen Gesichtspunkt*, Ann. Phys. **322** (1905) 132.
- [36] N. Bohr, *On the Constitution of Atoms and Molecules, Part I*, Philos. Mag. **26** (1913) 1.
- [37] N. Bohr, *On the Constitution of Atoms and Molecules, Part II, Systems Containing Only a Single Nucleus*, Philos. Mag. **26** (1913) 476.
- [38] N. Bohr, *On the Constitution of Atoms and Molecules, Part III, Systems Containing Several Nuclei*, Philos. Mag. **26** (1913) 857.
- [39] E. Schrödinger, *An Undulatory Theory of the Mechanics of Atoms and Molecules*, Phys. Rev. **28** (1926) 1049.
- [40] W. Heisenberg, *Über quantentheoretische Umdeutung kinematischer und mechanischer Beziehungen*, Z. Physik **33** (1925) 879.

BIBLIOGRAPHY

- [41] M. Born and P. Jordan, *Zur Quantenmechanik*, Z. Physik **34** (1925) 858.
- [42] M. Born, W. Heisenberg, and P. Jordan, *Zur Quantenmechanik II*, Z. Physik **35** (1926) 557.
- [43] C. D. Anderson, *The Positive Electron*, Phys. Rev. **43** (1933) 491.
- [44] P. A. M. Dirac, *The Quantum Theory of the Electron*, Proc. Roy. Soc. Lond. A **117** (1928) 610.
- [45] S. H. Neddermeyer and C. D. Anderson, *Note on the Nature of Cosmic-Ray Particles*, Phys. Rev. **51** (1937) 884.
- [46] E. Rutherford, *LIV. Collision of α particles with light atoms. IV. An anomalous effect in nitrogen*, Philos. Mag. **37** (1919) 581.
- [47] J. Chadwick, *The existence of a neutron*, Proc. R. Soc. Lond. A **136** (1932) 692.
- [48] L. W. Mo and Y. S. Tsai, *Radiative Corrections to Elastic and Inelastic ep and up Scattering*, Rev. Mod. Phys. **41** (1969) 205.
- [49] E. D. Bloom et al., *High-Energy Inelastic $e - p$ Scattering at 6° and 10°* , Phys. Rev. Lett. **23** (1969) 930.
- [50] M. Breidenbach et al., *Observed Behavior of Highly Inelastic Electron-Proton Scattering*, Phys. Rev. Lett. **23** (1969) 935.
- [51] G. Miller et al., *Inelastic Electron-Proton Scattering at Large Momentum Transfers and the Inelastic Structure Functions of the Proton*, Phys. Rev. D **5** (1972) 528.
- [52] A. Bodek et al., *The ratio of deep-inelastic e -n to e -p cross sections in the threshold region*, Phys. Lett. B **51** (1974) 417.
- [53] E. M. Riordan et al., *Extraction of $R = \frac{\sigma_L}{\sigma_T}$ from Deep Inelastic $e - p$ and $e - d$ Cross Sections*, Phys. Rev. Lett. **33** (1974) 561.
- [54] E. Riordan et al., *Tests of scaling of the proton electromagnetic structure functions*, Phys. Lett. B **52** (1974) 249.
- [55] M. Gell-Mann, *Symmetries of Baryons and Mesons*, Phys. Rev. **125** (1962) 1067.
- [56] M. Gell-Mann, *A Schematic Model of Baryons and Mesons*, Phys. Lett. **8** (1964) 214.
- [57] G. Zweig, *An SU_3 model for strong interaction symmetry and its breaking; Version 1*, Tech. Rep. CERN-TH-401, CERN, Geneva, 1964.
- [58] Y. Ne'eman, *Derivation of strong interactions from a gauge invariance*, Nuc. Phys. **26** (1961) 222.

- [59] S. S. Schweber, *QED and the men who made it: Dyson, Feynman, Schwinger and Tomonaga*. Princeton University Press, Princeton, New Jersey, 1994.
- [60] J. Schwinger, ed., *Selected Papers on Quantum Electrodynamics*. Dover Publications Inc., New York, 1958.
- [61] W. Gerlach and O. Stern, *Der experimentelle Nachweis der Richtungsquantelung im Magnetfeld*, Z. Physik **9** (1922) 349.
- [62] F. Englert and R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons*, Phys. Rev. Lett. **13** (1964) 321.
- [63] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, *Global Conservation Laws and Massless Particles*, Phys. Rev. Lett. **13** (1964) 585.
- [64] P. W. Higgs, *Broken symmetries, massless particles and gauge fields*, Phys. Lett. **12** (1964) 132.
- [65] P. W. Higgs, *Broken Symmetries and the Masses of Gauge Bosons*, Phys. Rev. Lett. **13** (1964) 508.
- [66] H. Yukawa, *On the Interaction of Elementary Particles. I*, Proc. Phys.-Math. Soc. Jpn. **17** (1935) 48.
- [67] W. Pauli, *Aufsätze und Vorträge über Physik und Erkenntnistheorie*. Friedr. Vieweg und Sohn, Braunschweig, 1961.
- [68] C. L. Cowan, et al., *Detection of the Free Neutrino: a Confirmation*, Science **124** (1956) 103.
- [69] S. L. Glashow, *Partial-symmetries of weak interactions*, Nuc. Phys. **22** (1961) 579.
- [70] S. Weinberg, *A Model of Leptons*, Phys. Rev. Lett. **19** (1967) 1264.
- [71] A. Salam and J. C. Ward, *Weak and electromagnetic interactions*, Il Nuovo Cimento **11** (1959) 568.
- [72] C. S. Wu et al., *Experimental Test of Parity Conservation in Beta Decay*, Phys. Rev. **105** (1957) 1413.
- [73] M. Goldhaber, L. Grodzins, and A. W. Sunyar, *Helicity of Neutrinos*, Phys. Rev. **109** (1958) 1015.
- [74] J. H. Christenson et al., *Evidence for the 2π Decay of the K_2^0 Meson*, Phys. Rev. Lett. **13** (1964) 138.
- [75] G. T. Danby, et al., *Observation of High-Energy Neutrino Reactions and the Existence of Two Kinds of Neutrinos*, Physical Review Letters **9** (1962) 36.

BIBLIOGRAPHY

- [76] N. Cabibbo, *Unitary Symmetry and Leptonic Decays*, Phys. Rev. Lett. **10** (1963) 531.
- [77] B. Björken and S. Glashow, *Elementary particles and $SU(4)$* , Phys. Lett. **11** (1964) 255.
- [78] S. L. Glashow, J. Iliopoulos, and L. Maiani, *Weak Interactions with Lepton-Hadron Symmetry*, Phys. Rev. D **2** (1970) 1285.
- [79] J. J. Aubert, et al., *Experimental Observation of a Heavy Particle J* , Phys. Rev. Lett. **33** (1974) 1404.
- [80] J. E. Augustin, et al., *Discovery of a Narrow Resonance in e^+e^- Annihilation*, Phys. Rev. Lett. **33** (1974) 1406.
- [81] F. Hasert, et al., *Observation of neutrino-like interactions without muon or electron in the gargamelle neutrino experiment*, Phys. Lett. B **46** (1973) 138.
- [82] F. Hasert, et al., *Search for elastic muon-neutrino electron scattering*, Phys. Lett. B **46** (1973) 121.
- [83] M. L. Perl, et al., *Evidence for Anomalous Lepton Production in $e^+ - e^-$ Annihilation*, Phys. Rev. Lett. **35** (1975) 1489.
- [84] BABAR Collaboration, *Observation of CP Violation in the B^0 Meson System*, Phys. Rev. Lett. **87** (2001) 091801.
- [85] Belle Collaboration, *Observation of Large CP Violation in the Neutral B Meson System*, Phys. Rev. Lett. **87** (2001) 091802.
- [86] LHCb Collaboration, *Observation of CP Violation in Charm Decays*, Phys. Rev. Lett. **122** (2019) 211803.
- [87] S. W. Herb, et al., *Observation of a Dimuon Resonance at 9.5 GeV in 400-GeV Proton-Nucleus Collisions*, Phys. Rev. Lett. **39** (1977) 252.
- [88] CDF Collaboration, *Observation of Top Quark Production in $\bar{p}p$ Collisions with the Collider Detector at Fermilab*, Phys. Rev. Lett. **74** (1995) 2626.
- [89] D0 Collaboration, *Search for High Mass Top Quark Production in $p\bar{p}$ Collisions at $\sqrt{s} = 1.8$ TeV*, Phys. Rev. Lett. **74** (1995) 2422.
- [90] R. Brandelik et al., *Evidence for planar events in e^+e^- annihilation at high energies*, Phys. Lett. B **86** (1979) 243.
- [91] UA1 Collaboration, *Experimental observation of isolated large transverse energy electrons with associated missing energy at $\sqrt{s}=540$ GeV*, Physics Letters B **122** (1983) 103.

- [92] UA2 Collaboration, *Observation of Single Isolated Electrons of High Transverse Momentum in Events with Missing Transverse Energy at the CERN anti-p p Collider*, Phys. Lett. B **122** (1983) 476.
- [93] UA1 Collaboration, *Experimental Observation of Lepton Pairs of Invariant Mass Around 95 GeV/c² at the CERN SPS Collider*, Phys. Lett. B **126** (1983) 398.
- [94] UA2 Collaboration, *Evidence for $Z^0 \rightarrow e^+e^-$ at the CERN $\bar{p}p$ Collider*, Phys. Lett. B **129** (1983) 130.
- [95] K. Kodama et al., *Observation of tau neutrino interactions*, Phys. Lett. B **504** (2001) 218.
- [96] B. Lemmer, *Measurement of Spin Correlations in $t\bar{t}$ Events from pp Collisions at $\sqrt{s} = 7$ TeV in the Lepton + Jets Final State with the ATLAS Detector*. PhD thesis, II.Physik-UniGö-Diss-2014-02, 2014.
- [97] B. Andersson, et al., *Parton fragmentation and string dynamics*, Phys. Rep. **97** (1983) 31–145.
- [98] S. Wolfram, *Parton and Hadron Production in e^+e^- Annihilation*, in *15th Rencontres de Moriond: Part I, High-Energy Hadronic Interactions*. 1980.
- [99] G. Lüders, *On the equivalence of invariance under time reversal and under particle-antiparticle conjugation for relativistic field theories*, Dan. Mat. Fys. Medd. **28** (1954) 1.
- [100] W. Pauli, *Exclusion principle, Lorentz group and reflexion of space-time and charge*, in *Niels Bohr and the Development of Physics*, W. Pauli, L. Rosenfeld, and V. Weisskopf, eds. McGraw-Hill, 1955.
- [101] J. S. Bell, *Time reversal in field theory*, Proc. R. Soc. Lond. A **231** (1955) 479.
- [102] J. D. Bekenstein, *Relativistic gravitation theory for the modified Newtonian dynamics paradigm*, Phys. Rev. D **70** (2004) 083509.
- [103] XENON Collaboration, *The XENON1T Dark Matter Experiment*, Eur. Phys. J. C **77** (2017) 881.
- [104] DAMIC Collaboration, *Measurement of radioactive contamination in the high-resistivity silicon CCDs of the DAMIC experiment*, JINST **10** (2015) P08014.
- [105] E. Hubble, *A relation between distance and radial velocity among extra-galactic nebulae*, PNAS **15** (1929) 168.
- [106] C. Abel et al., *Measurement of the Permanent Electric Dipole Moment of the Neutron*, Phys. Rev. Lett. **124** (2020).

BIBLIOGRAPHY

- [107] R. J. Crewther et al., *Chiral estimate of the electric dipole moment of the neutron in quantum chromodynamics*, Phys. Lett. B **88** (1979) 123.
- [108] Super-Kamiokande Collaboration, Y. Fukuda et al., *Evidence for Oscillation of Atmospheric Neutrinos*, Phys. Rev. Lett. **81** (1998) 1562.
- [109] E. Majorana and L. Maiani, *A symmetric theory of electrons and positrons*, in *Ettore Majorana Scientific Papers: On occasion of the centenary of his birth*, G. F. Bassani, ed. Springer, 2006.
- [110] P. Minkowski, $\mu \rightarrow e\gamma$ at a rate of one out of 109 muon decays?, Phys. Lett. B **67** (1977) 421.
- [111] KamLAND-Zen Collaboration, *Search for Majorana Neutrinos Near the Inverted Mass Hierarchy Region with KamLAND-Zen*, Phys. Rev. Lett. **117** (2016) 082503.
- [112] GERDA Collaboration, *Final Results of GERDA on the Search for Neutrinoless Double- β Decay*, Phys. Rev. Lett. **125** (2020) 252502.
- [113] E. Fermi, *Versuch einer Theorie der β -Strahlen. I*, Z. Phys. **88** (1934) 161.
- [114] B. Pontecorvo, *Neutrino Experiments and the Problem of Conservation of Leptonic Charge*, Zh. Eksp. Teor. Fiz. **53** (1967) 1717.
- [115] Z. Maki, M. Nakagawa, and S. Sakata, *Remarks on the Unified Model of Elementary Particles*, Prog. Theo. Phys. **28** (1962) 870.
- [116] T2K Collaboration, *Constraint on the matter–antimatter symmetry-violating phase in neutrino oscillations*, Nature **580** (2020) 339.
- [117] T2K Collaboration, *Improved constraints on neutrino mixing from the T2K experiment with 3.13×10^{21} protons on target*, Phys. Rev. D **103** (2021) 112008.
- [118] Y. L. Dokshitser, *Calculation of structure functions of deep-inelastic scattering and e^+e^- annihilation by perturbation theory in quantum chromodynamics*, JETP **73** (1977) 1216.
- [119] V. N. Gribov and L. N. Lipatov, *Deep inelastic $e p$ scattering in perturbation theory*, Sov. J. Nucl. Phys. **15** (1972) 438.
- [120] G. Altarelli and G. Parisi, *Asymptotic Freedom in Parton Language*, Nucl. Phys. B **126** (1977) 298.
- [121] S. Bailey et al., *Parton distributions from LHC, HERA, Tevatron and fixed target data: MSHT20 PDFs*, Eur. Phys. J. C **81** (2021) 342.
- [122] R. Ball et al., *Parton distributions from high-precision collider data*, Eur. Phys. J. C **77** (2017) 663.

- [123] T.-J. Hou, et al., *New CTEQ global analysis of quantum chromodynamics with high-precision data from the LHC*, Phys. Rev. D **103** (2021) 014013.
- [124] H. Abramowicz et al., *Combination of measurements of inclusive deep inelastic $e^\pm p$ scattering cross sections and QCD analysis of HERA data*, Eur. Phys. J. C **75** (2015) 580.
- [125] S. Alekhin et al., *Parton distribution functions, α_s , and heavy-quark masses for LHC Run II*, Phys. Rev. D **96** (2017) 014011.
- [126] S. Catani, D. de Florian, and G. Rodrigo, *Factorization violation in the collinear limit*, in *Proceedings of “Loops and Legs in Quantum Field Theory” 11th DESY Workshop on Elementary Particle Physics — PoS(LL2012)*. 2013.
- [127] J. C. Collins, D. E. Soper, and G. Sterman, *Factorization for short distance hadron-hadron scattering*, Nuc. Phys. B **261** (1985) 104.
- [128] J. Collins and D. Soper, *The Theorems of Perturbative QCD*, Annual Review of Nuclear and Particle Science **37** (1987) 383.
- [129] J. C. Collins, D. E. Soper, and G. F. Sterman, *Factorization of Hard Processes in QCD*, in *Perturbative Quantum Chromodynamics, Adv. Ser. Direct. High Energy Phys. - Vol. 5*, A. H. Mueller, ed. 1989.
- [130] ATLAS and CDF and CMS and D0 Collaborations, *First combination of Tevatron and LHC measurements of the top-quark mass*, 2014. arXiv:1403.4427.
- [131] A. Czarnecki and K. Melnikov, *Two-loop QCD corrections to top quark width*, Nuc. Phys. B **544** (1999) 520.
- [132] K. G. Chetyrkin et al., *Second order QCD corrections to $\Gamma(t \rightarrow Wb)$* , Phys. Rev. D **60** (1999) 114015.
- [133] I. Bigi, et al., *Production and decay properties of ultra-heavy quarks*, Phys. Lett. B **181** (1986) 157.
- [134] LHC Top Working Group, *LHC Top Working Group Summary Plots*, 2023. <https://twiki.cern.ch/twiki/bin/view/LHCPhysics/LHCTopWGSummaryPlots>. Accessed: 22.9.2023.
- [135] ATLAS Collaboration, *Top working group cross-section summary plots - June 2023*, 2023. <https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PUBNOTES/ATL-PHYS-PUB-2023-014/>. Accessed: 22.9.2023.
- [136] ATLAS Collaboration, *Direct top quark decay width measurement in the $t\bar{t}$ lepton+jets channel at $\sqrt{s} = 8$ TeV with the ATLAS experiment*, Eur. Phys. J. C **78** (2018) 129.

BIBLIOGRAPHY

- [137] ATLAS Collaboration Collaboration,, *Measurement of the top-quark decay width in top-quark pair events in the dilepton channel at $\sqrt{s} = 13$ TeV with the ATLAS detector*, Tech. Rep. ATLAS-CONF-2019-038, CERN, Geneva, 2019.
- [138] CMS Collaboration Collaboration,, *Measurement of the ratio $B(t \rightarrow Wb)/B(t \rightarrow Wq)$ in pp collisions at $\sqrt{s} = 8$ TeV*, Phys. Lett. B **736** (2014) 33.
- [139] W. Bernreuther and Z.-G. Si, *Top quark and leptonic charge asymmetries for the Tevatron and LHC*, Phys. Rev. D **86** (2012) 034026.
- [140] J. Aguilar-Saavedra et al., *Shedding light on the $t\bar{t}$ asymmetry: the photon handle*, JHEP **4** (2014) 188.
- [141] F. Maltoni, et al., *Top-quark charge asymmetry and polarization in $t\bar{t}W^\pm$ production at the LHC*, Phys. Lett. B **736** (2014) 252.
- [142] S. Berge and S. Westhoff, *Top-quark charge asymmetry goes forward: two new observables for hadron colliders*, JHEP **7** (2013) 179.
- [143] ATLAS Collaboration, *Evidence for the charge asymmetry in $pp \rightarrow t\bar{t}$ production at $\sqrt{s}=13$ TeV with the ATLAS detector*, JHEP **8** (2023) 77.
- [144] ATLAS Collaboration, *Measurement of the energy asymmetry in $t\bar{t}j$ production at 13TeV with the ATLAS experiment and interpretation in the SMEFT framework*, Eur. Phys. J. C **82** (2022) 374.
- [145] ATLAS Collaboration, *Search for leptonic charge asymmetry in $t\bar{t}W$ production in final states with three leptons at $\sqrt{s} = 13$ TeV*, JHEP **7** (2023) 33.
- [146] ATLAS collaboration, *Measurement of the charge asymmetry in top-quark pair production in association with a photon with the ATLAS experiment*, Phys. Lett. B **843** (2023) 137848.
- [147] W. Bernreuther, et al., *Top-Quark Spin Correlations at Hadron Colliders: Predictions at Next-to-Leading Order QCD*, Phys. Rev. Lett. **87** (2001) 242002.
- [148] ATLAS Collaboration, *Measurements of top-quark pair spin correlations in the $e\mu$ channel at $\sqrt{s}=13$ TeV using pp collisions in the ATLAS detector*, Eur. Phys. J. C **80** (2020) 754.
- [149] CMS Collaboration, *Measurement of the top quark polarization and $t\bar{t}$ spin correlations using dilepton final states in proton-proton collisions at $\sqrt{s} = 13$ TeV*, Phys. Rev. D **100** (2019) 072002.
- [150] Y. Afik and J. de Nova, *Entanglement and quantum tomography with top quarks at the LHC*, Eur. Phys. J. Plus **136** (2021) 907.

- [151] ATLAS Collaboration, *Observation of quantum entanglement in top-quark pairs using the ATLAS detector*, 2023. arXiv:2311.07288. Submitted to Nature.
- [152] A. Czarnecki, J. G. Körner, and J. H. Piclum, *Helicity fractions of W bosons from top quark decays at next-to-next-to-leading order in QCD*, Phys. Rev. D **81** (2010) 111503.
- [153] ATLAS Collaboration, *Measurement of the W boson polarization in top quark decays with the ATLAS detector*, JHEP **6** (2012) 88.
- [154] ATLAS Collaboration, *Measurement of the W boson polarisation in $t\bar{t}$ events from pp collisions at $\sqrt{s}=8$ TeV in the lepton + jets channel with ATLAS*, Eur. Phys. J. C **77** (2017) 264.
- [155] ATLAS Collaboration, *Measurement of the polarisation of W bosons produced in top-quark decays using dilepton events at $\sqrt{s}=13$ TeV with the ATLAS experiment*, Phys. Lett. B **843** (2023) 137829.
- [156] CMS Collaboration, *Measurement of the W -boson helicity in top-quark decays from $t\bar{t}$ production in lepton+jets events in pp collisions at $\sqrt{s}=7$ TeV*, JHEP **10** (2013) 167.
- [157] CMS Collaboration, *Measurement of the W boson helicity in events with a single reconstructed top quark in pp collisions at $\sqrt{s}=8$ TeV*, JHEP **1** (2015) 53.
- [158] CMS Collaboration, *Measurement of the W boson helicity fractions in the decays of top quark pairs to lepton+jets final states produced in pp collisions at $\sqrt{s}=8$ TeV*, Phys. Lett. B **762** (2016) 512.
- [159] ATLAS and CMS Collaboration, *Combination of the W boson polarization measurements in top quark decays using ATLAS and CMS data at $\sqrt{s}=8$ TeV*, JHEP **8** (2020) 51.
- [160] ATLAS Collaboration, *Top working group top+ X summary plots - June 2023*, 2023. <https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PUBNOTES/ATL-PHYS-PUB-2023-013/>. Accessed: 26.9.2023.
- [161] D. Atwood and A. Soni, *Analysis for magnetic moment and electric dipole moment form factors of the top quark via $e^+e^- \rightarrow t\bar{t}$* , Phys. Rev. D **45** (1992) 2405.
- [162] J. Papavassiliou and C. Parrinello, *Gauge-invariant top quark form factors from e^+e^- experiments*, Phys. Rev. D **50** (1994) 3059.
- [163] W. Hollik et al., *Top dipole form-factors and loop induced CP violation in supersymmetry*, [Erratum: Nucl. Phys. B557 (1999),407].
- [164] U. Baur et al., *Probing electroweak top quark couplings at hadron colliders*, Phys. Rev. D **71** (2005) 054013.

BIBLIOGRAPHY

- [165] W. Bernreuther et al., *QCD Corrections to Static Heavy-Quark Form Factors*, Phys. Rev. Lett. **95** (2005) 261802.
- [166] M. Fael and T. Gehrmann, *Probing top quark electromagnetic dipole moments in single-top-plus-photon production*, Phys. Rev. D **88** (2013) 033003.
- [167] J. Aguilar-Saavedra, *A minimal set of top anomalous couplings*, Nucl. Phys. B **812** (2009) 181–204.
- [168] M. Schulze and Y. Soreq, *Pinning down electroweak dipole operators of the top quark*, Eur. Phys. J. C **76** (2016) 466.
- [169] S. M. Etesamia, S. Khatibi, and M. M. Najafabadi, *Measuring anomalous $WW\gamma$ and $t\bar{t}\gamma$ couplings using $top+\gamma$ production at the LHC*, Eur. Phys. J. C **76** (2016) 533.
- [170] T. Ibrahim and P. Nath, *Top quark electric dipole moment in a minimal supersymmetric standard model extension with vectorlike multiplets*, Phys. Rev. D **82** (2010) 055001.
- [171] T. Ibrahim and P. Nath, *Chromoelectric dipole moment of the top quark in models with vectorlike multiplets*, Phys. Rev. D **84** (2011) 015003.
- [172] O. B. Bylund, et al., *Probing top quark neutral couplings in the Standard Model Effective Field Theory at NLO in QCD*, JHEP **5** (2016) 52.
- [173] W. Buchmüller and D. Wyler, *Effective lagrangian analysis of new interactions and flavour conservation*, Nucl. Phys. B **268** (1986) 621–653.
- [174] C. Burges and H. J. Schnitzer, *Virtual effects of excited quarks as probes of a possible new hardonic mass scale*, Nucl. Phys. B **228** (1983) 464–500.
- [175] C. Leung, S. Love, and S. Rao, *Low-energy manifestations of a new interactions scale: Operator analysis*, Z. Phys. C **31** (1986) 433.
- [176] K. G. Wilson, *Non-Lagrangian Models of Current Algebra*, Phys. Rev. **179** (1969) 1499.
- [177] U. Baur, M. Buice, and L. H. Orr, *Direct measurement of the top quark charge at hadron colliders*, Phys. Rev. D **64** (2001) 094019.
- [178] G. Bevilacqua et al., *Off-shell vs on-shell modelling of top quarks in photon associated production*, JHEP **3** (2020) 154.
- [179] K. Zoch, *Cross-section measurements of top-quark pair production in association with a hard photon at 13 TeV with the ATLAS detector*. PhD thesis, II.Physik-UniGö-Diss-2020/01, 2020.

- [180] G. Bevilacqua et al., *Hard photons in hadroproduction of top quarks with realistic final states*, JHEP **10** (2018) 158.
- [181] P.-F. Duan, et al., *QCD corrections to associated production of $t\bar{t}\gamma$ at hadron colliders*, Phys. Rev. D **80** (2009) 014022.
- [182] P.-F. Duan et al., *Next-to-Leading Order QCD Corrections to $t\bar{t}\gamma$ Production at the 7 TeV LHC*, Chin. Phys. Lett. **28** (2011) 111401.
- [183] F. Maltoni, D. Pagani, and I. Tsinikos, *Associated production of a top-quark pair with vector bosons at NLO in QCD: impact on $t\bar{t}H$ searches at the LHC*, JHEP **2** (2016) 113.
- [184] P.-F. Duan et al., *Electroweak corrections to top quark pair production in association with a hard photon at hadron colliders*, Phys. Lett. B **766** (2017) 102.
- [185] K. Melnikov, M. Schulze, and A. Scharf, *QCD corrections to top quark pair production in association with a photon at hadron colliders*, Phys. Rev. D **83** (2011) 074013.
- [186] CDF Collaboration, *Evidence for $t\bar{t}\gamma$ production and measurement of $\sigma_{t\bar{t}\gamma}/\sigma_{t\bar{t}}$* , Phys. Rev. D **84** (2011) 031104.
- [187] D. Chang, W.-F. Chang, and E. Ma, *Alternative interpretation of the Fermilab Tevatron top events*, Phys. Rev. D **59** (1999) 091503.
- [188] D. Chang, W.-F. Chang, and E. Ma, *Fitting precision electroweak data with exotic heavy quarks*, Phys. Rev. D **61** (2000) 037301.
- [189] CDF Collaboration, *Exclusion of exotic top-like quarks with $-4/3$ electric charge using jet-charge tagging in single-lepton $t\bar{t}$ events at CDF*, Phys. Rev. D **88** (2013) 032003.
- [190] D0 Collaboration, *Measurement of the electric charge of the top quark in $t\bar{t}$ events*, Phys. Rev. D **90** (2014) 051101.
- [191] CMS Collaboration, *Constraints on the Top-Quark Charge from Top-PairEvents*, Tech. Rep. CMS-PAS-TOP-11-031, CERN, Geneva, 2012.
- [192] ATLAS collaboration, *Measurement of the top quark charge in pp collisions at $\sqrt{s}=7$ TeV with the ATLAS detector*, JHEP **11** (2013) 31.
- [193] CERN, *Where did it all begin?*, <https://www.home.cern/about/who-we-are/our-history>. Including the video "CERN and science for peace". Accessed: 29.9.2023.
- [194] CERN, *Convention for the Establishment of a European Organization for Nuclear Research*, <https://council.web.cern.ch/en/content/convention-establishment-european-organization-nuclear-research>. Accessed: 29.9.2023.

BIBLIOGRAPHY

- [195] United Nations, *Intergovernmental and other organizations*, <https://www.un.org/en/about-us/intergovernmental-and-other-organizations>. Accessed: 29.9.2023.
- [196] CERN, *Our Mission*, <https://www.home.cern/about/who-we-are/our-mission>. Accessed: 29.9.2023.
- [197] CERN, *Proceedings of the ECFA-CERN Workshop : Large Hadron Collider in the LEP tunnel: Lausanne, Switzerland 21 - 27 Mar 1984. ECFA-CERN Workshop on Large Hadron Collider in the LEP tunnel*. CERN, Geneva, 1984.
- [198] CERN Council, *Funding and Approval of the LHC (Resolutions : Annex I - Double Majority, Annex II - Contribution of Germany, Annex III Approval of LHC). Financement et approbation du LHC. 217th Meeting of Committee of Council*, 1994.
- [199] T. Schoerner-Sadenius, ed., *The Large Hadron Collider : Harvest of Run 1*. Springer International Publishing, 2015.
- [200] L. Evans and P. Bryant, *LHC Machine*, JINST **3** (2008) S08001.
- [201] TOTEM Collaboration, *The TOTEM Experiment at the CERN Large Hadron Collider*, JINST **3** (2008) S08007.
- [202] LHCf Collaboration, *The LHCf detector at the CERN Large Hadron Collider*, JINST **3** (2008) S08006.
- [203] ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, JINST **3** (2008) S08003.
- [204] CMS Collaboration, *The CMS experiment at the CERN LHC*, JINST **3** (2008) S08004.
- [205] LHCb Collaboration, *The LHCb Detector at the LHC*, JINST **3** (2008) S08005.
- [206] ALICE Collaboration, *The ALICE experiment at the CERN LHC*, JINST **3** (2008) S08002.
- [207] R. Bruce, et al., *Review of LHC Run 2 Machine Configurations*, Proceedings of the 2019 Evian workshop on LHC beam operations (2019) 187.
- [208] K. Fuchsberger et al., *Operational experience during initial beam commissioning of the LHC*, Proceedings of the 1st International Particle Accelerator Conference (IPAC 2010) (2010) 456.
- [209] M. Lamont, *The First Years of LHC Operation for Luminosity Production*, Proceedings of the 4th International Particle Accelerator Conference (IPAC 2013) (2013) 6.

- [210] ATLAS Collaboration, *Luminosity determination in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector at the LHC*, Tech. Rep. CERN-EP-2022-281, CERN, Geneva, 2022. Submitted to Eur. Phys. J. C.
- [211] ATLAS Collaboration, *Online Luminosity Summary Plots over Multiple Years including full Run 2*, 2023. <https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2>. Accessed: 4.10.2023.
- [212] ATLAS Collaboration, *ATLAS Insertable B-Layer Technical Design Report*, Tech. Rep. CERN-LHCC-2010-013, ATLAS-TDR-19, CERN, Geneva, 2010.
- [213] ATLAS Collaboration, *Production and integration of the ATLAS Insertable B-Layer*, JINST **13** (2018) T05008.
- [214] J. Pequeno, *Computer generated image of the whole ATLAS detector*, 2008. <https://cds.cern.ch/record/1095924>. Accessed: 4.10.2023.
- [215] J. Pequeno, *Computer generated image of the ATLAS inner detector*, 2008. <https://cds.cern.ch/record/1095926>. Accessed: 9.10.2023.
- [216] J. Pequeno, *Computer Generated image of the ATLAS calorimeter*, 2008. <https://cds.cern.ch/record/1095927>. Accessed: 10.10.2023.
- [217] J. Pequeno, *Computer generated image of the ATLAS Muons subsystem*, 2008. <https://cds.cern.ch/record/1095929>. Accessed: 10.10.2023.
- [218] ATLAS Collaboration, *Performance of the ATLAS trigger system in 2015*, Eur. Phys. J. C **77** (2017) 317.
- [219] ATLAS Collaboration, *Operation of the ATLAS trigger system in Run 2*, JINST **15** (2020) P10004.
- [220] S. van der Meer, *Calibration of the effective beam height in the ISR*, Tech. Rep. CERN-ISR-PO-68-31, ISR-PO-68-31, CERN, Geneva, 1968.
- [221] ATLAS Collaboration, *Luminosity determination in pp collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector at the LHC*, Eur. Phys. J. C **71** (2011) 1630.
- [222] P. Grafström and W. Kozanecki, *Luminosity determination at proton colliders*, Prog. Part. Nucl. Phys. **81** (2015) 97–148.
- [223] G. Avoni et al., *The new LUCID-2 detector for luminosity measurement and monitoring in ATLAS*, JINST **13** (2018) P07017.
- [224] ATLAS Collaboration, *The ATLAS Simulation Infrastructure*, Eur. Phys. J. C **70** (2010) 823.
- [225] J. Alwall et al., *The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations*, JHEP **7** (2014) 79.

BIBLIOGRAPHY

- [226] R. Ball et al., *Parton distributions for the LHC run II*, JHEP **4** (2015) 40.
- [227] T. Sjöstrand, S. Mrenna, and P. Skands, *A brief introduction to PYTHIA 8.1*, Comput. Phys. Commun. **178** (2008) 852.
- [228] ATLAS Collaboration, *ATLAS Pythia 8 tunes to 7 TeV data*, Tech. Rep. ATL-PHYS-PUB-2014-021, CERN, Geneva, 2014.
- [229] S. Frixione, et al., *Angular correlations of lepton pairs from vector boson and top quark decays in Monte Carlo simulations*, JHEP **4** (2007) 81.
- [230] P. Artoisenet et al., *Automatic spin-entangled decays of heavy resonances in Monte Carlo simulations*, JHEP **4** (2013) 15.
- [231] D. J. Lange, *The EvtGen particle decay simulation package*, Nucl. Instrum. Meth. A **462** (2001) 152.
- [232] S. Frixione, *Isolated photons in perturbative QCD*, Phys. Lett. B **429** (1998) 369.
- [233] J. Bellm et al., *Herwig 7.2 release note*, Eur. Phys. J. C **80** (2020) 452.
- [234] T. Gleisberg, et al., *Event generation with SHERPA 1.1*, JHEP **02** (2009) 007.
- [235] S. Schumann and F. Krauss, *A parton shower algorithm based on Catani-Seymour dipole factorisation*, JHEP **3** (2008) 038.
- [236] S. Catani and M. Seymour, *A general algorithm for calculating jet cross sections in NLO QCD*, Nucl. Phys. B **485** (1997) 291, [Erratum: Nucl. Phys. B **510** (1998), 503].
- [237] S. Catani, et al., *The dipole formalism for next-to-leading order QCD calculations with massive partons*, Nucl. Phys. B **627** (2002) 189.
- [238] S. Höche et al., *QCD matrix elements + parton showers. The NLO case*, JHEP **4** (2013) 27.
- [239] F. Cascioli, P. Maierhöfer, and S. Pozzorini, *Scattering Amplitudes with Open Loops*, Phys. Rev. Lett. **108** (2012) 111601.
- [240] A. Denner, S. Dittmaier, and L. Hofer, *Collier: A fortran-based complex one-loop library in extended regularizations*, Comput. Phys. Commun. **212** (2017) 220.
- [241] P. Nason, *A New method for combining NLO QCD with shower Monte Carlo algorithms*, JHEP **11** (2004) 40.
- [242] S. Frixione, P. Nason, and C. Oleari, *Matching NLO QCD computations with Parton Shower simulations: the POWHEG method*, JHEP **11** (2007) 70.
- [243] S. Alioli, et al., *A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX*, JHEP **6** (2010) 43.

- [244] M. Czakon and A. Mitov, *Top++: A program for the calculation of the top-pair cross-section at hadron colliders*, Comput. Phys. Commun. **185** (2014) 2930.
- [245] LHC Higgs Cross Section Working Group, *Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector*. CERN Yellow Reports: Monographs. CERN, Geneva, 2017.
- [246] N. Kidonakis, *Next-to-next-to-leading logarithm resummation for s-channel single top quark production*, Phys. Rev. D **81** (2010) 054028.
- [247] N. Kidonakis, *Two-loop soft anomalous dimensions for single top quark associated production with a W^- or H^-* , Phys. Rev. D **82** (2010) 054018.
- [248] N. Kidonakis, *Next-to-next-to-leading-order collinear and soft gluon corrections for t-channel single top quark production*, Phys. Rev. D **83** (2011) 091503.
- [249] ATLAS Collaboration, *Measurement of $W^\pm$ and Z-boson production cross sections in pp collisions at $s=13$ TeV with the ATLAS detector*, Phys. Lett. B **759** (2016) 601.
- [250] J. M. Campbell and R. K. Ellis, *Update on vector boson pair production at hadron colliders*, Phys. Rev. D **60** (1999) 113006.
- [251] ATLAS Collaboration, *Inclusive and differential cross section measurements of $t\bar{t}Z$ production in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector, including EFT and spin correlations interpretations*, Tech. Rep. ATLAS-CONF-2023-065, CERN, Geneva, 2023.
- [252] ATLAS Collaboration, *Measurement of the photon identification efficiencies with the ATLAS detector using LHC Run 2 data collected in 2015 and 2016*, Eur. Phys. J. C **79** (2019) 205.
- [253] ATLAS Collaboration, *Electron and photon energy calibration with the ATLAS detector using 2015–2016 LHC proton-proton collision data*, JINST **14** (2019) P03017.
- [254] ATLAS Collaboration, *Electron reconstruction and identification in the ATLAS experiment using the 2015 and 2016 LHC proton–proton collision data at $\sqrt{s} = 13$ TeV*, Eur. Phys. J. C **79** (2019) 639.
- [255] ATLAS Collaboration, *Tagging non-prompt electrons and muons*, Tech. Rep. ATL-COM-PHYS-2016-1444, CERN, Geneva, 2016.
- [256] ATLAS Collaboration, *Muon reconstruction performance of the ATLAS detector in proton–proton collision data at $\sqrt{s}=13$ TeV*, Eur. Phys. J. C **76** (2016) 292.
- [257] M. Cacciari, G. P. Salam, and G. Soyez, *The anti- k_t jet clustering algorithm*, JHEP **4** (2008) 63.

BIBLIOGRAPHY

- [258] ATLAS Collaboration, *Jet reconstruction and performance using particle flow with the ATLAS Detector*, Eur. Phys. J. C **77** (2017) 466.
- [259] ATLAS Collaboration, *Jet energy scale measurements and their systematic uncertainties in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, Phys. Rev. D **96** (2017) 072002.
- [260] Atlas Collaboration, *Tagging and suppression of pileup jets with the ATLAS detector*, Tech. Rep. ATLAS-CONF-2014-018, CERN, Geneva, May, 2014.
- [261] ATLAS Collaboration, *Optimisation and performance studies of the ATLAS b -tagging algorithms for the 2017-18 LHC run*, Tech. Rep. ATL-PHYS-PUB-2017-013, CERN, Geneva, 2017.
- [262] ATLAS Collaboration, *ATLAS b -jet identification performance and efficiency measurement with $t\bar{t}$ events in pp collisions at $\sqrt{s} = 13$ TeV*, Eur. Phys. J. C **79** (2019) 970.
- [263] ATLAS Collaboration, *Performance of missing transverse momentum reconstruction with the ATLAS detector using proton-proton collisions at $\sqrt{s} = 13$ TeV*, Eur. Phys. J. C **78** (2018) 903.
- [264] ATLAS Collaboration, *E_T^{miss} performance in the ATLAS detector using 2015-2016 LHC p - p collisions*, Tech. Rep. ATLAS-CONF-2018-023, CERN, Geneva, 2018.
- [265] M. Cacciari, G. P. Salam, and G. Soyez, *The catchment area of jets*, JHEP **4** (2008) 5.
- [266] ATLAS Collaboration, *Search for Scalar Diphoton Resonances in the Mass Range 65–600 GeV with the ATLAS Detector in pp Collision Data at $\sqrt{s} = 8$ TeV*, Phys. Rev. Lett. **113** (2014) 171801.
- [267] M. Oreglia, *A study of the reactions $\psi' \rightarrow \gamma\gamma\psi$* . PhD thesis, SLAC Report No. SLAC-R-0236, 1980.
- [268] S. Bernstein, *Démonstration du théorème de Weierstrass fondée sur le calcul des probabilités*, Comm. Kharkov Math. Soc. **13** (1912) 1.
- [269] ATLAS Collaboration, *Estimation of fake lepton background for top analyses using the Matrix Method with the 2015 dataset at $\sqrt{s}=13$ TeV with AnalysisTop-2.3.41*, Tech. Rep. ATL-COM-PHYS-2016-198, CERN, Geneva, 2016.
- [270] F. Chollet et al., *Keras*, <https://keras.io>, 2015.
- [271] M. Abadi et al., *TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems*, 2015. <https://www.tensorflow.org/>. Software available from tensorflow.org.

- [272] M. Minsky and S. Papert, *Perceptrons*. MIT Press, Cambridge, 1969.
- [273] F. Rosenblatt, *The Perceptron: A Probabilistic Model for Information Storage and Organization in The Brain*, Psychol. Rev. **65** (1958) 386.
- [274] D. Rumelhart, G. Hinton, and R. Williams, *Learning representations by back-propagating errors*, Nature **323** (1986) 533.
- [275] D. P. Kingma and J. Ba, *Adam: A Method for Stochastic Optimization*, 2017. arXiv:1412.6980.
- [276] I. Sutskever, et al., *On the importance of initialization and momentum in deep learning*, in *Proceedings of the 30th International Conference on Machine Learning*. 2013.
- [277] Y. Nesterov, *A method for solving the convex programming problem with convergence rate $O(1/k^2)$* , Proc. USSR Acad. Sci. **269** (1983) 543.
- [278] T. Dozat, *Incorporating Nesterov Momentum into Adam*, in *Proceedings of the 4th International Conference on Learning Representations*. 2016.
- [279] N. Srivastava, et al., *Dropout: A Simple Way to Prevent Neural Networks from Overfitting*, JMLR **15** (2014) 1929.
- [280] A. Kirchhoff, *Studies to classify prompt photons in the $t\bar{t}\gamma$ -process*, II.Physik-UniGö-MSc-2018/04 (2018).
- [281] J. H. Friedman, *Data analysis techniques for high energy particle physics*, in *Proceedings of the 1974 CERN School of Computing: Godøysund, Norway 11 - 24 Aug 1974. 3rd CERN School of Computing*. 1974.
- [282] J. Butterworth, et al., *PDF4LHC recommendations for LHC Run II*, J. Phys. G **43** (2016) 023001.
- [283] S. Catani, et al., *QCD Matrix Elements + Parton Showers*, JHEP **2001** (2002) 063.
- [284] T. Sjöstrand, et al., *PYTHIA 6.3 Physics and Manual*, 2003. arXiv:hep-ph/0308153.
- [285] S. Hoeche, et al., *Matching parton showers and matrix elements*, in *HERA and the LHC: A Workshop on the Implications of HERA for LHC Physics: CERN - DESY Workshop 2004/2005 (Midterm Meeting, CERN, 11-13 October 2004; Final Meeting, DESY, 17-21 January 2005)*. 2005.
- [286] P. Nason and C. Oleari, *Generation cuts and Born suppression in POWHEG*, 2013. arXiv:1303.3922.

BIBLIOGRAPHY

- [287] ATLAS Collaboration, *Expected performance of the ATLAS b-tagging algorithms in Run-2*, Tech. Rep. ATL-PHYS-PUB-2015-022, CERN, Geneva, Jul, 2015.
- [288] ATLAS Collaboration, *Optimisation of the ATLAS b-tagging performance for the 2016 LHC Run*, Tech. Rep. ATL-PHYS-PUB-2016-012, CERN, Geneva, Jun, 2016.
- [289] ATLAS Collaboration, *Using Event Weights to account for differences in Instantaneous Luminosity and Trigger Prescale in Monte Carlo and Data*, Tech. Rep. ATL-COM-SOFT-2015-119, CERN, Geneva, 2015.
- [290] ATLAS Collaboration, *Luminosity determination in pp collisions at $\sqrt{s}=8$ TeV using the ATLAS detector at the LHC*, Eur. Phys. J. C **76** (2016) 653.
- [291] M. Czakon, P. Fiedler, and A. Mitov, *Total Top-Quark Pair-Production Cross Section at Hadron Colliders Through $\mathcal{O}(\alpha_S^4)$* , Phys. Rev. Lett. **110** (2013) 252004.
- [292] R. Barlow and C. Beeston, *Fitting using finite Monte Carlo samples*, Computer Physics Communications **77** (1993) 219–228.
- [293] R. Brun and F. Rademakers, *ROOT - An Object Oriented Data Analysis Framework*, Nucl. Instrum. Meth. A **389** (1997) 81.
- [294] ROOT Collaboration, K. Cranmer, et al., *HistFactory: A tool for creating statistical models for use with RooFit and RooStats*, Tech. Rep. CERN-OPEN-2012-016, Geneva, 2012.
- [295] W. Verkerke and D. P. Kirkby, *The RooFit toolkit for data modeling*, Tech. Rep. CHEP-2003-MOLT007, SLAC, La Jolla, CA, 2003.
- [296] L. Moneta, et al., *The RooStats Project*, in *Proceedings, 13th International Workshop on Advanced computing and analysis techniques in physics research (ACAT2010) : Jaipur, India, February 22-27, 2010*. 2010.
- [297] F. James and M. Roos, *Minuit: A System for Function Minimization and Analysis of the Parameter Errors and Correlations*, Comput. Phys. Commun. **10** (1975) 343.
- [298] N. Castro, et al., *EFTfitter: a tool for interpreting measurements in the context of effective field theories*, Eur. Phys. J. C **76** (2016) 432.
- [299] T. Ježo, et al., *An NLO+PS generator for $t\bar{t}$ and Wt production and decay including non-resonant and interference effects*, Eur. Phys. J. C **76** (2016) 691.

Appendices

APPENDIX A

Monte Carlo Samples

The following tables¹ give an overview of the Monte-Carlo samples used for this analysis.

Table A.1.: Monte-Carlo samples with top quarks used for the analysis. Some of them contain filters for exactly one lepton, at least one lepton or two leptons. The stated cross-sections include k -factors and possible generator filter efficiencies.

Process	Generator	Tune	Xsec [pb]
$t\bar{t}\gamma$ (NLO $2 \rightarrow 3$, inclusive)	aMC@NLO + PYTHIA8	A14 NNPDF2.3LO	1.572
$t\bar{t}\gamma$ (<i>var3c</i> up)	aMC@NLO + PYTHIA8	A14 NNPDF2.3LO	1.572
$t\bar{t}\gamma$ (<i>var3c</i> down)	aMC@NLO + PYTHIA8	A14 NNPDF2.3LO	1.572
$t\bar{t}\gamma$ (NLO $2 \rightarrow 3$)	aMC@NLO + HERWIG7	H7.1-Default	1.575
$t\bar{t} + \gamma$ from decay (LO $2 \rightarrow 2$)	MG5 + PYTHIA8	A14 NNPDF2.3LO	2.261
$t\bar{t} + \gamma$ from decay (shower variation)	MG5 + HERWIG7	H7.1-Default	2.280
$Wt\gamma$ (γ from production)	MG5 + PYTHIA8	A14 NNPDF2.3LO	0.25
$Wt\gamma$ (γ from decay)	MG5 + PYTHIA8	A14 NNPDF2.3LO	0.29
$Wt\gamma$ (γ from production, shower variation)	aMC@NLO + HERWIG7	MMHT2014	0.25
$Wt\gamma$ (γ from decay, shower variation))	aMC@NLO + HERWIG7	MMHT2014	0.30
$t\bar{t}W$	aMC@NLO + PYTHIA8	A14 NNPDF2.3LO	0.60
$t\bar{t}\nu\nu$	aMC@NLO + PYTHIA8	A14 NNPDF2.3LO	0.17
$t\bar{t}qq$	aMC@NLO + PYTHIA8	A14 NNPDF2.3LO	0.59
$t\bar{t}ee$	aMC@NLO + PYTHIA8	A14 NNPDF2.3LO	0.04
$t\bar{t}\mu\mu$	aMC@NLO + PYTHIA8	A14 NNPDF2.3LO	0.04
$t\bar{t}\tau\tau$	aMC@NLO + PYTHIA8	A14 NNPDF2.3LO	0.04

¹The tables are taken from the internal ATLAS documentation. Some internal information has been removed.

A. Monte Carlo Samples

Table A.2.: Further Monte-Carlo samples with top quarks used for the analysis. Some of them contain filters for exactly one lepton, at least one lepton or two leptons. The stated cross-sections include k -factors and possible generator filter efficiencies.

Process	Generator	Tune	Xsec [pb]
$t\bar{t}$ (inclusive)	POWHEG + PYTHIA8	A14 NNPDF2.3LO	452.34
$t\bar{t}$ (inclusive, AFII)	POWHEG + PYTHIA8	A14 NNPDF2.3LO	452.34
$t\bar{t}$ (1+jets, hdamp variation)	POWHEG + PYTHIA8	A14 NNPDF2.3LO	364.74
$t\bar{t}$ (1+jets, p_T hard)	POWHEG + PYTHIA8	A14 NNPDF2.3LO	320.12
$t\bar{t}$ (1+jets, shower variation)	POWHEG + HERWIG7	H7.1-Default	320.12
$t\bar{t}$ (dilepton)	POWHEG + PYTHIA8	A14 NNPDF2.3LO	87.71
$t\bar{t}$ (dilepton, AFII)	POWHEG + PYTHIA8	A14 NNPDF2.3LO	87.71
$t\bar{t}$ (dilepton, hdamp variation)	POWHEG + PYTHIA8	A14 NNPDF2.3LO	87.72
$t\bar{t}$ (dilepton, p_T hard)	POWHEG + PYTHIA8	A14 NNPDF2.3LO	76.96
$t\bar{t}$ (dilepton, shower variation)	POWHEG + HERWIG7	H7.1-Default	76.95
Wt (inclusive, t , DR)	POWHEG + PYTHIA8	A14 NNPDF2.3LO	35.86
Wt (inclusive, $\bar{t}$, DR)	POWHEG + PYTHIA8	A14 NNPDF2.3LO	35.85
Wt (inclusive, t , shower variation)	POWHEG + HERWIG7	MMHT2014	35.85
Wt (inclusive, $\bar{t}$, shower variation)	POWHEG + HERWIG7	MMHT2014	35.85
Wt (dilepton, t , DR)	POWHEG + PYTHIA8	A14 NNPDF2.3LO	3.78
Wt (dilepton, $\bar{t}$, DR)	POWHEG + PYTHIA8	A14 NNPDF2.3LO	3.78
Wt (dilepton, t , shower variation)	POWHEG + HERWIG7	MMHT2014	3.78
Wt (dilepton, $\bar{t}$, shower variation)	POWHEG + HERWIG7	MMHT2014	3.78
s-channel, top	POWHEG + PYTHIA8	A14 NNPDF2.3LO	2.06
s-channel, anti-top	POWHEG + PYTHIA8	A14 NNPDF2.3LO	1.29
t-channel, top	POWHEG + PYTHIA8	A14 NNPDF2.3LO	44.15
t-channel, anti-top	POWHEG + PYTHIA8	A14 NNPDF2.3LO	26.28

Table A.3.: Monte-Carlo samples for $V\gamma$ and VV production used for the analysis. The samples are split into slices of photon transverse momenta. The stated cross-sections include k -factors and possible generator filter efficiencies.

Process	Generator	Tune	Xsec [pb]
$e\nu\gamma$	SHERPA 2.2.8 (NLO)	NNPDF3.0NNLO	356.31
$\mu\nu\gamma$	SHERPA 2.2.8 (NLO)	NNPDF3.0NNLO	358.25
$\tau\nu\gamma$	SHERPA 2.2.8 (NLO)	NNPDF3.0NNLO	356.36
$ee\gamma$	SHERPA 2.2.8 (NLO)	NNPDF3.0NNLO	98.705
$\mu\mu\gamma$	SHERPA 2.2.8 (NLO)	NNPDF3.0NNLO	98.690
$\tau\tau\gamma$	SHERPA 2.2.8 (NLO)	NNPDF3.0NNLO	98.715
$\nu\nu\gamma$	SHERPA 2.2.8 (NLO)	NNPDF3.0NNLO	57.023
$ZZ \rightarrow \ell\ell\ell\ell$	SHERPA 2.2.2	NNPDF3.0NNLO	1.25
$ZW \rightarrow \ell\ell\nu$	SHERPA 2.2.2	NNPDF3.0NNLO	4.58
$WW/ZZ \rightarrow \ell\ell\nu\nu$	SHERPA 2.2.2	NNPDF3.0NNLO	12.50
$WZ \rightarrow \ell\nu\nu$	SHERPA 2.2.2	NNPDF3.0NNLO	3.23
$ZZ \rightarrow qq\nu\nu$	SHERPA 2.2.1	NNPDF3.0NNLO	4.31
$ZZ \rightarrow qq\ell\ell$	SHERPA 2.2.1	NNPDF3.0NNLO	2.20
$WZ \rightarrow qq\nu\nu$	SHERPA 2.2.1	NNPDF3.0NNLO	6.80
$WZ \rightarrow qq\ell\ell$	SHERPA 2.2.1	NNPDF3.0NNLO	3.43
$WW \rightarrow qq\ell^-\nu$	SHERPA 2.2.1	NNPDF3.0NNLO	24.71
$WW \rightarrow qq\ell^+\nu$	SHERPA 2.2.1	NNPDF3.0NNLO	24.72
$WZ \rightarrow \ell\nu qq$	SHERPA 2.2.1	NNPDF3.0NNLO	11.42

Table A.4.: Monte-Carlo samples for Z +jets production used for the analysis. The samples are split into slices where max refers to $\max(H_T, p_T^Z)$, and the samples contain vetoes/filters on b -quarks and c -quarks to optimise statistics. The stated cross-sections include k -factors and possible generator filter efficiencies.

Process	Generator	Tune	Xsec [pb]
$Z \rightarrow ee$, max $\in [0, 70)$ GeV, b -veto, c -veto	SHERPA 2.2.1	NNPDF3.0NLO	1587.02
$Z \rightarrow ee$, max $\in [0, 70)$ GeV, b -veto, c -filter	SHERPA 2.2.1	NNPDF3.0NLO	220.00
$Z \rightarrow ee$, max $\in [0, 70)$ GeV, b -filter	SHERPA 2.2.1	NNPDF3.0NLO	127.09
$Z \rightarrow ee$, max $\in [70, 140)$ GeV, b -veto, c -veto	SHERPA 2.2.1	NNPDF3.0NLO	74.90
$Z \rightarrow ee$, max $\in [70, 140)$ GeV, b -veto, c -filter	SHERPA 2.2.1	NNPDF3.0NLO	20.32
$Z \rightarrow ee$, max $\in [70, 140)$ GeV, b -filter	SHERPA 2.2.1	NNPDF3.0NLO	12.74
$Z \rightarrow ee$, max $\in [140, 280)$ GeV, b -veto, c -veto	SHERPA 2.2.1	NNPDF3.0NLO	24.42
$Z \rightarrow ee$, max $\in [140, 280)$ GeV, b -veto, c -filter	SHERPA 2.2.1	NNPDF3.0NLO	9.23
$Z \rightarrow ee$, max $\in [140, 280)$ GeV, b -filter	SHERPA 2.2.1	NNPDF3.0NLO	6.08
$Z \rightarrow ee$, max $\in [280, 500)$ GeV, b -veto, c -veto	SHERPA 2.2.1	NNPDF3.0NLO	4.80
$Z \rightarrow ee$, max $\in [280, 500)$ GeV, b -veto, c -filter	SHERPA 2.2.1	NNPDF3.0NLO	2.25
$Z \rightarrow ee$, max $\in [280, 500)$ GeV, b -filter	SHERPA 2.2.1	NNPDF3.0NLO	1.49
$Z \rightarrow ee$, max $\in [500, 1000)$ GeV	SHERPA 2.2.1	NNPDF3.0NLO	1.76
$Z \rightarrow ee$, max ≥ 1000 GeV	SHERPA 2.2.1	NNPDF3.0NLO	0.15
$Z \rightarrow \mu\mu$, max $\in [0, 70)$ GeV, b -veto, c -veto	SHERPA 2.2.1	NNPDF3.0NLO	1588.47
$Z \rightarrow \mu\mu$, max $\in [0, 70)$ GeV, b -veto, c -filter	SHERPA 2.2.1	NNPDF3.0NLO	219.52
$Z \rightarrow \mu\mu$, max $\in [0, 70)$ GeV, b -filter	SHERPA 2.2.1	NNPDF3.0NLO	127.13
$Z \rightarrow \mu\mu$, max $\in [70, 140)$ GeV, b -veto, c -veto	SHERPA 2.2.1	NNPDF3.0NLO	73.42
$Z \rightarrow \mu\mu$, max $\in [70, 140)$ GeV, b -veto, c -filter	SHERPA 2.2.1	NNPDF3.0NLO	20.91
$Z \rightarrow \mu\mu$, max $\in [70, 140)$ GeV, b -filter	SHERPA 2.2.1	NNPDF3.0NLO	12.51
$Z \rightarrow \mu\mu$, max $\in [140, 280)$ GeV, b -veto, c -veto	SHERPA 2.2.1	NNPDF3.0NLO	23.30
$Z \rightarrow \mu\mu$, max $\in [140, 280)$ GeV, b -veto, c -filter	SHERPA 2.2.1	NNPDF3.0NLO	9.14
$Z \rightarrow \mu\mu$, max $\in [140, 280)$ GeV, b -filter	SHERPA 2.2.1	NNPDF3.0NLO	6.08
$Z \rightarrow \mu\mu$, max $\in [280, 500)$ GeV, b -veto, c -veto	SHERPA 2.2.1	NNPDF3.0NLO	4.66
$Z \rightarrow \mu\mu$, max $\in [280, 500)$ GeV, b -veto, c -filter	SHERPA 2.2.1	NNPDF3.0NLO	2.21
$Z \rightarrow \mu\mu$, max $\in [280, 500)$ GeV, b -filter	SHERPA 2.2.1	NNPDF3.0NLO	1.47
$Z \rightarrow \mu\mu$, max $\in [500, 1000)$ GeV	SHERPA 2.2.1	NNPDF3.0NLO	1.74
$Z \rightarrow \mu\mu$, max ≥ 1000 GeV	SHERPA 2.2.1	NNPDF3.0NLO	0.14
$Z \rightarrow \tau\tau$, max $\in [0, 70)$ GeV, b -veto, c -veto	SHERPA 2.2.1	NNPDF3.0NLO	1612.53
$Z \rightarrow \tau\tau$, max $\in [0, 70)$ GeV, b -veto, c -filter	SHERPA 2.2.1	NNPDF3.0NLO	211.71
$Z \rightarrow \tau\tau$, max $\in [0, 70)$ GeV, b -filter	SHERPA 2.2.1	NNPDF3.0NLO	127.09
$Z \rightarrow \tau\tau$, max $\in [70, 140)$ GeV, b -veto, c -veto	SHERPA 2.2.1	NNPDF3.0NLO	74.71
$Z \rightarrow \tau\tau$, max $\in [70, 140)$ GeV, b -veto, c -filter	SHERPA 2.2.1	NNPDF3.0NLO	20.53
$Z \rightarrow \tau\tau$, max $\in [70, 140)$ GeV, b -filter	SHERPA 2.2.1	NNPDF3.0NLO	12.77
$Z \rightarrow \tau\tau$, max $\in [140, 280)$ GeV, b -veto, c -veto	SHERPA 2.2.1	NNPDF3.0NLO	24.59
$Z \rightarrow \tau\tau$, max $\in [140, 280)$ GeV, b -veto, c -filter	SHERPA 2.2.1	NNPDF3.0NLO	9.30
$Z \rightarrow \tau\tau$, max $\in [140, 280)$ GeV, b -filter	SHERPA 2.2.1	NNPDF3.0NLO	6.17
$Z \rightarrow \tau\tau$, max $\in [280, 500)$ GeV, b -veto, c -veto	SHERPA 2.2.1	NNPDF3.0NLO	4.76
$Z \rightarrow \tau\tau$, max $\in [280, 500)$ GeV, b -veto, c -filter	SHERPA 2.2.1	NNPDF3.0NLO	2.24
$Z \rightarrow \tau\tau$, max $\in [280, 500)$ GeV, b -filter	SHERPA 2.2.1	NNPDF3.0NLO	1.49
$Z \rightarrow \tau\tau$, max $\in [500, 1000)$ GeV	SHERPA 2.2.1	NNPDF3.0NLO	1.76
$Z \rightarrow \tau\tau$, max ≥ 1000 GeV	SHERPA 2.2.1	NNPDF3.0NLO	0.14

A. Monte Carlo Samples

Table A.5.: Monte-Carlo samples for W +jets production used for the analysis. The samples are split into slices where max refers to $\max(H_T, p_T^W)$, and the samples contain vetoes/filters on b -quarks and c -quarks to optimise statistics. The stated cross-sections include k -factors and possible generator filter efficiencies.

Process	Generator	Tune	Xsec [pb]
$W \rightarrow e\nu$, max $\in [0, 70)$ GeV, b -veto, c -veto	SHERPA 2.2.1	NNPDF3.0NLO	15324.22
$W \rightarrow e\nu$, max $\in [0, 70)$ GeV, b -veto, c -filter	SHERPA 2.2.1	NNPDF3.0NLO	2430.66
$W \rightarrow e\nu$, max $\in [0, 70)$ GeV, b -filter	SHERPA 2.2.1	NNPDF3.0NLO	832.20
$W \rightarrow e\nu$, max $\in [70, 140)$ GeV, b -veto, c -veto	SHERPA 2.2.1	NNPDF3.0NLO	618.70
$W \rightarrow e\nu$, max $\in [70, 140)$ GeV, b -veto, c -filter	SHERPA 2.2.1	NNPDF3.0NLO	224.00
$W \rightarrow e\nu$, max $\in [70, 140)$ GeV, b -filter	SHERPA 2.2.1	NNPDF3.0NLO	77.93
$W \rightarrow e\nu$, max $\in [140, 280)$ GeV, b -veto, c -veto	SHERPA 2.2.1	NNPDF3.0NLO	197.31
$W \rightarrow e\nu$, max $\in [140, 280)$ GeV, b -veto, c -filter	SHERPA 2.2.1	NNPDF3.0NLO	94.96
$W \rightarrow e\nu$, max $\in [140, 280)$ GeV, b -filter	SHERPA 2.2.1	NNPDF3.0NLO	36.54
$W \rightarrow e\nu$, max $\in [280, 500)$ GeV, b -veto, c -veto	SHERPA 2.2.1	NNPDF3.0NLO	38.34
$W \rightarrow e\nu$, max $\in [280, 500)$ GeV, b -veto, c -filter	SHERPA 2.2.1	NNPDF3.0NLO	22.36
$W \rightarrow e\nu$, max $\in [280, 500)$ GeV, b -filter	SHERPA 2.2.1	NNPDF3.0NLO	9.59
$W \rightarrow e\nu$, max $\in [500, 1000)$ GeV	SHERPA 2.2.1	NNPDF3.0NLO	14.60
$W \rightarrow e\nu$, max ≥ 1000 GeV	SHERPA 2.2.1	NNPDF3.0NLO	1.20
$W \rightarrow \mu\nu$, max $\in [0, 70)$ GeV, b -veto, c -veto	SHERPA 2.2.1	NNPDF3.0NLO	15317.86
$W \rightarrow \mu\nu$, max $\in [0, 70)$ GeV, b -veto, c -filter	SHERPA 2.2.1	NNPDF3.0NLO	2431.20
$W \rightarrow \mu\nu$, max $\in [0, 70)$ GeV, b -filter	SHERPA 2.2.1	NNPDF3.0NLO	828.88
$W \rightarrow \mu\nu$, max $\in [70, 140)$ GeV, b -veto, c -veto	SHERPA 2.2.1	NNPDF3.0NLO	618.56
$W \rightarrow \mu\nu$, max $\in [70, 140)$ GeV, b -veto, c -filter	SHERPA 2.2.1	NNPDF3.0NLO	223.36
$W \rightarrow \mu\nu$, max $\in [70, 140)$ GeV, b -filter	SHERPA 2.2.1	NNPDF3.0NLO	76.73
$W \rightarrow \mu\nu$, max $\in [140, 280)$ GeV, b -veto, c -veto	SHERPA 2.2.1	NNPDF3.0NLO	197.79
$W \rightarrow \mu\nu$, max $\in [140, 280)$ GeV, b -veto, c -filter	SHERPA 2.2.1	NNPDF3.0NLO	96.45
$W \rightarrow \mu\nu$, max $\in [140, 280)$ GeV, b -filter	SHERPA 2.2.1	NNPDF3.0NLO	36.51
$W \rightarrow \mu\nu$, max $\in [280, 500)$ GeV, b -veto, c -veto	SHERPA 2.2.1	NNPDF3.0NLO	38.30
$W \rightarrow \mu\nu$, max $\in [280, 500)$ GeV, b -veto, c -filter	SHERPA 2.2.1	NNPDF3.0NLO	22.40
$W \rightarrow \mu\nu$, max $\in [280, 500)$ GeV, b -filter	SHERPA 2.2.1	NNPDF3.0NLO	8.77
$W \rightarrow \mu\nu$, max $\in [500, 1000)$ GeV	SHERPA 2.2.1	NNPDF3.0NLO	14.56
$W \rightarrow \mu\nu$, max ≥ 1000 GeV	SHERPA 2.2.1	NNPDF3.0NLO	1.20
$W \rightarrow \tau\nu$, max $\in [0, 70)$ GeV, b -veto, c -veto	SHERPA 2.2.1	NNPDF3.0NLO	15324.89
$W \rightarrow \tau\nu$, max $\in [0, 70)$ GeV, b -veto, c -filter	SHERPA 2.2.1	NNPDF3.0NLO	2443.43
$W \rightarrow \tau\nu$, max $\in [0, 70)$ GeV, b -filter	SHERPA 2.2.1	NNPDF3.0NLO	837.63
$W \rightarrow \tau\nu$, max $\in [70, 140)$ GeV, b -veto, c -veto	SHERPA 2.2.1	NNPDF3.0NLO	619.45
$W \rightarrow \tau\nu$, max $\in [70, 140)$ GeV, b -veto, c -filter	SHERPA 2.2.1	NNPDF3.0NLO	222.60
$W \rightarrow \tau\nu$, max $\in [70, 140)$ GeV, b -filter	SHERPA 2.2.1	NNPDF3.0NLO	76.99
$W \rightarrow \tau\nu$, max $\in [140, 280)$ GeV, b -veto, c -veto	SHERPA 2.2.1	NNPDF3.0NLO	197.36
$W \rightarrow \tau\nu$, max $\in [140, 280)$ GeV, b -veto, c -filter	SHERPA 2.2.1	NNPDF3.0NLO	93.90
$W \rightarrow \tau\nu$, max $\in [140, 280)$ GeV, b -filter	SHERPA 2.2.1	NNPDF3.0NLO	34.82
$W \rightarrow \tau\nu$, max $\in [280, 500)$ GeV, b -veto, c -veto	SHERPA 2.2.1	NNPDF3.0NLO	39.28
$W \rightarrow \tau\nu$, max $\in [280, 500)$ GeV, b -veto, c -filter	SHERPA 2.2.1	NNPDF3.0NLO	22.25
$W \rightarrow \tau\nu$, max $\in [280, 500)$ GeV, b -filter	SHERPA 2.2.1	NNPDF3.0NLO	9.49
$W \rightarrow \tau\nu$, max $\in [500, 1000)$ GeV	SHERPA 2.2.1	NNPDF3.0NLO	14.60
$W \rightarrow \tau\nu$, max ≥ 1000 GeV	SHERPA 2.2.1	NNPDF3.0NLO	1.20

Additional information on the top quark reconstruction

B.1. Quadratic equation for neutrino p_z

The starting point to derive the neutrino $p_{z,\nu}$ by using the W -boson's pole mass m_W as a constraint is the following equation:

$$m_W^2 = \mathbf{p}_W^2 = (\mathbf{p}_l + \mathbf{p}_\nu)^2, \quad (\text{B.1})$$

where $\mathbf{p}_W$, $\mathbf{p}_l$ and $\mathbf{p}_\nu$ are the 4-vectors of the W -boson and the lepton and neutrino it decays into. This can be written as

$$m_W^2 - m_l^2 - m_\nu^2 \approx m_W^2 = 2\mathbf{p}_l\mathbf{p}_\nu, \quad (\text{B.2})$$

with m_l and m_ν being the lepton and neutrino masses, which are much smaller than m_W . Expanding this into its components gives

$$\frac{m_W^2}{2} = E_l E_\nu - (p_{x,l}p_{x,\nu} + p_{y,l}p_{y,\nu} + p_{z,l}p_{z,\nu}). \quad (\text{B.3})$$

Given that neutrinos are almost massless one gets

$$\frac{m_W^2}{2} = E_l \sqrt{p_{x,\nu}^2 + p_{y,\nu}^2 + p_{z,\nu}^2} - (p_{x,l}p_{x,\nu} + p_{y,l}p_{y,\nu} + p_{z,l}p_{z,\nu}). \quad (\text{B.4})$$

Moving $(p_{x,l}p_{x,\nu} + p_{y,l}p_{y,\nu} + p_{z,l}p_{z,\nu})$ to the left side of the equation, defining

$$\frac{m_W^2}{2} + p_{x,l}p_{x,\nu} + p_{y,l}p_{y,\nu} = C \quad (\text{B.5})$$

and squaring the equation gives

$$C^2 + p_{z,l}^2 p_{z,\nu}^2 + 2C p_{z,l} p_{z,\nu} = E_l^2 (p_{x,\nu}^2 + p_{y,\nu}^2 + p_{z,\nu}^2). \quad (\text{B.6})$$

B. Additional information on the top quark reconstruction

Ordering this by the powers of $p_{z,\nu}$ and using $p_{x,\nu}^2 + p_{y,\nu}^2 = p_{T,\nu}^2$ gives

$$0 = (E_l^2 - p_{z,l}^2)p_{z,\nu}^2 - 2Cp_{z,l}p_{z,\nu} + E_l^2p_{T,\nu}^2 - C^2. \quad (\text{B.7})$$

Defining

$$a = E_l^2 - p_{z,l}^2 \quad (\text{B.8})$$

$$b = 2Cp_{z,l} = 2 \left(\frac{m_W^2}{2} + p_{x,l}p_{x,\nu} + p_{y,l}p_{y,\nu} \right) p_{z,l} \quad (\text{B.9})$$

$$c = E_l^2p_{T,\nu}^2 - \left(\frac{m_W^2}{2} + p_{x,l}p_{x,\nu} + p_{y,l}p_{y,\nu} \right)^2 \quad (\text{B.10})$$

and using $p_{T,\nu} = E_T^{miss}$, $p_{x,\nu} = E_T^{miss} \cos(\phi_{E_T^{miss}})$ and $p_{y,\nu} = E_T^{miss} \sin(\phi_{E_T^{miss}})$ leads to expressions of a , b and c that are only dependent on the measurable quantities E_l , $p_{l,x}$, $p_{l,y}$, $p_{l,z}$, E_T^{miss} and $\phi(E_T^{miss})$. The solutions for the $p_{z,\nu}$ are given by the standard formula to solve quadratic equations

$$p_{z,\nu} = -\frac{b}{2a} \pm \sqrt{\frac{b^2}{4a^2} - \frac{c}{a}}. \quad (\text{B.11})$$

B.2. Plots on top-quark reconstruction

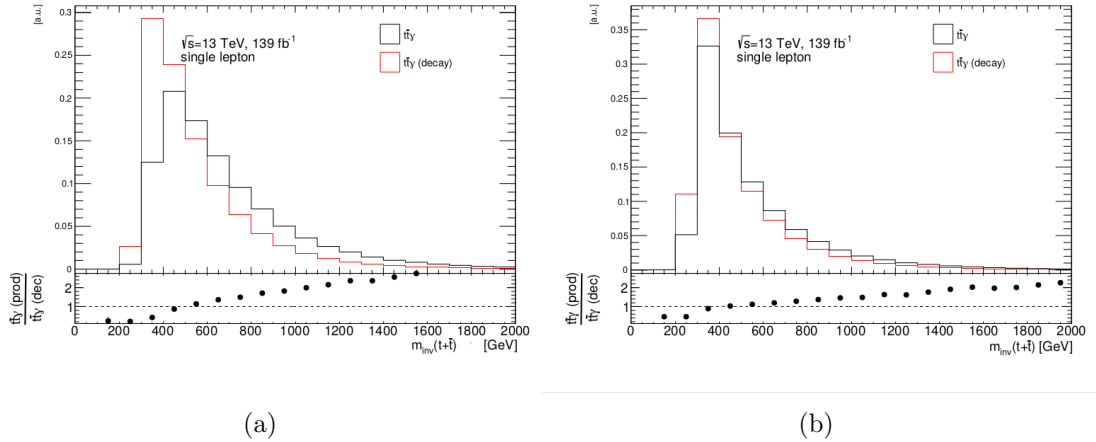


Figure B.1.: The distributions for the summed mass of both reconstructed top quarks (best permutation as described in Section 5.3.2) for $t\bar{t}\gamma$ (dec.) and $t\bar{t}\gamma$ (prod.) at reco-level (left), while on the right the same distributions at particle level are shown.

B.2. Plots on top-quark reconstruction

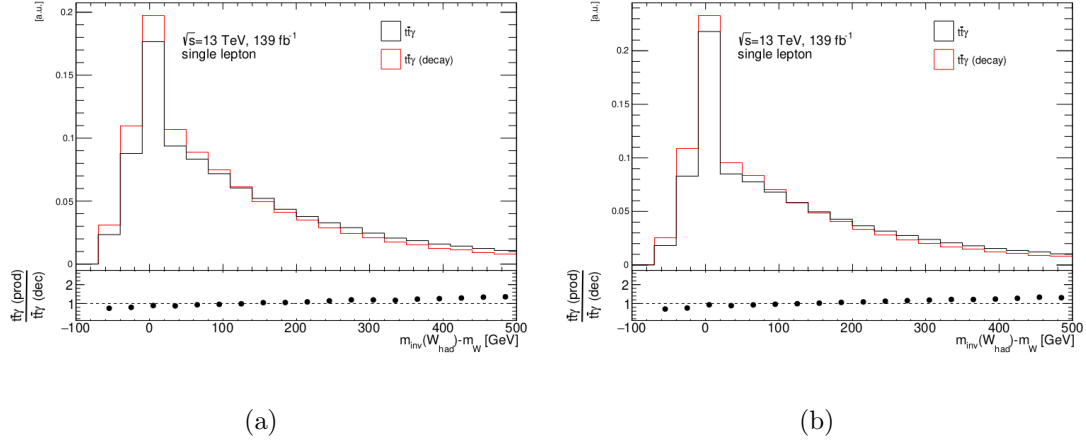


Figure B.2.: The distributions for the mass difference of the reconstructed hadronically decaying W boson and the W pole mass for $t\bar{t}\gamma$ (dec.) and $t\bar{t}\gamma$ (prod.) at reco-level (left), while on the right the same distributions at particle level are shown.

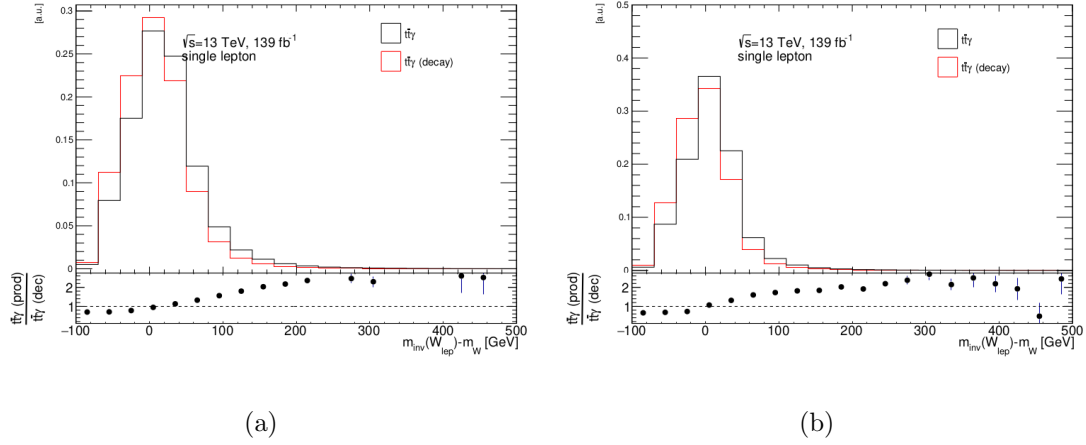


Figure B.3.: The distributions for the mass difference of the reconstructed leptonically decaying W boson and the W pole mass for $t\bar{t}\gamma$ (dec.) and $t\bar{t}\gamma$ (prod.) at reco-level (left), while on the right the same distributions at particle level are shown.

Additional plots on prompt photon background control regions

C.1. $W\gamma$ control region

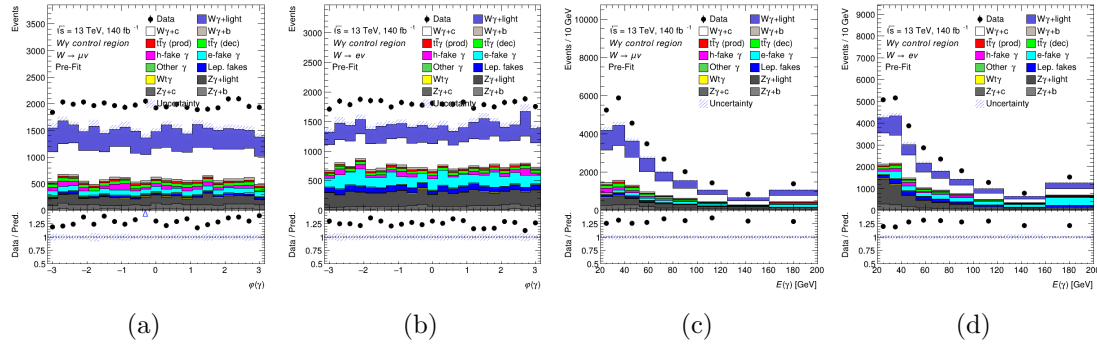


Figure C.1.: Distributions for the $W\gamma$ control region, showing the ϕ and E distributions of the photon for μ +jets ((a) and (c)) and e +jets ((b) and (d)) subchannels. The uncertainty band shows statistical uncertainty only.

C.1. $W\gamma$ control region

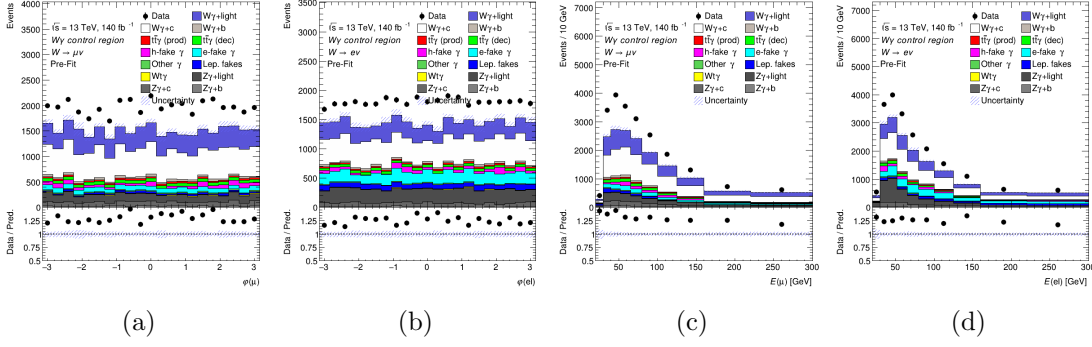


Figure C.5.: Distributions for the $W\gamma$ control region, showing the ϕ and E distributions of the lepton for μ +jets ((a) and (c)) and e +jets ((b) and (d)) subchannels. The uncertainty band shows statistical uncertainty only.

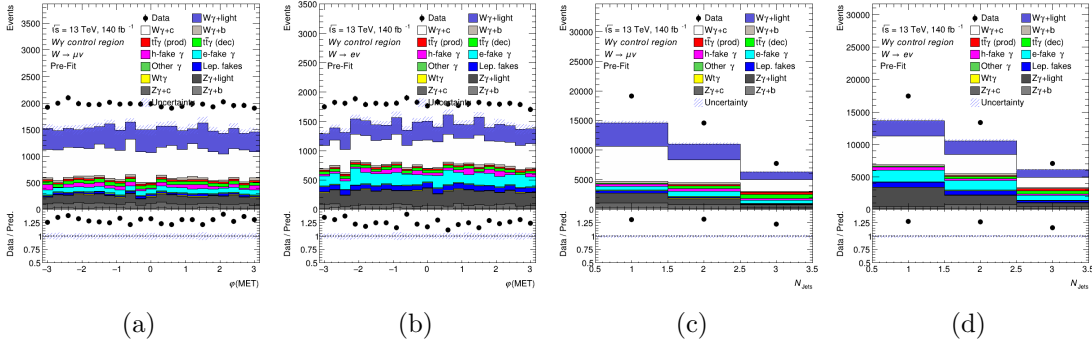


Figure C.6.: Distributions for the $W\gamma$ control region, showing the distributions for $\phi(E_T^{miss})$ and the number of jets for μ +jets ((a) and (c)) and e +jets ((b) and (d)) subchannels. The uncertainty band shows statistical uncertainty only.

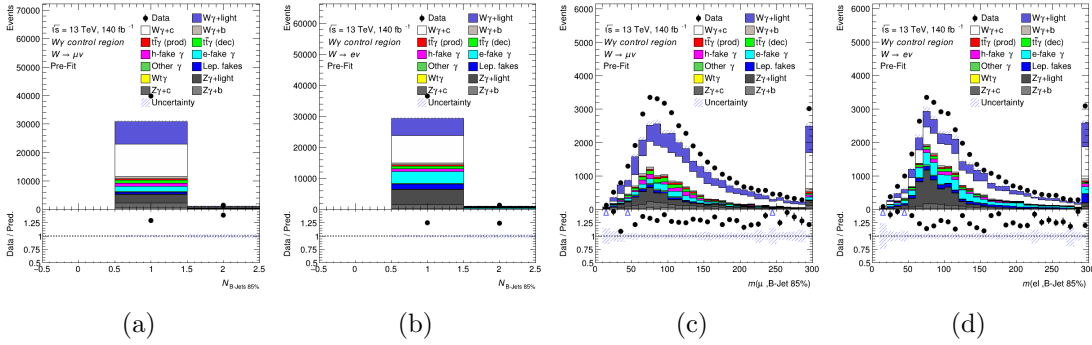


Figure C.7.: Distributions for the $W\gamma$ control region, showing the b -jet multiplicity distributions at the 85% working point and the invariant mass of the b -tagged jet and the lepton for μ +jets ((a) and (c)) and e +jets ((b) and (d)) subchannels. The uncertainty band shows statistical uncertainty only.

C.2. $Z\gamma$ control region

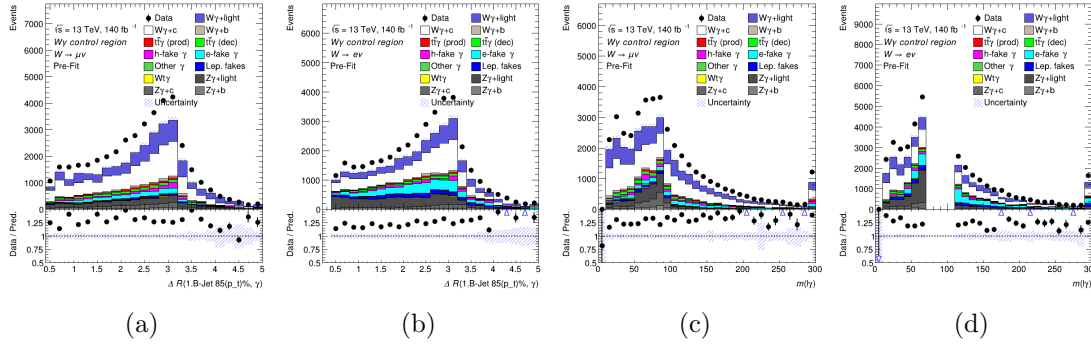


Figure C.11.: Distributions for the $W\gamma$ control region, showing the ΔR distributions of the photon and the leading p_T b -tagged jet, as well as the invariant mass of the lepton and the photon for μ +jets ((a) and (c)) and e +jets ((b) and (d)) subchannels. The uncertainty band shows statistical uncertainty only.

C.2. $Z\gamma$ control region

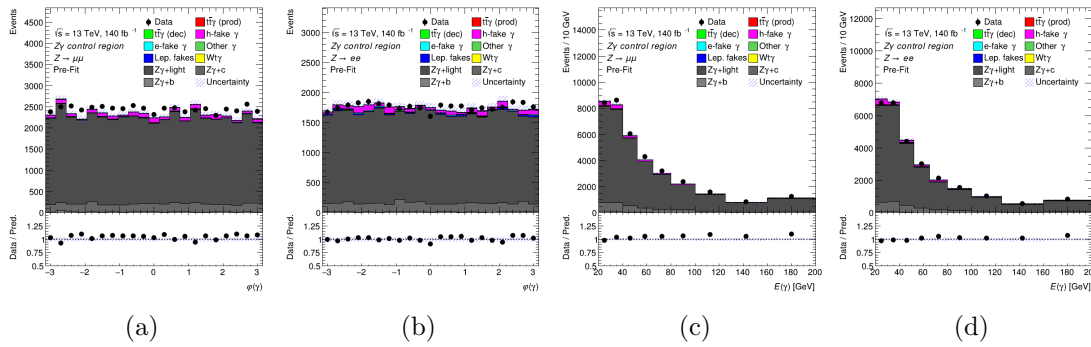


Figure C.12.: Distributions for the $Z\gamma$ control region, showing the ϕ and E distributions of the photon for $\mu\mu$ ((a) and (c)) and ee ((b) and (d)) subchannels. The uncertainty band shows statistical uncertainty only.

C. Additional plots on prompt photon background control regions

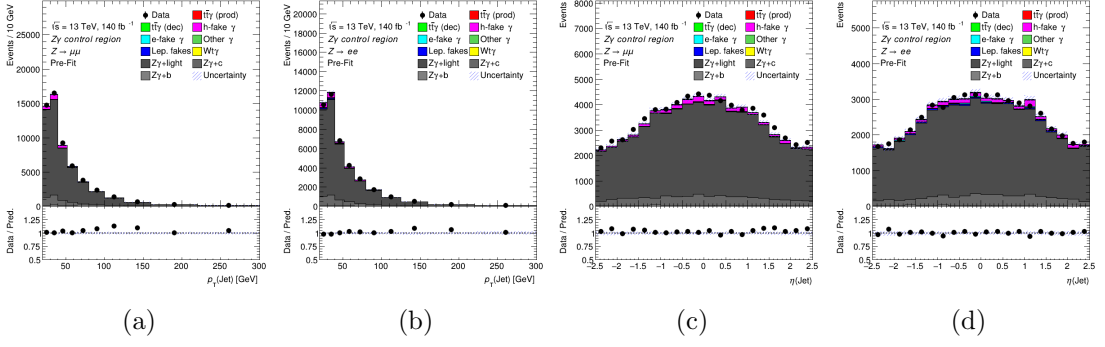


Figure C.13.: Distributions for the $Z\gamma$ control region, showing the p_T and η distributions of all jets for $\mu\mu$ ((a) and (c)) and ee ((b) and (d)) subchannels. The uncertainty band shows statistical uncertainty only.

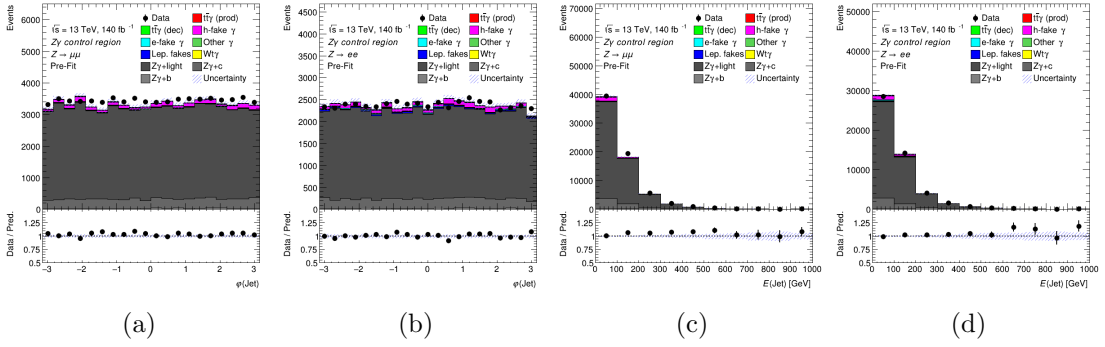


Figure C.14.: Distributions for the $Z\gamma$ control region, showing the ϕ and E distributions of all jets for $\mu\mu$ ((a) and (c)) and ee ((b) and (d)) subchannels. The uncertainty band shows statistical uncertainty only.

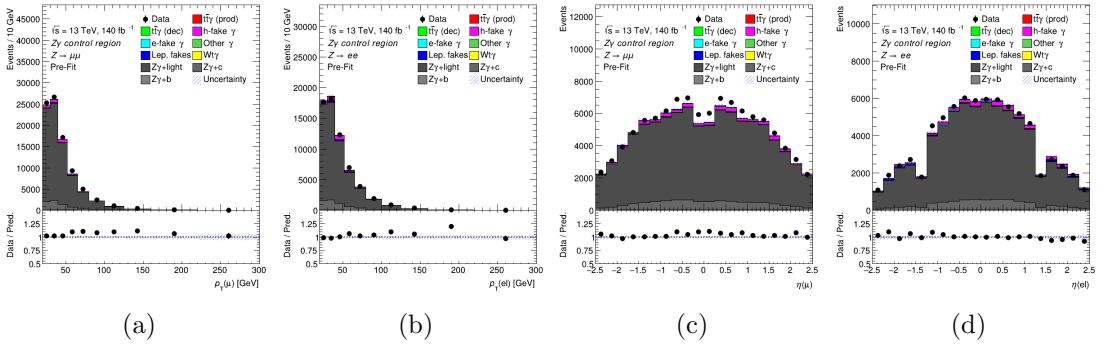


Figure C.15.: Distributions for the $Z\gamma$ control region, showing the p_T and η distributions of all leptons for $\mu\mu$ ((a) and (c)) and ee ((b) and (d)) subchannels. The uncertainty band shows statistical uncertainty only.

C.2. $Z\gamma$ control region

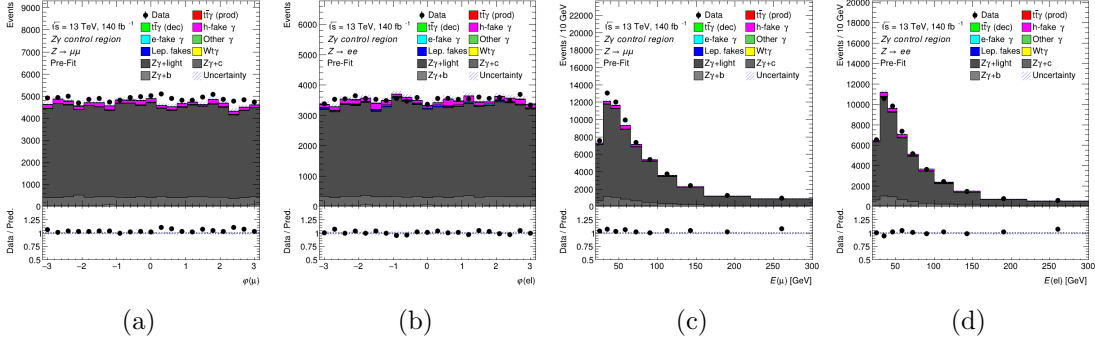


Figure C.16.: Distributions for the $Z\gamma$ control region, showing the ϕ and E distributions of all leptons for $\mu\mu$ ((a) and (c)) and ee ((b) and (d)) subchannels. The uncertainty band shows statistical uncertainty only.

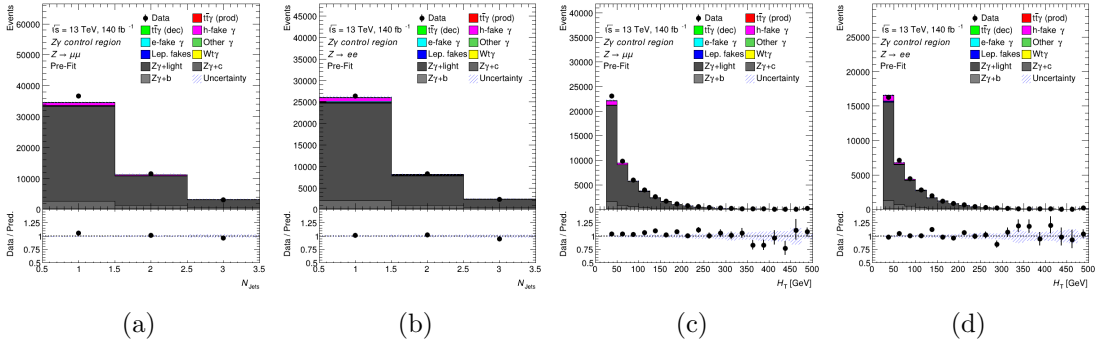


Figure C.17.: Distributions for the $Z\gamma$ control region, showing the jet multiplicity and H_T distributions for $\mu\mu$ ((a) and (c)) and ee ((b) and (d)) subchannels. The uncertainty band shows statistical uncertainty only.

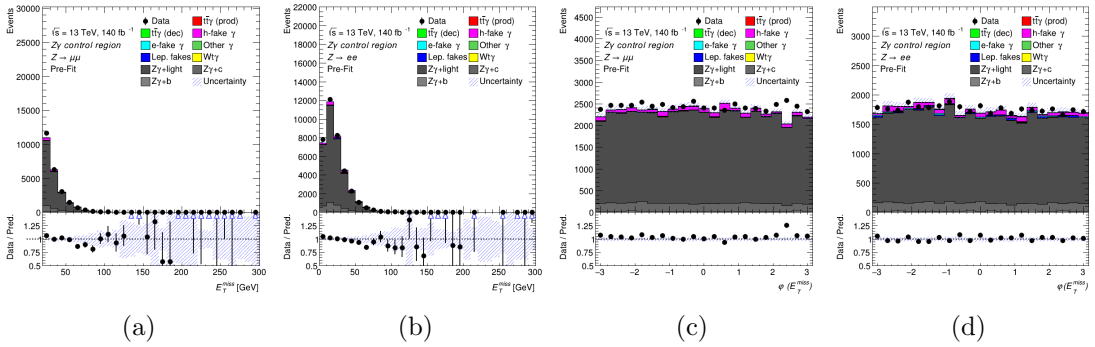


Figure C.18.: Distributions for the $Z\gamma$ control region, showing the E_T^{miss} and $\phi(E_T^{\text{miss}})$ distributions for $\mu\mu$ ((a) and (c)) and ee ((b) and (d)) subchannels. The uncertainty band shows statistical uncertainty only.

Additional plots on the neural network training

In this Appendix additional material is shown related to the neural network training. For both channels the input distributions, the separation values, distributions showing the impact of negative weights and distributions showing the comparison between training and testing are presented. For the lepton+jets channel the permutation importance is shown and the full set of correlation plots used for the definition of the fit regions.

D.1. L+jets channel

D.1.1. Input distributions

D. Additional plots on the neural network training

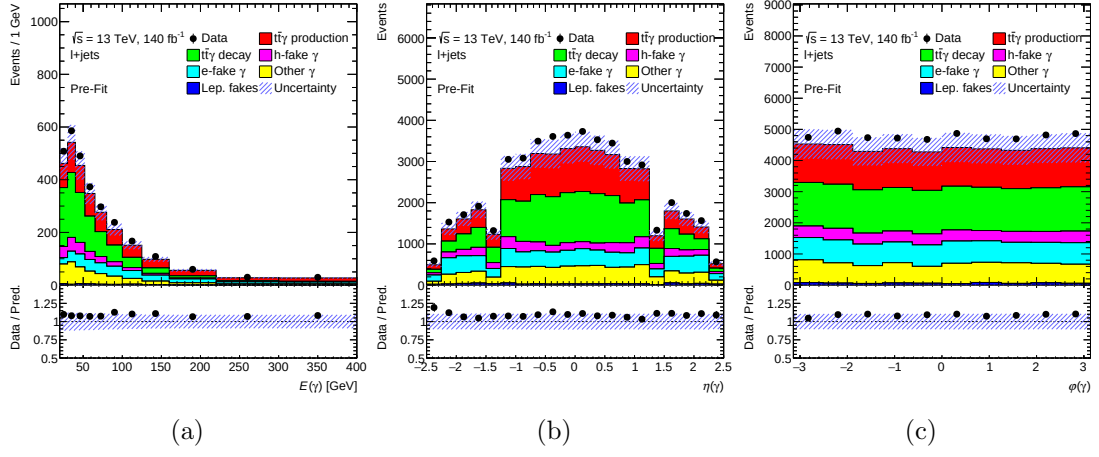


Figure D.1.: The E , η and ϕ distributions (from left to right) of the photon inserted into the NN training. The uncertainty band corresponds to the total uncertainty, including stat. and syst. uncertainties.

D.1. $L+\text{jets}$ channel

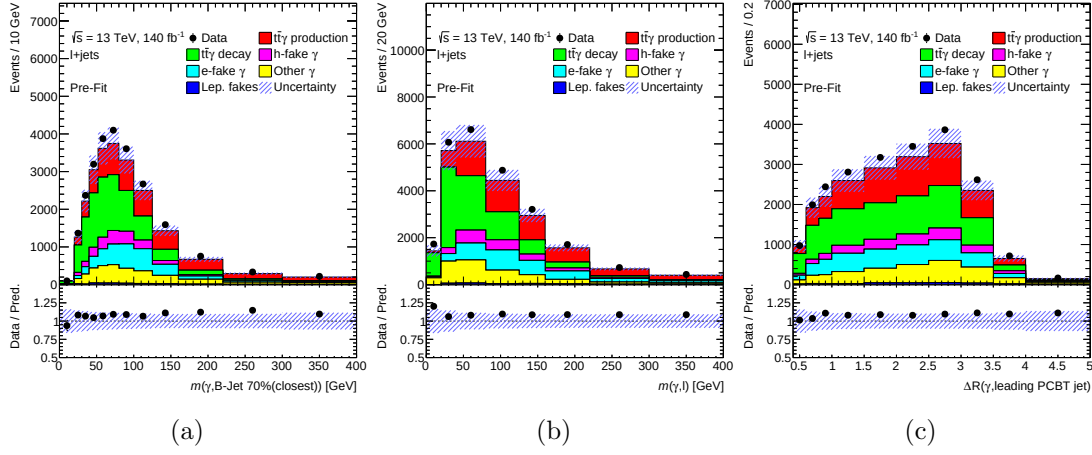


Figure D.2.: The invariant mass distributions of the photon and leading b-jet, and the photon and the lepton, as well as the ΔR distribution between the photon and the leading PCBT jet (from left to right) inserted into the NN training. The uncertainty band corresponds to the total uncertainty, including stat. and syst. uncertainties.

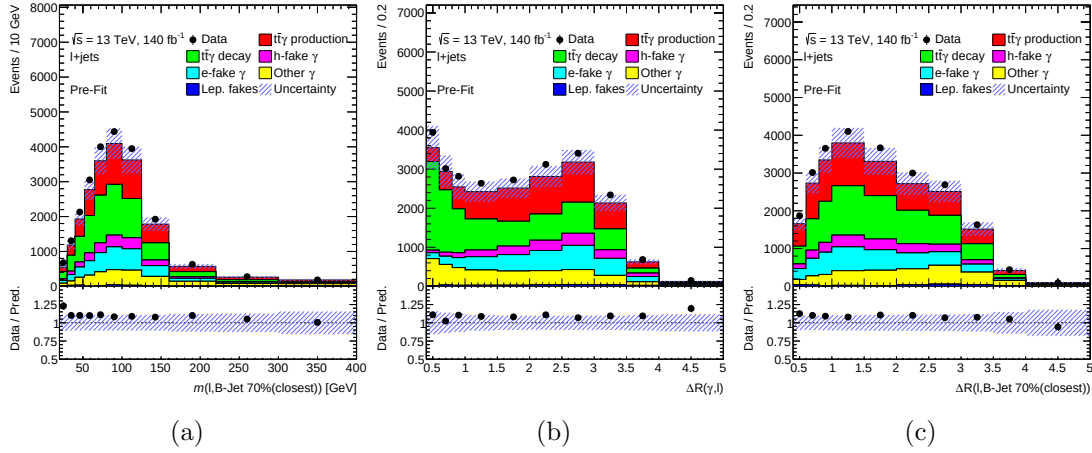


Figure D.3.: The invariant mass distribution of the lepton and the closest b-jet, and the ΔR distributions between the photon and the leading b-jet, and between the lepton and the closest b-jet (from left to right) inserted into the NN training. The uncertainty band corresponds to the total uncertainty, including stat. and syst. uncertainties.

D. Additional plots on the neural network training

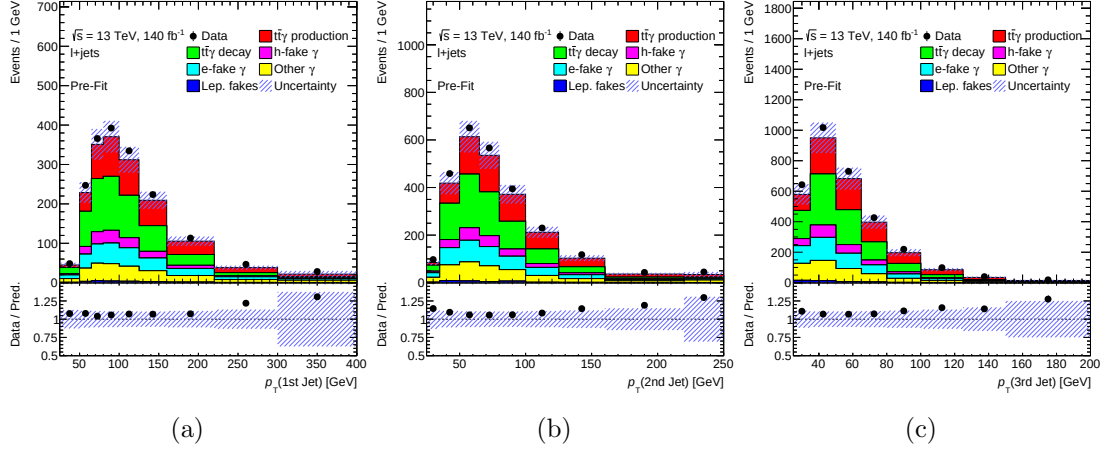


Figure D.4.: The p_T distributions of the first, second and third jet (from left to right) inserted into the NN training. The uncertainty band corresponds to the total uncertainty, including stat. and syst. uncertainties.

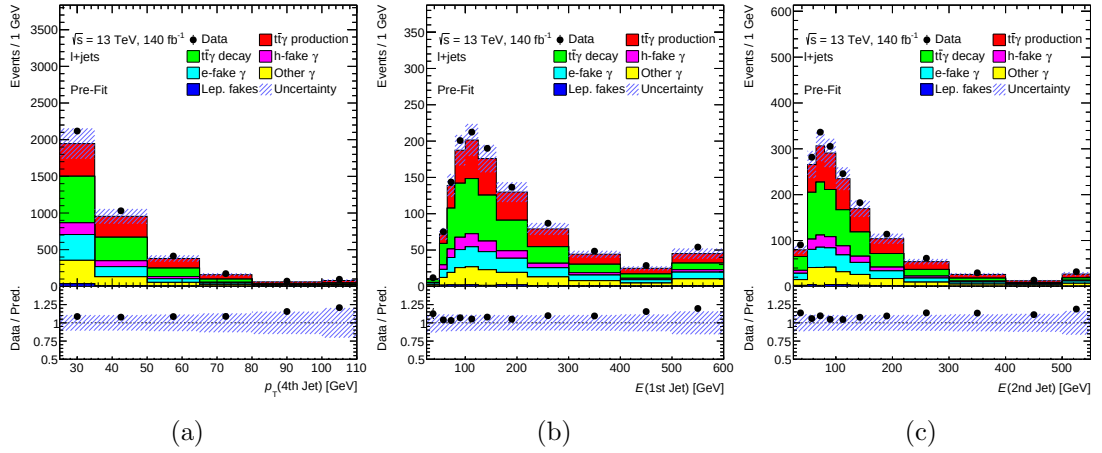


Figure D.5.: The p_T distribution of the fourth jet and the E distributions of the first and second jet (from left to right) inserted into the NN training. The uncertainty band corresponds to the total uncertainty, including stat. and syst. uncertainties.

D.1. L +jets channel

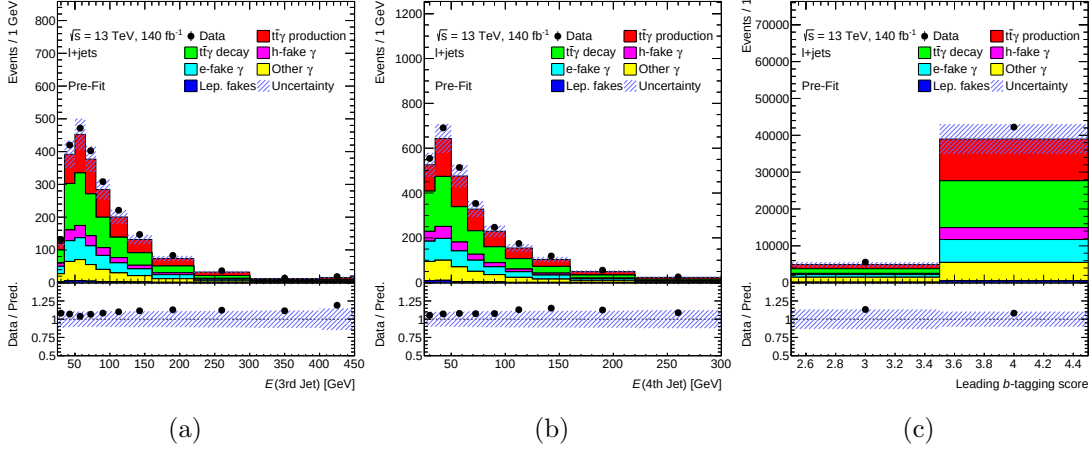


Figure D.6.: The E distributions of third and fourth jet and the PCBT distribution of the leading PCBT jet (from left to right) inserted into the NN training. The uncertainty band corresponds to the total uncertainty, including stat. and syst. uncertainties.

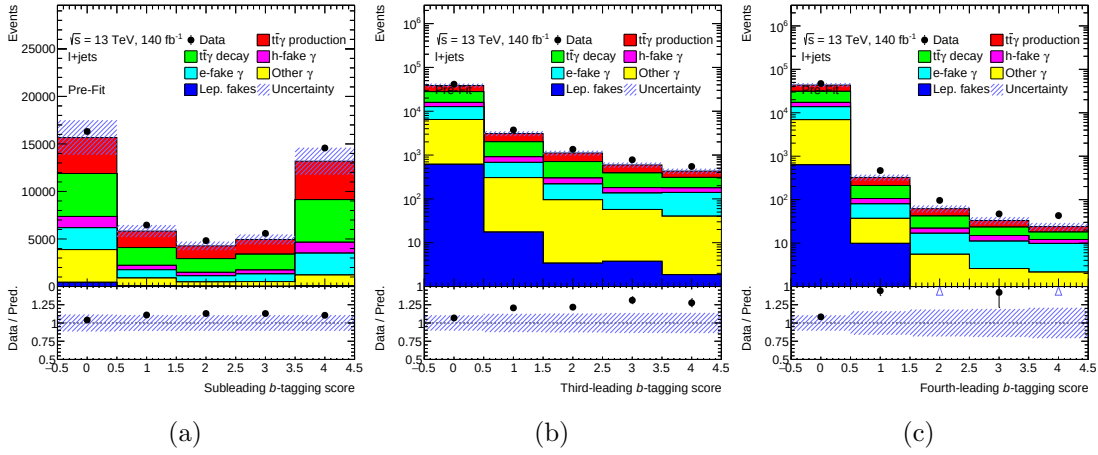


Figure D.7.: The PCBT distributions of the sub-, third- and fourth-leading PCBT jet (from left to right) inserted into the NN training. The uncertainty band corresponds to the total uncertainty, including stat. and syst. uncertainties.

D. Additional plots on the neural network training

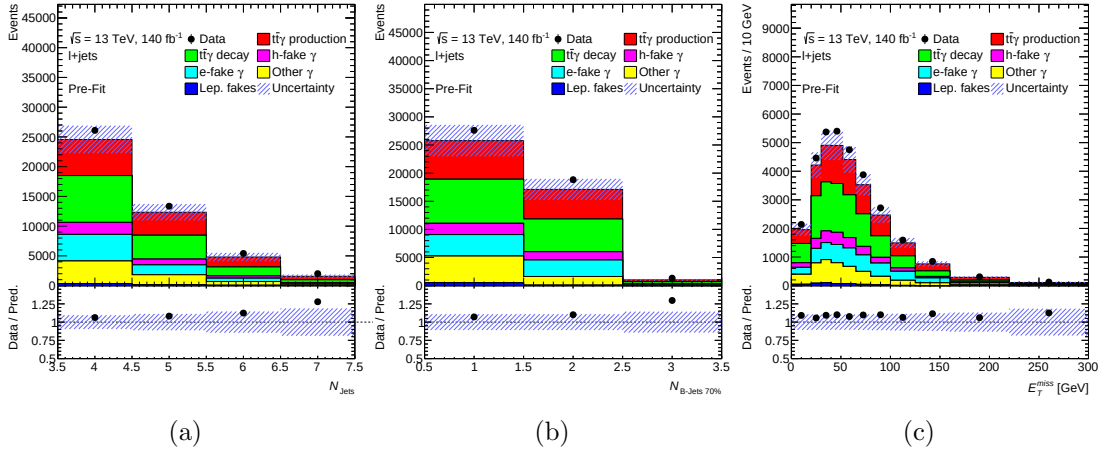


Figure D.8.: The multiplicity distributions of jets and b -jets per event and the E_T^{miss} distribution (from left to right) inserted into the NN training. The uncertainty band corresponds to the total uncertainty, including stat. and syst. uncertainties.

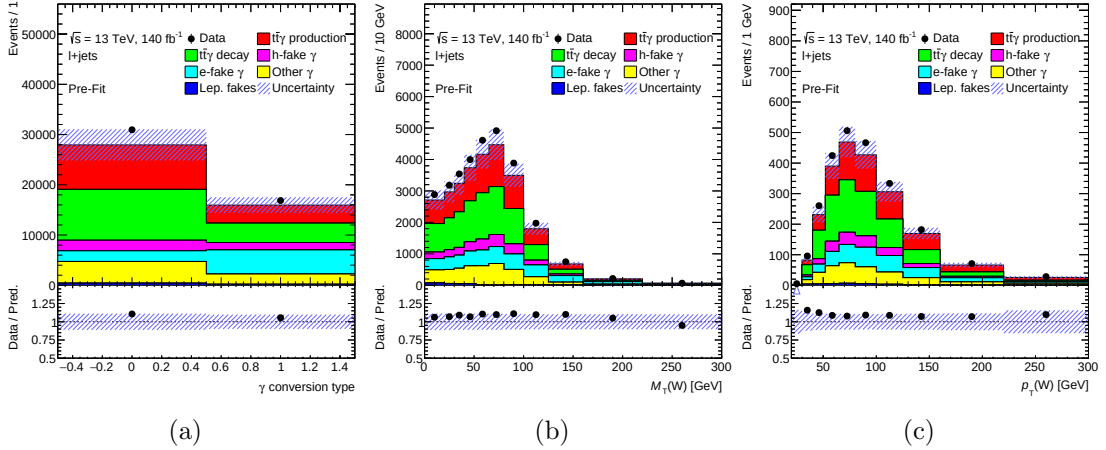


Figure D.9.: The photon conversion type (0 means unconverted, 1 means converted) distribution and the transverse mass and p_T distributions of the leptonically decaying W-boson (from left to right) inserted into the NN training. The uncertainty band corresponds to the total uncertainty, including stat. and syst. uncertainties.

D.1. $L+jets$ channel

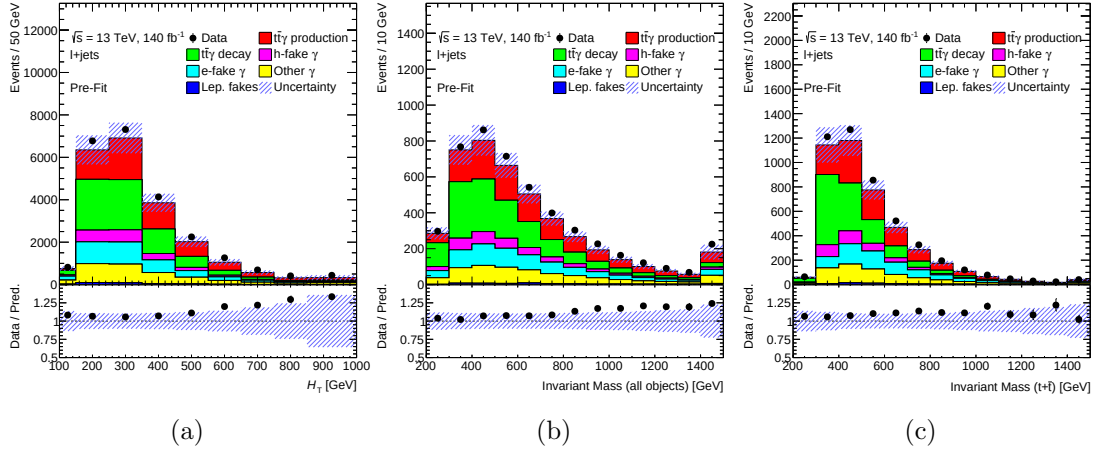


Figure D.10.: The H_T distribution, the invariant mass of all reconstructed objects and sum of the invariant masses of top and anti-top quark of the best permutation to reconstruct them (from left tot right) inserted into the NN training. The uncertainty band corresponds to the total uncertainty, including stat. and syst. uncertainties.

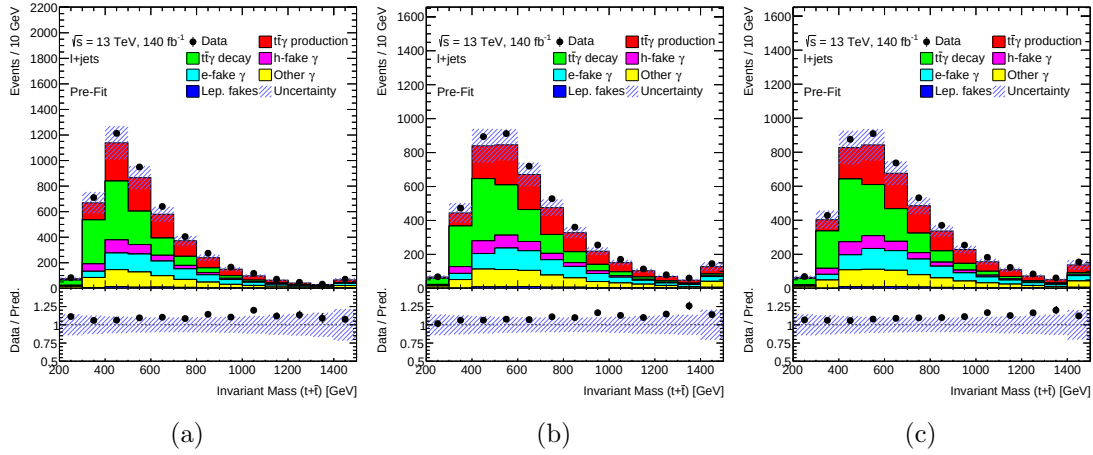


Figure D.11.: The distributions for the sum of the invariant masses of top- and anti-top-quark of the second, third and fourth best permutation to reconstruct them (from left to right) inserted into the NN training. The uncertainty band corresponds to the total uncertainty, including stat. and syst. uncertainties.

D. Additional plots on the neural network training

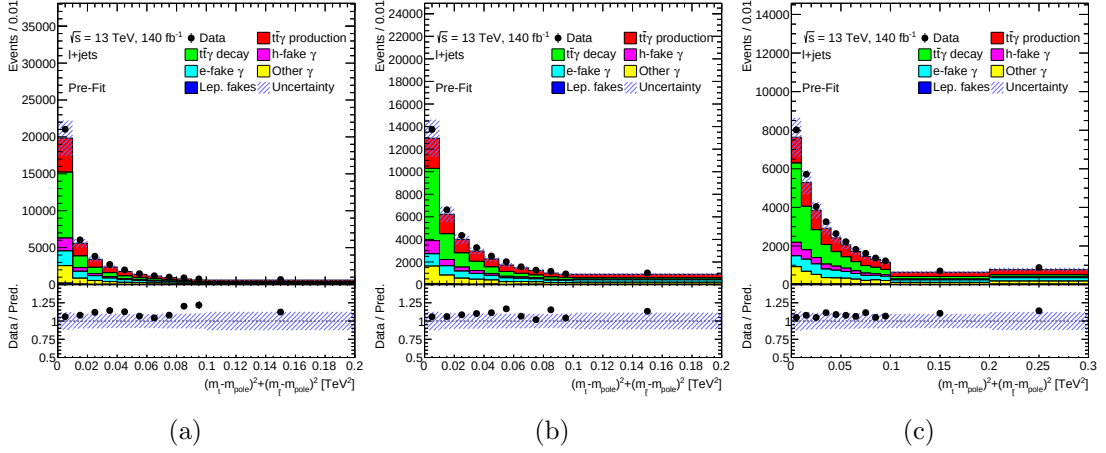


Figure D.12.: The distributions for the sum of the squared differences between the pole-mass and the reconstructed masses of top- and anti-top-quark for the best, second and third best permutation (from left tot right) inserted into the NN training. The uncertainty band corresponds to the total uncertainty, including stat. and syst. uncertainties.

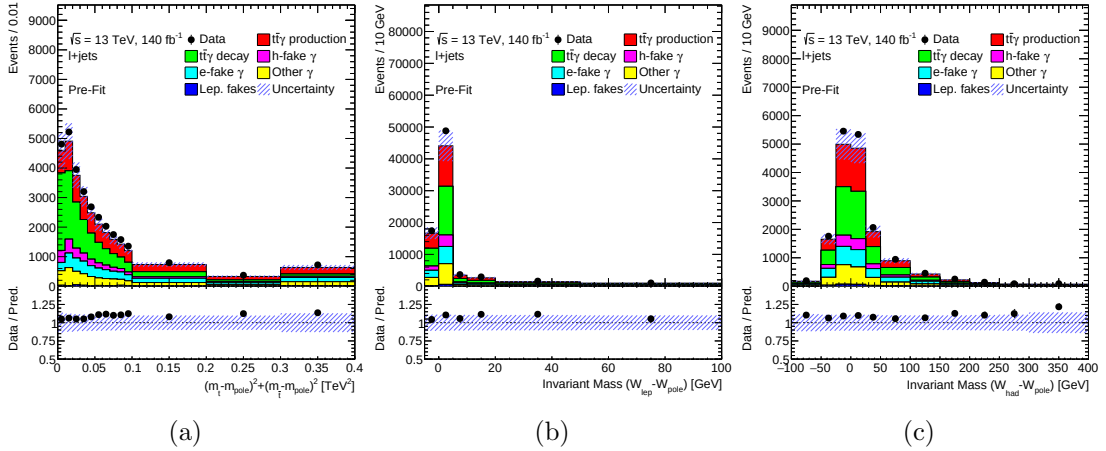
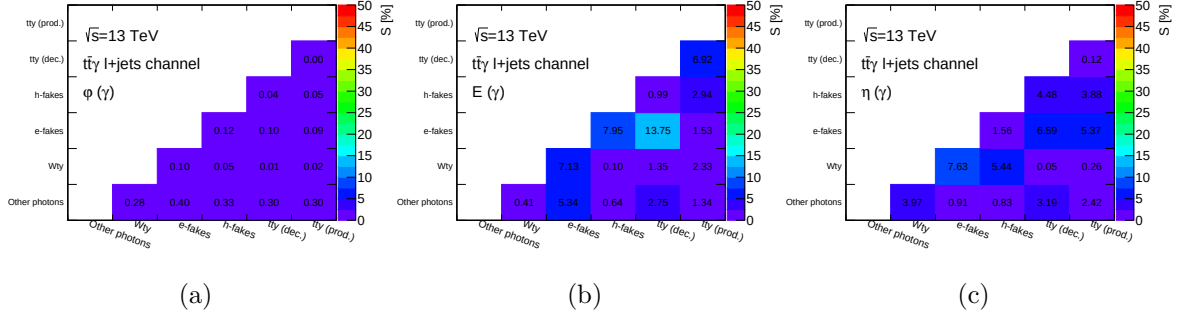
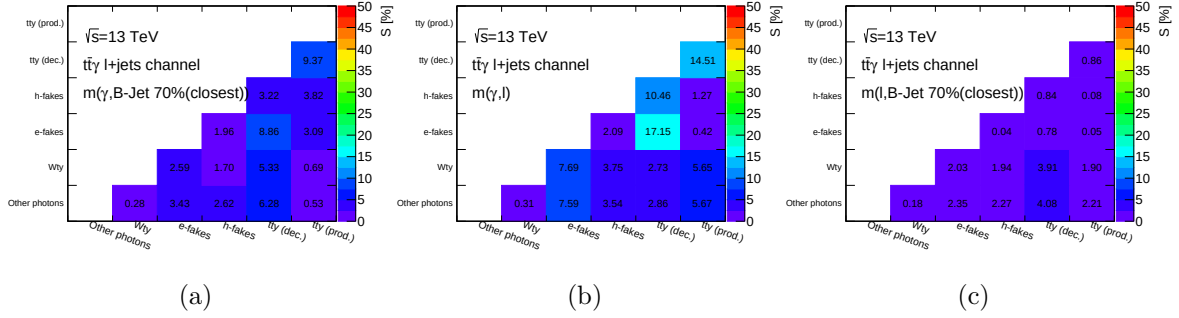
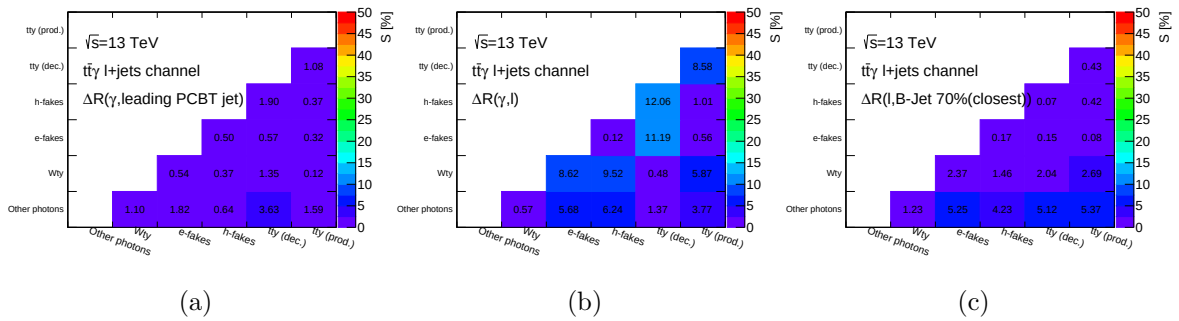


Figure D.13.: The distributions for the sum of the squared differences between the pole-mass and the reconstructed masses of top- and anti-top-quark for the fourth best permutation and for the differences between the leptonically and hadronically reconstrcuted W -masses and the W pole mass (from left to right) inserted into the NN training. The uncertainty band corresponds to the total uncertainty, including stat. and syst. uncertainties.

D.1.2. Separation plots

Figure D.14.: The separation values for the ϕ , E and η distributions of the photon (from left to right) for the different event categories.Figure D.15.: The separation values for the invariant mass distributions of the photon and the closest b -jet, of the photon and the lepton and of the lepton and the closest b -jet (from left to right) for the different event categories.Figure D.16.: The separation values for the ΔR distributions of the photon and the leading PCBT jet, of the photon and the lepton and of the lepton and the closest b -jet (from left to right) for the different event categories.

D. Additional plots on the neural network training

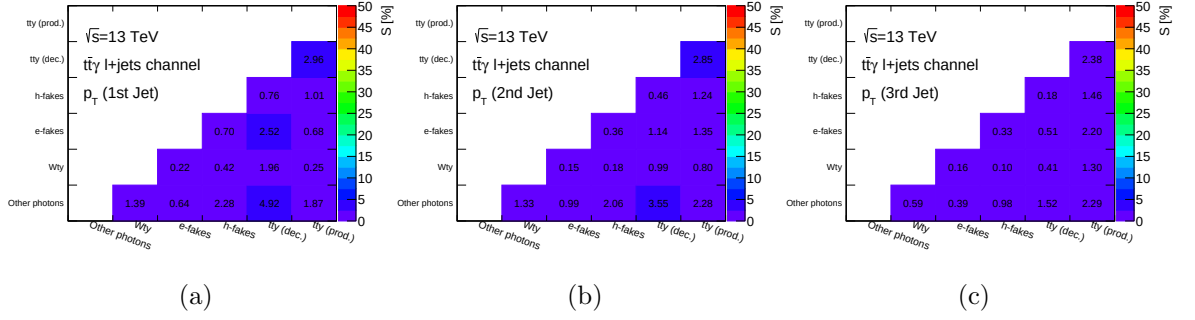


Figure D.17.: The separation values for the p_T distributions of the first, second and third jet (from left to right) for the different event categories.

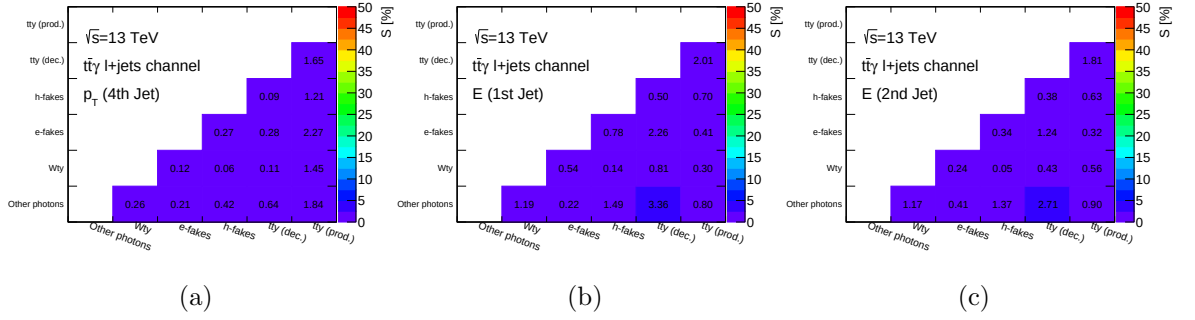


Figure D.18.: The separation values for the p_T distribution of the fourth jet and the E distributions for the first and second jet (from left to right) for the different event categories.

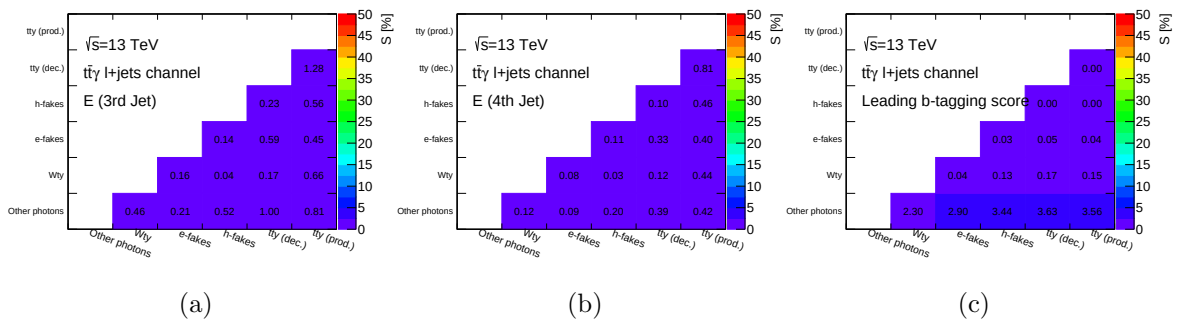


Figure D.19.: The separation values for the E distributions of the third and fourth jet and the PCBT distribution for the leading PCBT jet (from left to right) for the different event categories.

D.1. $L+jets$ channel

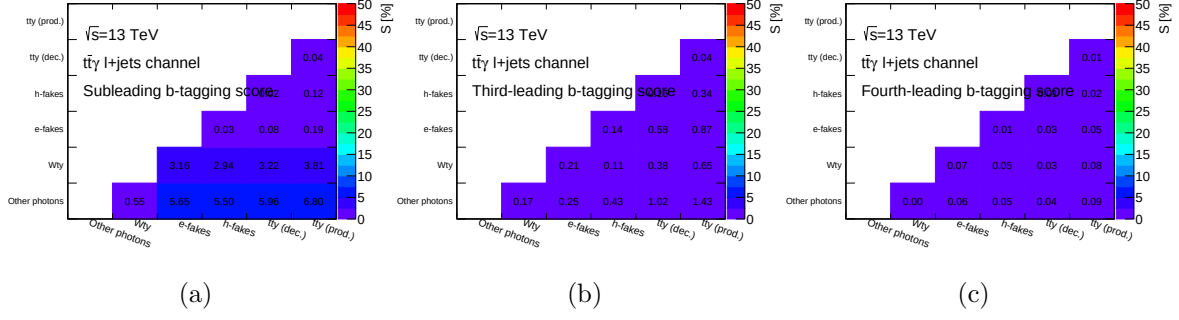


Figure D.20.: The separation values for the PCBT distributions for the sub-, third- and fourth-leading PCBT jet (from left to right) for the different event categories.

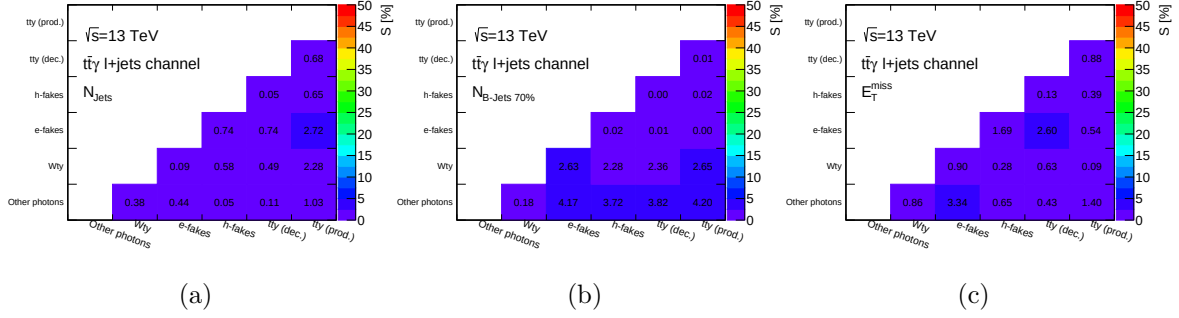


Figure D.21.: The separation values for the distributions showing the number of jets and b -jets (at the 70% WP) and for the E_T^{miss} distribution (from left to right) for the different event categories.

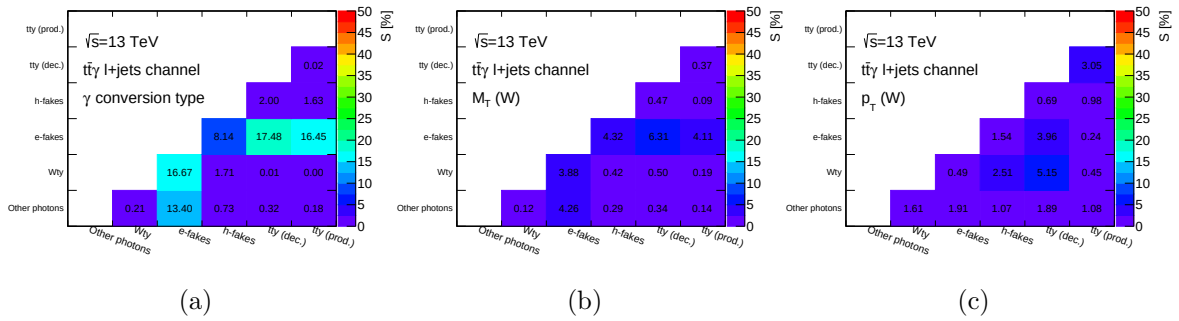


Figure D.22.: The separation values for the photon conversion type distribution and the transverse mass and momentum distributions of the leptonic W (from left to right) for the different event categories.

D. Additional plots on the neural network training

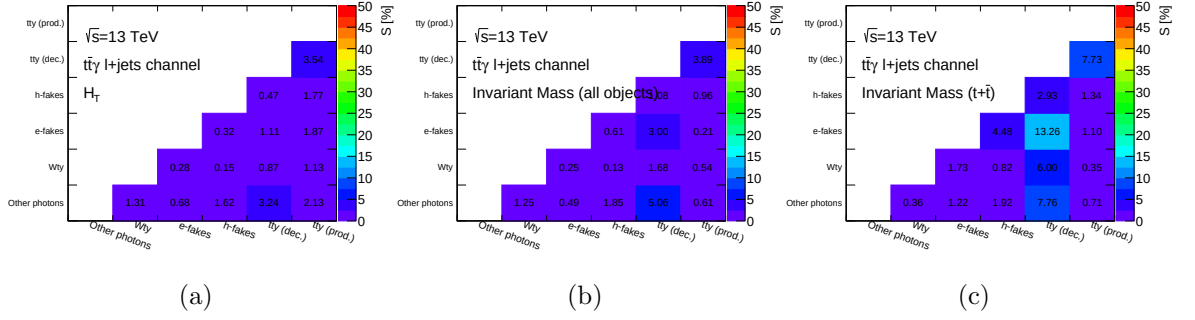


Figure D.23.: The separation values for the H_T distribution, for the distributions of the invariant mass of all reconstructed objects and for the distribution of the sum of the invariant masses of top and anti-top quark for the best permutation (from left tot right) for the different event categories.

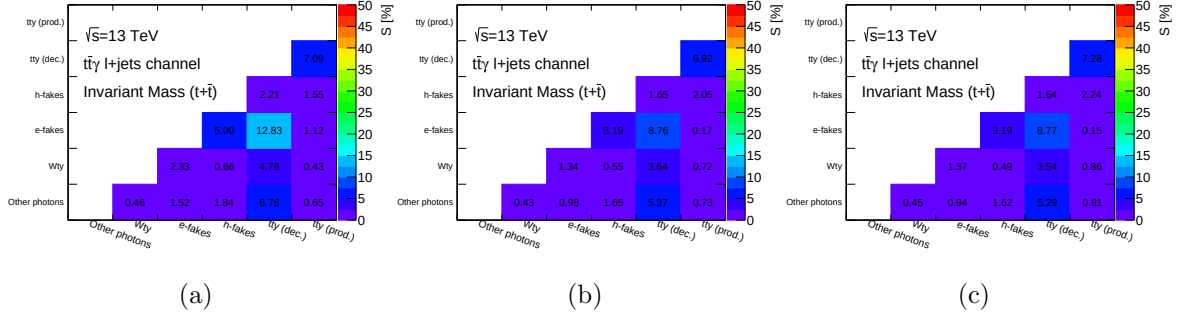


Figure D.24.: The separation values for the distributions of the sum of the invariant masses of top- and anti-top-quark for the second, third and fourth best permutation (from left to right) for the different event categories.

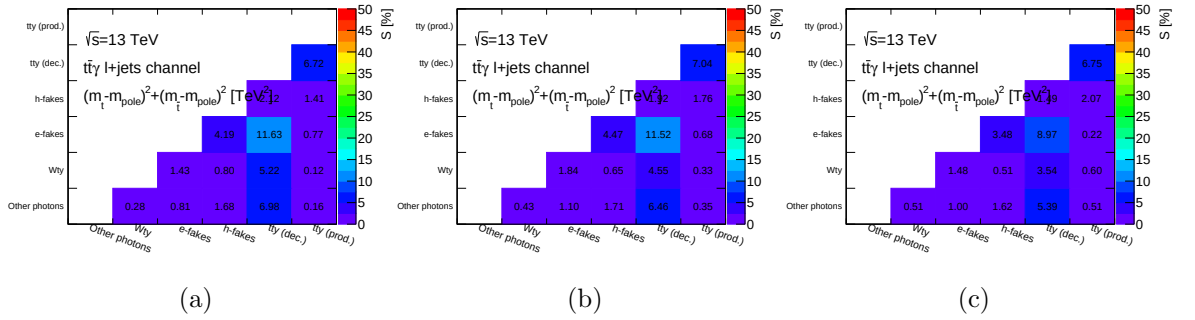


Figure D.25.: The separation values for the distributions of the sum of the squared differences between the pole-mass and the reconstructed masses of top and anti-top-quark for the best, second best and third best permutation (from left to right) for the different event categories.

D.1. $L+jets$ channel

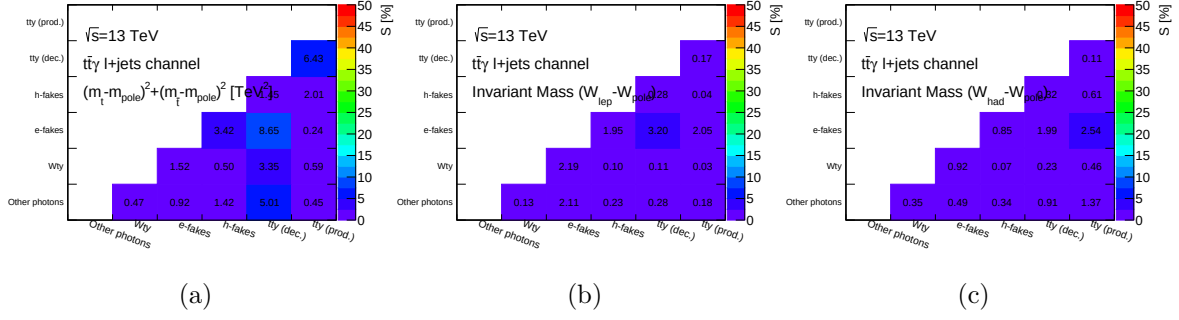


Figure D.26.: The separation values for the distributions of the sum of the squared differences between the pole-mass and the reconstructed masses of top and anti-top-quark for the fourth best permutation and for the differences between the reconstructed W -boson mass and the pole mass for the (hypothetical) leptonically and (hypothetical) hadronically decaying W -boson (from left to right) for the different event categories.

D.1.3. Negative weight comparisons

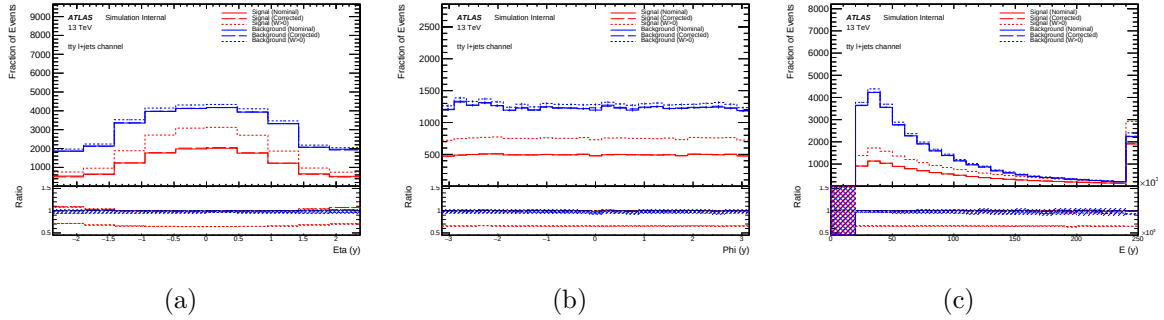


Figure D.27.: The impact of negative weights on the η , ϕ and E distributions of the photon (from left to right) inserted into the NN training.

D. Additional plots on the neural network training

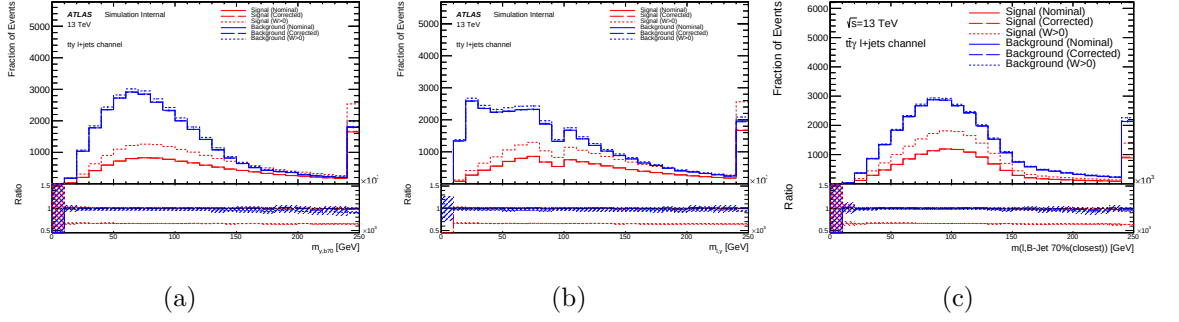


Figure D.28.: The impact of negative weights on the invariant mass distributions of the photon and closest b -jet, of the photon and the lepton and of the lepton and the closest b -jet (from left to right) inserted into the NN training.

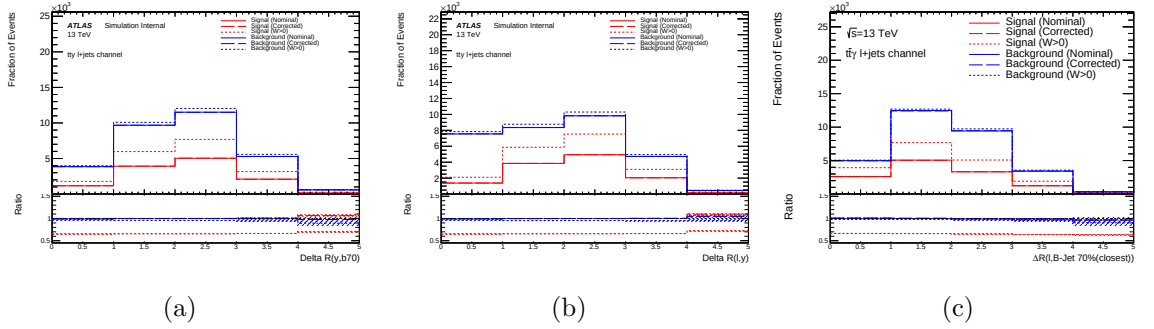


Figure D.29.: The impact of negative weights on the ΔR distributions of the photon and the leading PCBT jet, of the photon and the lepton and of the lepton and the closest b -jet (from left to right) inserted into the NN training.

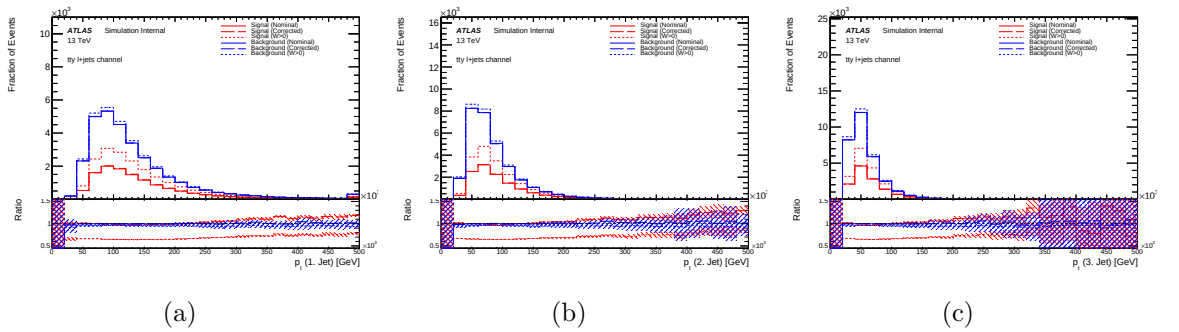


Figure D.30.: The impact of negative weights on the p_T distributions of the first, second and third jet (from left to right) inserted into the NN training.

D.1. $L+jets$ channel

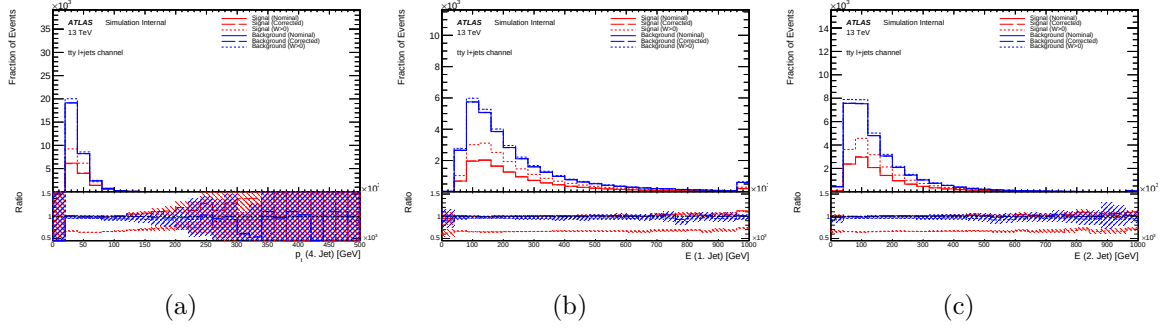


Figure D.31.: The impact of negative weights on the p_T distribution of the fourth jet and on the E distributions of the first and second jet (from left to right) inserted into the NN training.

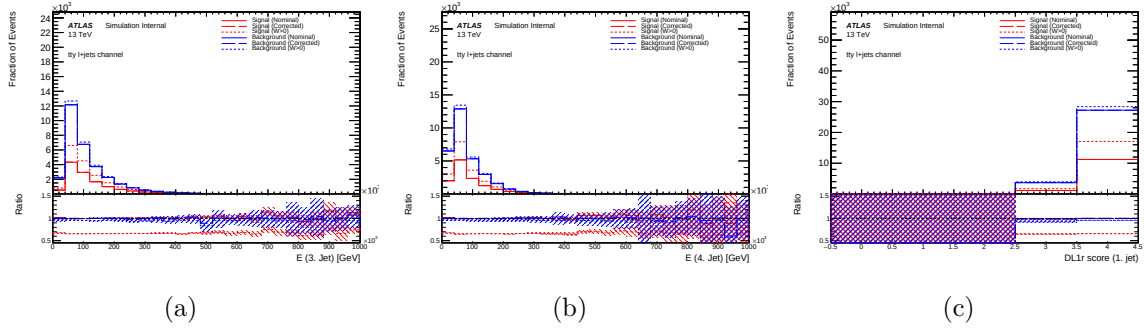


Figure D.32.: The impact of negative weights on the E distributions of the third and fourth jet and on PCBT distribution of the leading PCBT jet (from left to right) inserted into the NN training.

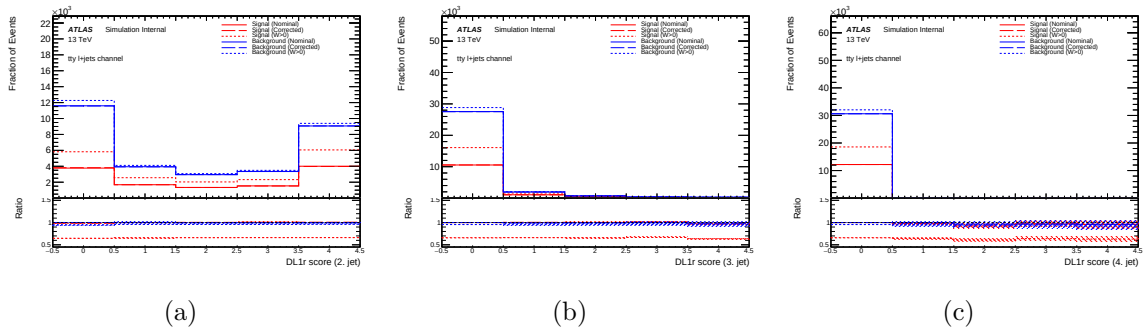


Figure D.33.: The impact of negative weights on the PCBT distributions of the sub-, third- and fourth-leading PCBT jet (from left to right) inserted into the NN training.

D. Additional plots on the neural network training

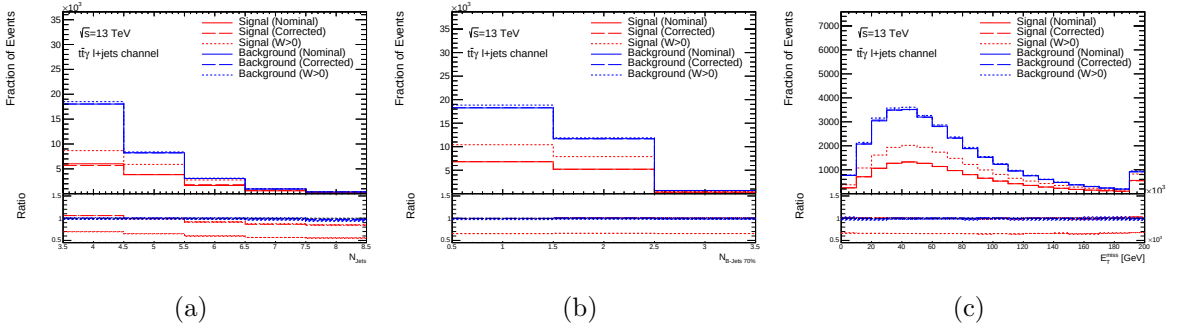


Figure D.34.: The impact of negative weights on the distributions of the number of jets and b -jets per event and on the E_T^{miss} distribution (from left to right) inserted into the NN training.

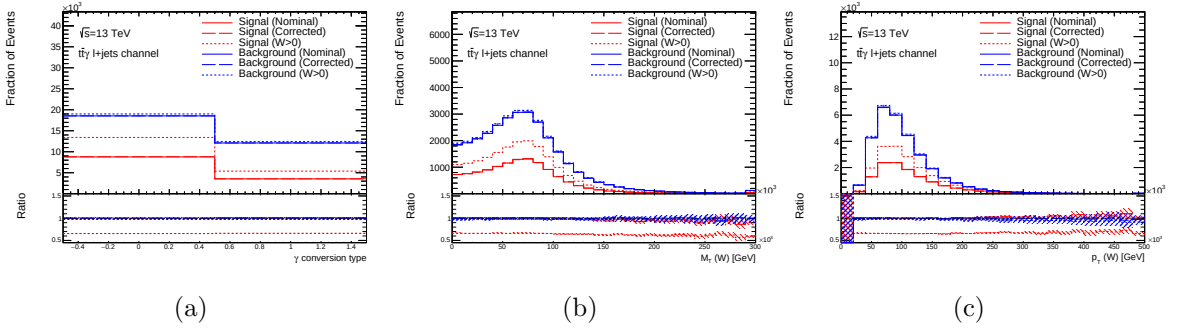


Figure D.35.: The impact of negative weights on the photon conversion type (0 means unconverted, 1 means converted) distribution and on the distributions of the transverse mass and momentum of the leptonically decaying W -boson (from left to right) inserted into the NN training.

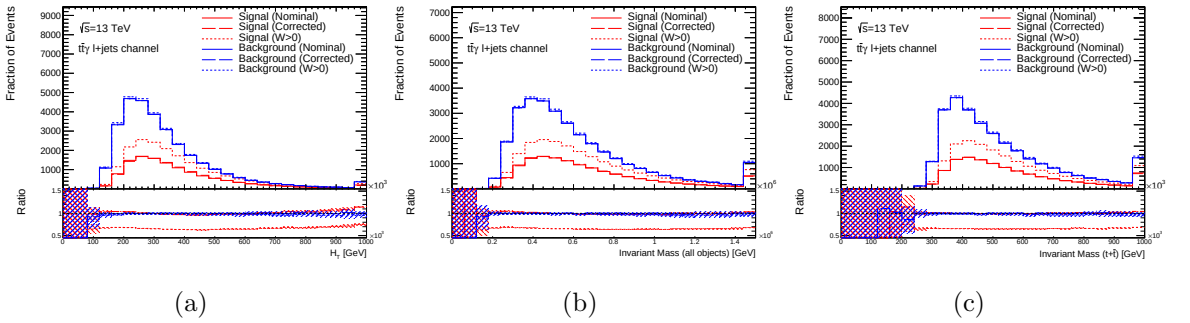


Figure D.36.: The impact of negative weights on the H_T distribution and on the distribution of the invariant mass of all reconstructed objects, as well as on the distribution of the sum of the invariant masses of top and anti-top quark of the best permutation to reconstruct them (from left to right) inserted into the NN training.

D.1. $L+jets$ channel

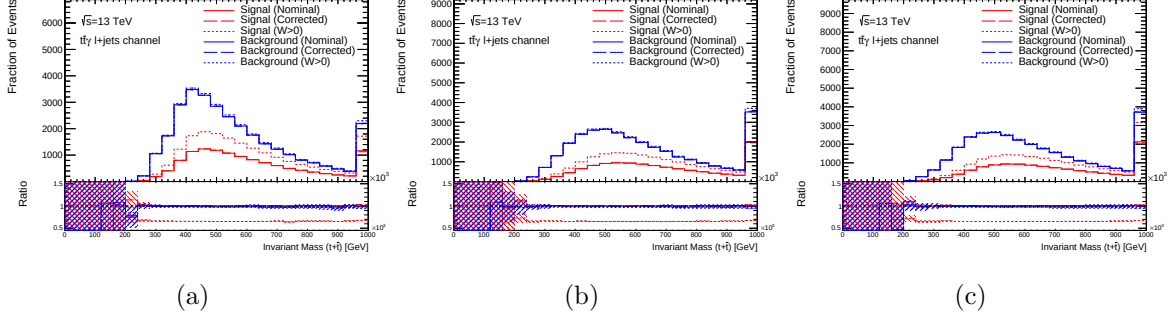


Figure D.37.: The impact of negative weights on the distributions of the sum of the invariant masses of top- and anti-top-quark of second, third and fourth best permutation to reconstruct them (from left to right) inserted into the NN training.

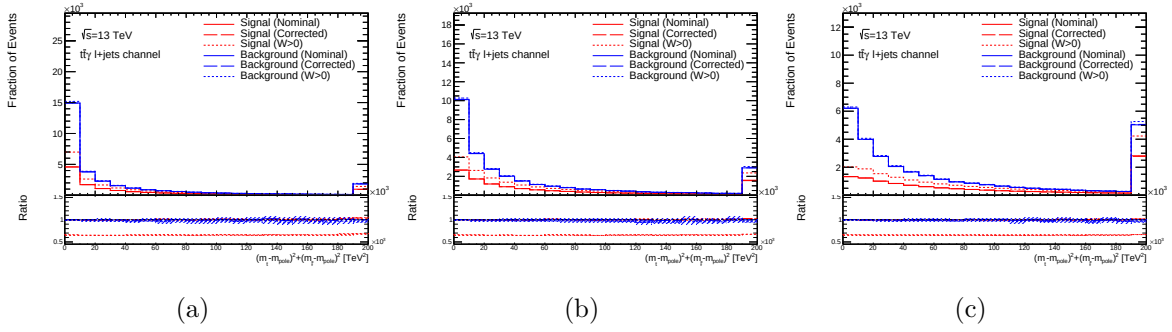


Figure D.38.: The impact of negative weights on the distributions of the sum of the squared differences between the pole-mass and the reconstructed mass of top and anti-top-quark for the best, second best and third best permutation (from left to right) inserted into the NN training.

D. Additional plots on the neural network training

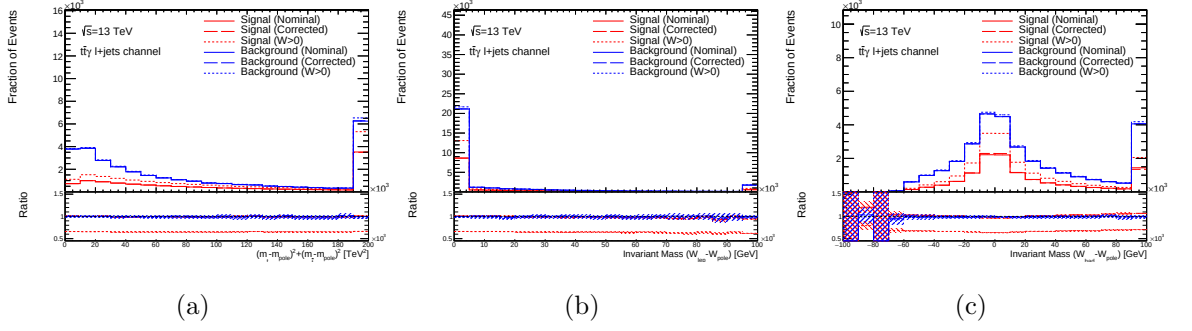


Figure D.39.: The impact of negative weights on the distribution of the sum of the squared differences between the pole-mass and the reconstructed masses of top and anti-top-quark for the fourth best permutation and on the distributions of the invariant masses of the leptonically and hadronically decaying W -boson shifted by the W -pole-mass (from left to right) inserted into the NN training.

D.1.4. Comparison of training and testing output distributions

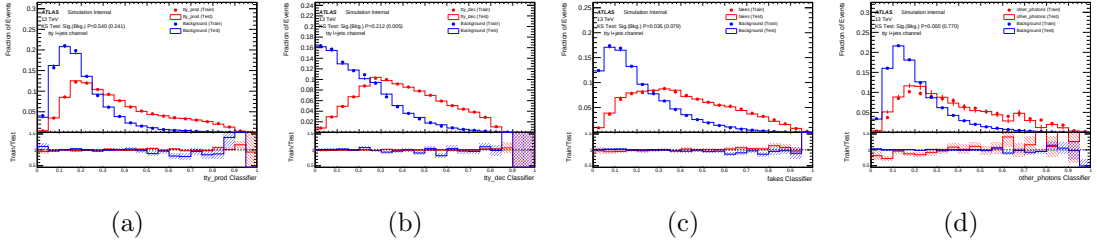


Figure D.40.: Output distributions for signal ((a): $t\bar{t}\gamma$ (prod.), (b): $t\bar{t}\gamma$ (dec.), (c): photon fakes and (d):other photons) and background (everything not considered as signal in the respective plot) for the $t\bar{t}\gamma$ (prod.) classifier (a), the $t\bar{t}\gamma$ (dec.) classifier (b), the photon fakes classifier (c) and the other photon classifier (d) for fold 1.

D.1. L +jets channel

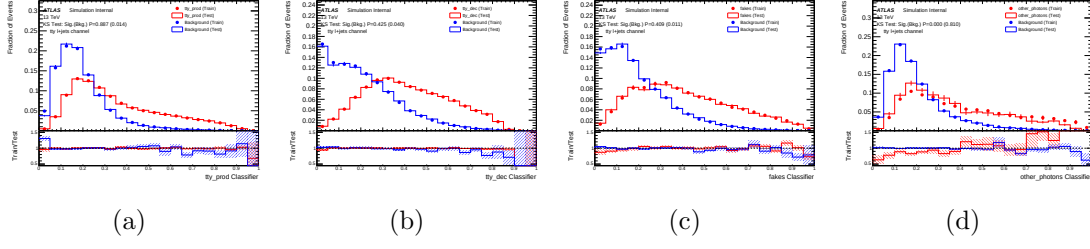


Figure D.41.: Output distributions for signal ((a): $t\bar{t}\gamma$ (prod.), (b): $t\bar{t}\gamma$ (dec.), (c): photon fakes and (d):other photons) and background (everything not considered as signal in the respective plot) for the $t\bar{t}\gamma$ (prod.) classifier (a), the $t\bar{t}\gamma$ (dec.) classifier (b), the photon fakes classifier (c) and the other photon classifier (d) for fold 2.

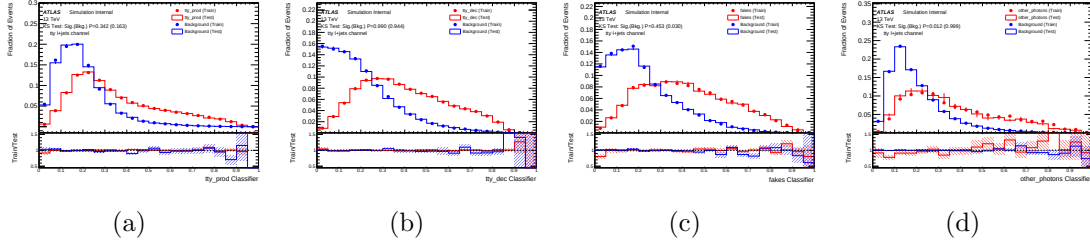


Figure D.42.: Output distributions for signal ((a): $t\bar{t}\gamma$ (prod.), (b): $t\bar{t}\gamma$ (dec.), (c): photon fakes and (d):other photons) and background (everything not considered as signal in the respective plot) for the $t\bar{t}\gamma$ (prod.) classifier (a), the $t\bar{t}\gamma$ (dec.) classifier (b), the photon fakes classifier (c) and the other photon classifier (d) for fold 3.

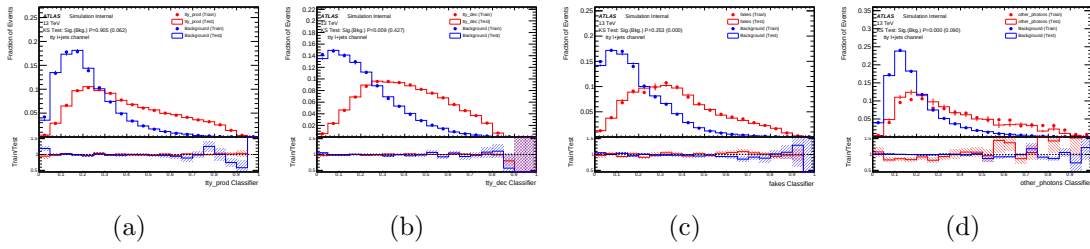


Figure D.43.: Output distributions for signal ((a): $t\bar{t}\gamma$ (prod.), (b): $t\bar{t}\gamma$ (dec.), (c): photon fakes and (d):other photons) and background (everything not considered as signal in the respective plot) for the $t\bar{t}\gamma$ (prod.) classifier (a), the $t\bar{t}\gamma$ (dec.) classifier (b), the photon fakes classifier (c) and the other photon classifier (d) for fold 4.

D. Additional plots on the neural network training

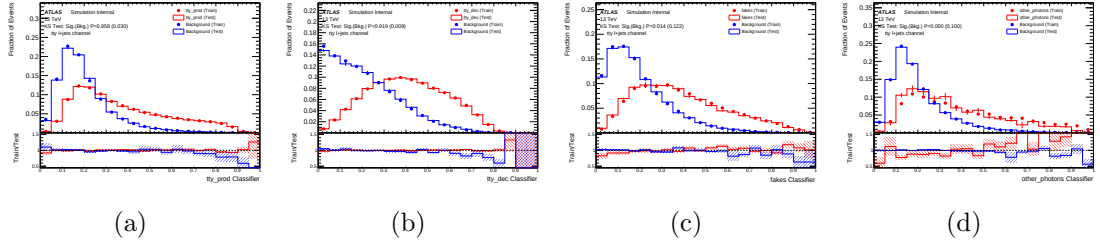
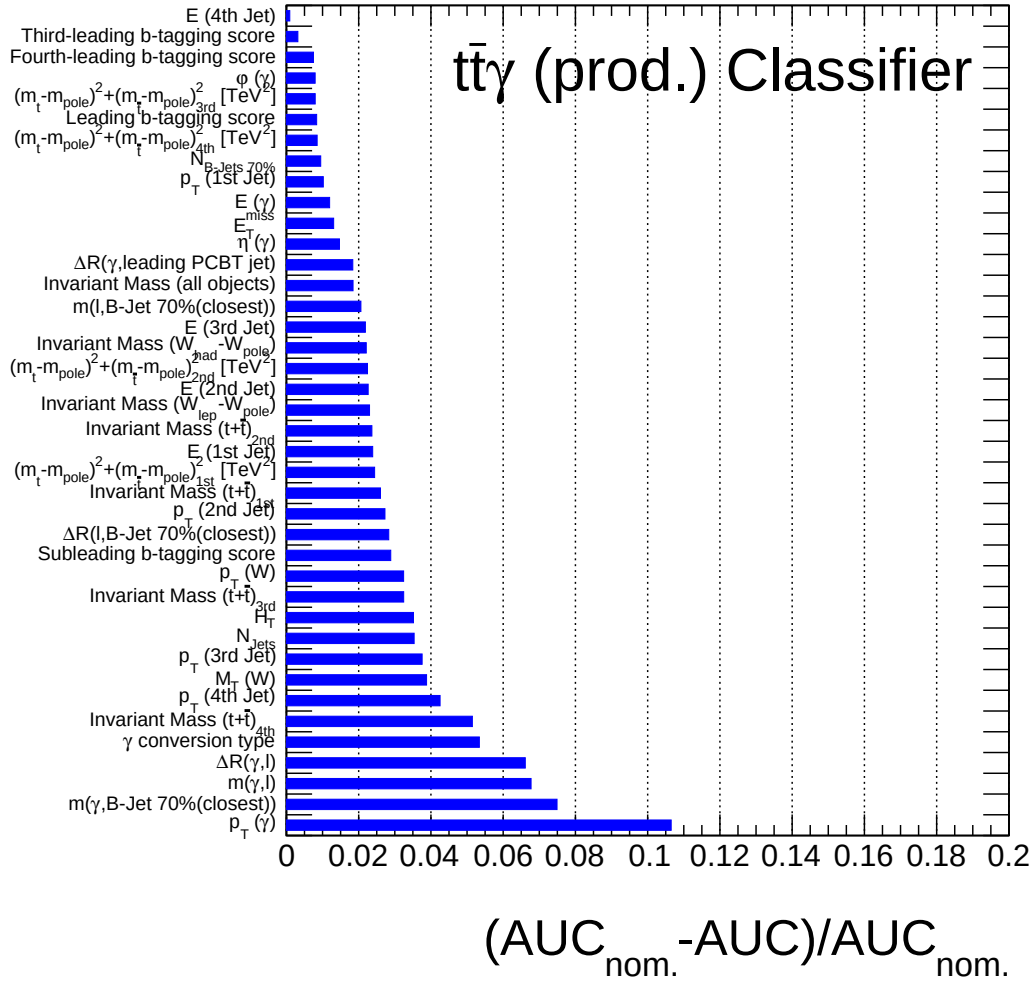


Figure D.44.: Output distributions for signal ((a): $t\bar{t}\gamma$ (prod.), (b): $t\bar{t}\gamma$ (dec.), (c): photon fakes and (d):other photons) and background (everything not considered as signal in the respective plot) for the $t\bar{t}\gamma$ (prod.) classifier (a), the $t\bar{t}\gamma$ (dec.) classifier (b), the photon fakes classifier (c) and the other photon classifier (d) for fold 5.

D.1.5. Permutation importance

Figure D.45.: Permutation importance for the $t\bar{t}\gamma$ (prod.) classifier.

D. Additional plots on the neural network training

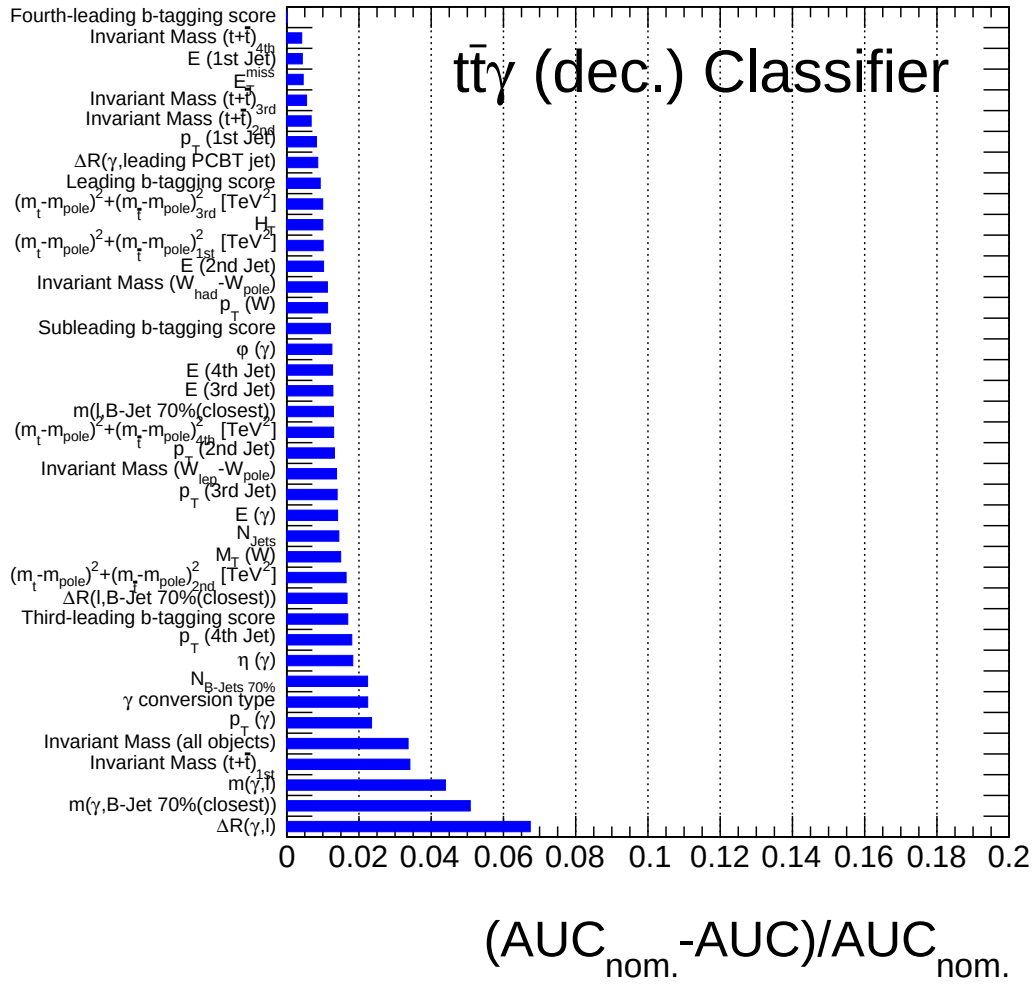


Figure D.46.: Permutation importance for the $t\bar{t}\gamma$ (dec.) classifier.

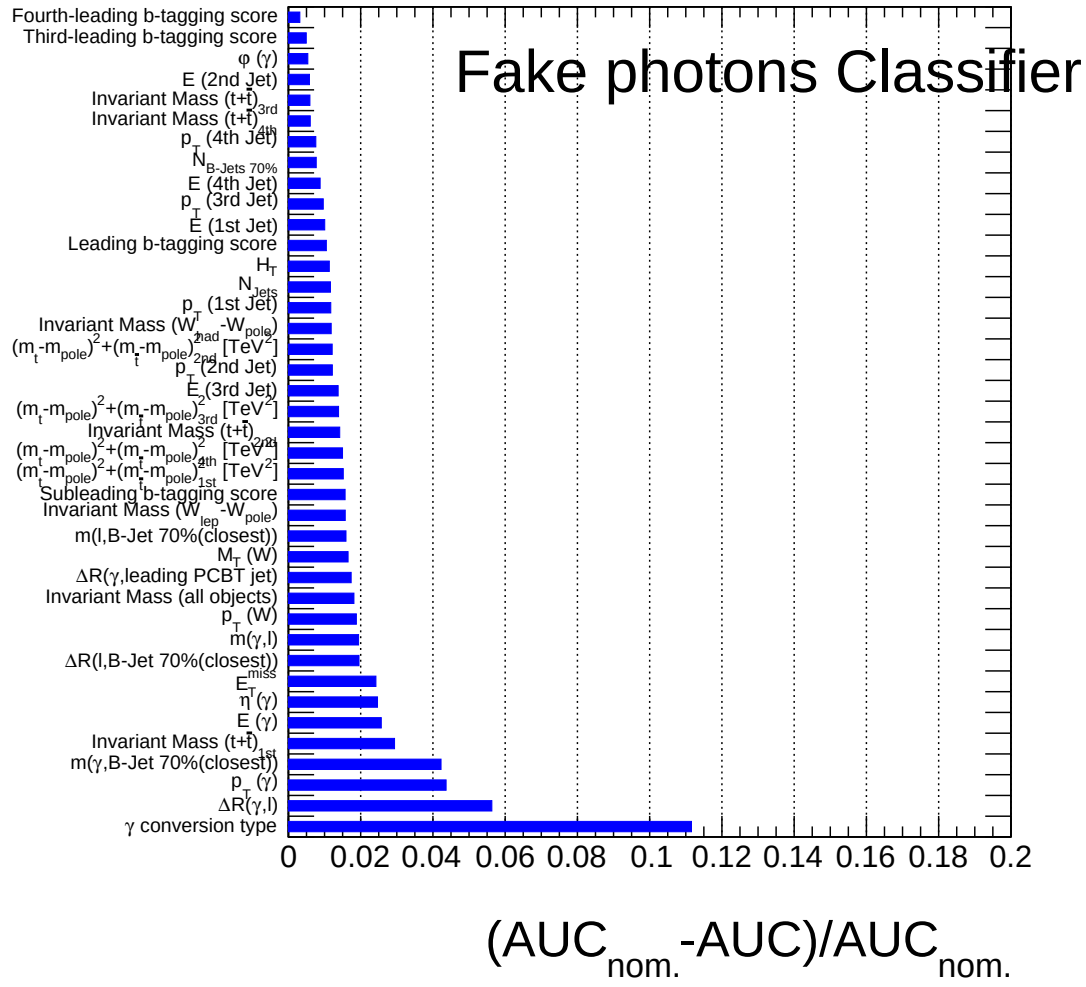


Figure D.47.: Permutation importance for the photon fakes classifier.

D. Additional plots on the neural network training

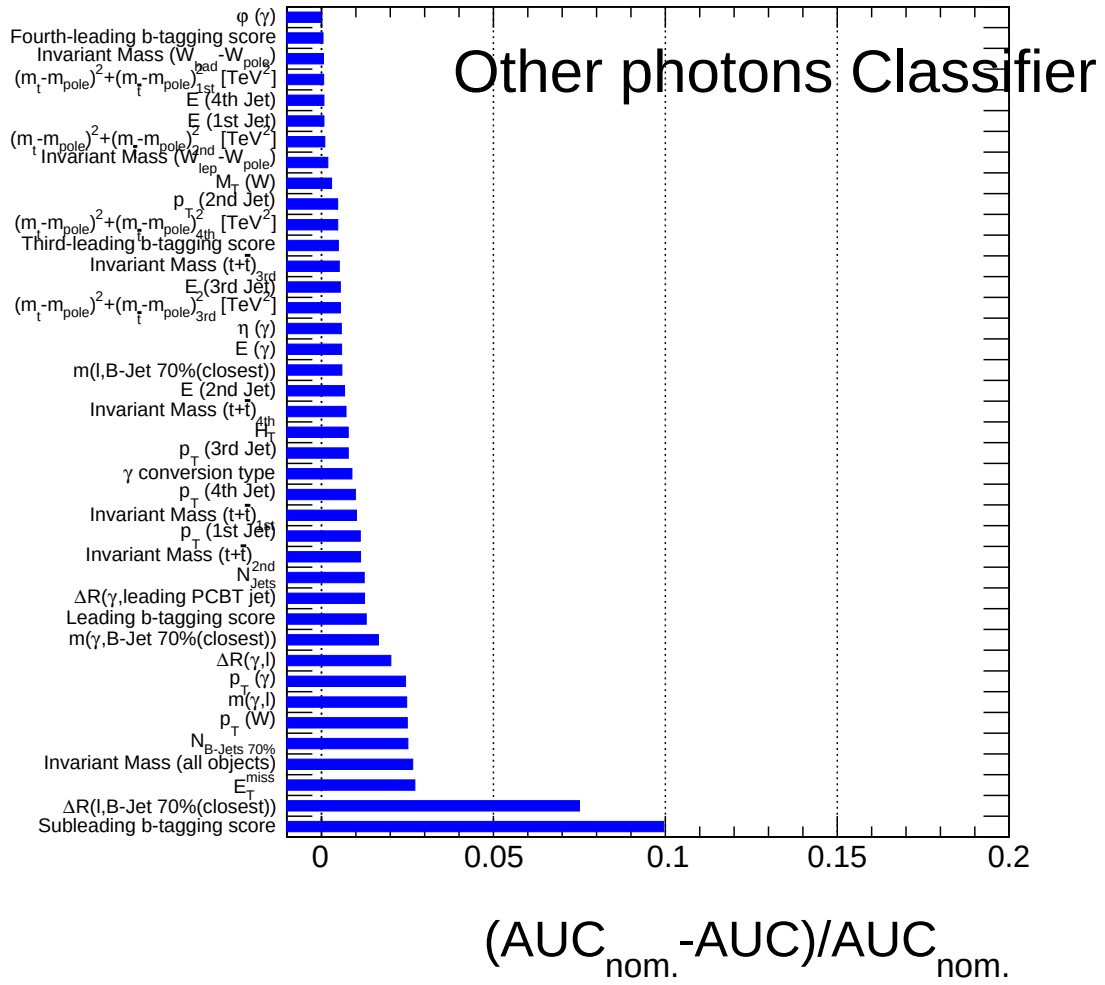


Figure D.48.: Permutation importance for the other photon classifier.

D.1.6. Output correlation plots

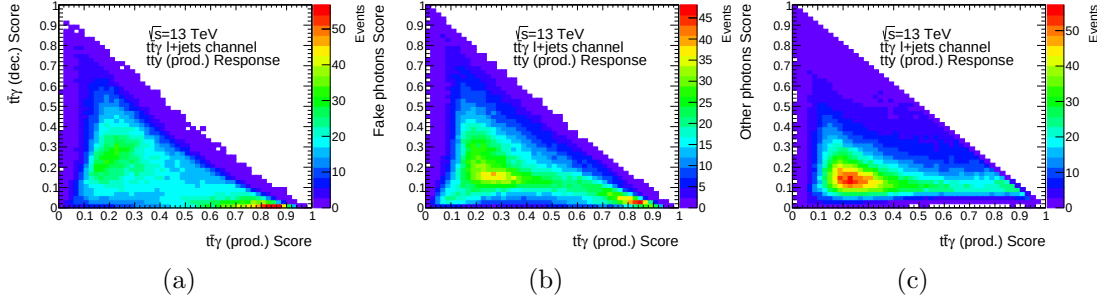


Figure D.49.: The distributions of the $t\bar{t}\gamma$ (prod.) classifier vs. the $t\bar{t}\gamma$ (dec.) classifier (a), of the $t\bar{t}\gamma$ (prod.) classifier vs. the photon fake classifier (b) and of the $t\bar{t}\gamma$ (prod.) classifier vs. the other photon classifier (c) for $t\bar{t}\gamma$ (prod.) events.

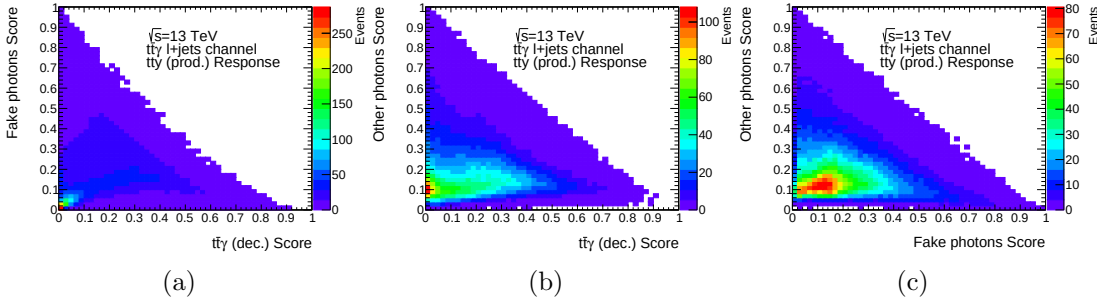


Figure D.50.: The distributions of the $t\bar{t}\gamma$ (dec.) classifier vs. the photon fake classifier (a), of the $t\bar{t}\gamma$ (dec.) classifier vs. the other photon classifier (b) and of the photon fake classifier vs. the other photon classifier (c) for $t\bar{t}\gamma$ (prod.) events.

D. Additional plots on the neural network training

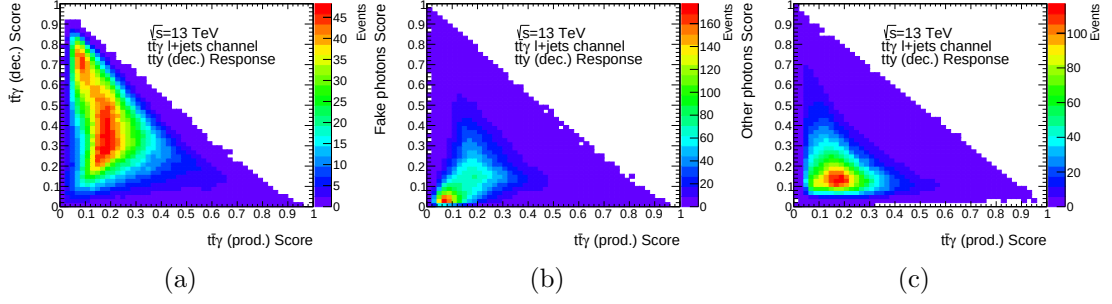


Figure D.51.: The distributions of the $t\bar{t}\gamma$ (prod.) classifier vs. the $t\bar{t}\gamma$ (dec.) classifier (a), of the $t\bar{t}\gamma$ (prod.) classifier vs. the photon fake classifier (b) and of the $t\bar{t}\gamma$ (prod.) classifier vs. the other photon classifier (c) for $t\bar{t}\gamma$ (dec.) events.

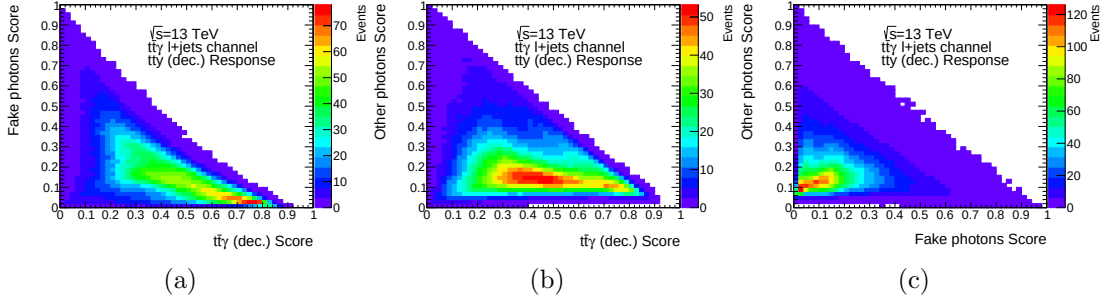


Figure D.52.: The distributions of the $t\bar{t}\gamma$ (dec.) classifier vs. the photon fake classifier (a), of the $t\bar{t}\gamma$ (dec.) classifier vs. the other photon classifier (b) and of the photon fake classifier vs. the other photon classifier (c) for $t\bar{t}\gamma$ (dec.) events.

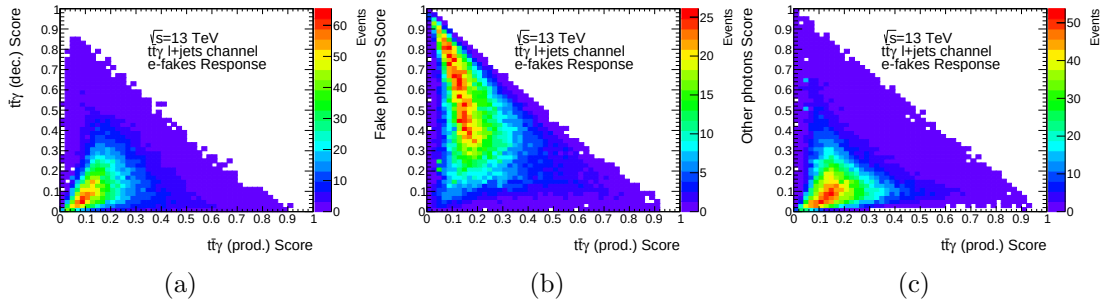


Figure D.53.: The distributions of the $t\bar{t}\gamma$ (prod.) classifier vs. the $t\bar{t}\gamma$ (dec.) classifier (a), of the $t\bar{t}\gamma$ (prod.) classifier vs. the photon fake classifier (b) and of the $t\bar{t}\gamma$ (prod.) classifier vs. the other photon classifier (c) for $e \rightarrow \gamma$ fake events.

D.1. $L+jets$ channel

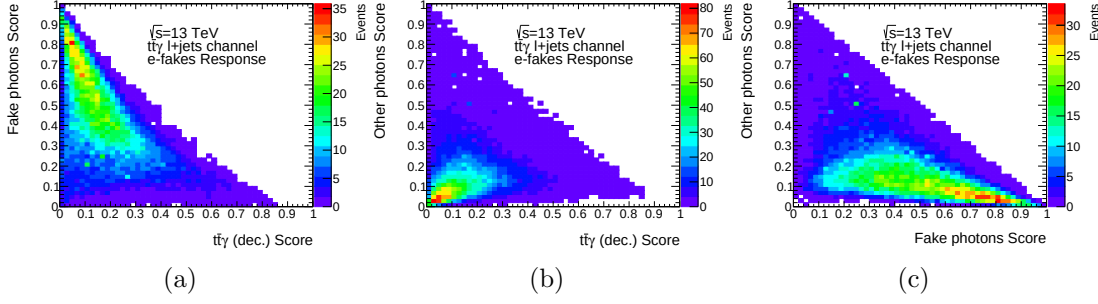


Figure D.54.: The distributions of the $t\bar{t}\gamma$ (dec.) classifier vs. the photon fake classifier (a), of the $t\bar{t}\gamma$ (dec.) classifier vs. the other photon classifier (b) and of the photon fake classifier vs. the other photon classifier (c) for $e \rightarrow \gamma$ fake events.

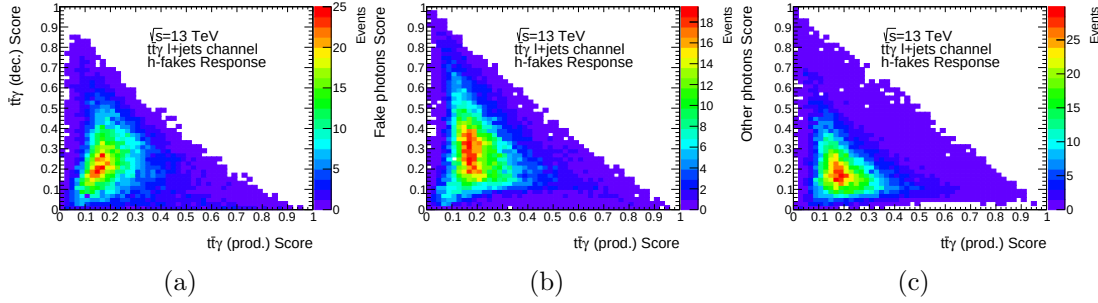


Figure D.55.: The distributions of the $t\bar{t}\gamma$ (prod.) classifier vs. the $t\bar{t}\gamma$ (dec.) classifier (a), of the $t\bar{t}\gamma$ (prod.) classifier vs. the photon fake classifier (b) and of the $t\bar{t}\gamma$ (prod.) classifier vs. the other photon classifier (c) for $h \rightarrow \gamma$ fake events.

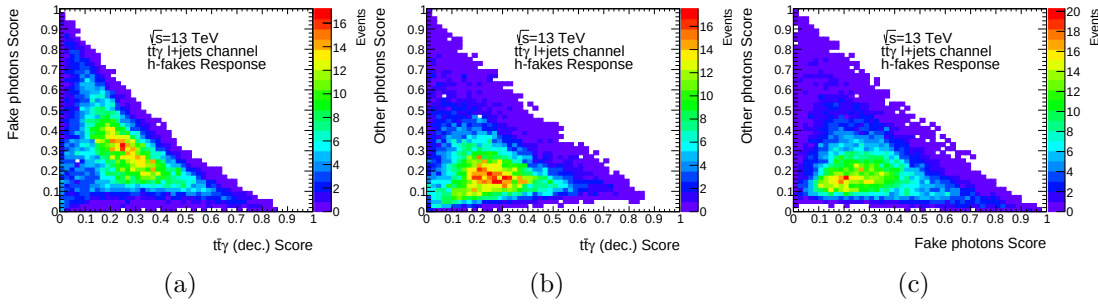


Figure D.56.: The distributions of the $t\bar{t}\gamma$ (dec.) classifier vs. the photon fake classifier (a), of the $t\bar{t}\gamma$ (dec.) classifier vs. the other photon classifier (b) and of the photon fake classifier vs. the other photon classifier (c) for $h \rightarrow \gamma$ fake events.

D. Additional plots on the neural network training

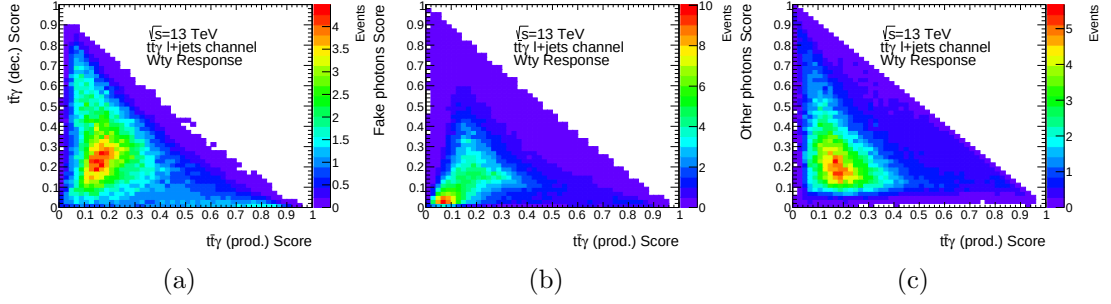


Figure D.57.: The distributions of the $t\bar{t}\gamma$ (prod.) classifier vs. the $t\bar{t}\gamma$ (dec.) classifier (a), of the $t\bar{t}\gamma$ (prod.) classifier vs. the photon fake classifier (b) and of the $t\bar{t}\gamma$ (prod.) classifier vs. the other photon classifier (c) for $Wt\gamma$ events.

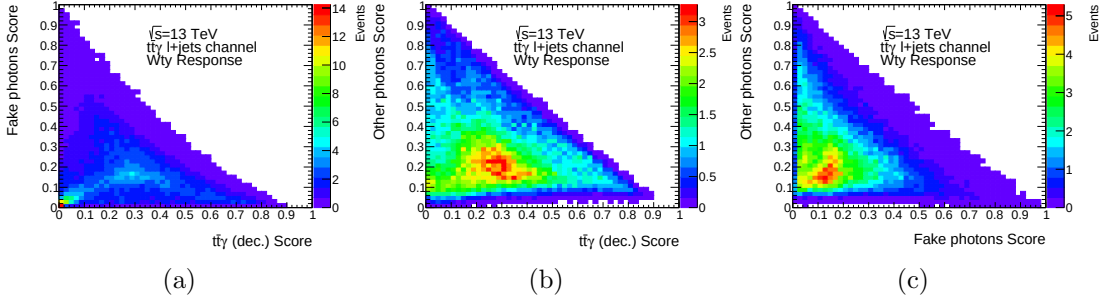


Figure D.58.: The distributions of the $t\bar{t}\gamma$ (dec.) classifier vs. the photon fake classifier (a), of the $t\bar{t}\gamma$ (dec.) classifier vs. the other photon classifier (b) and of the photon fake classifier vs. the other photon classifier (c) for $Wt\gamma$ events.

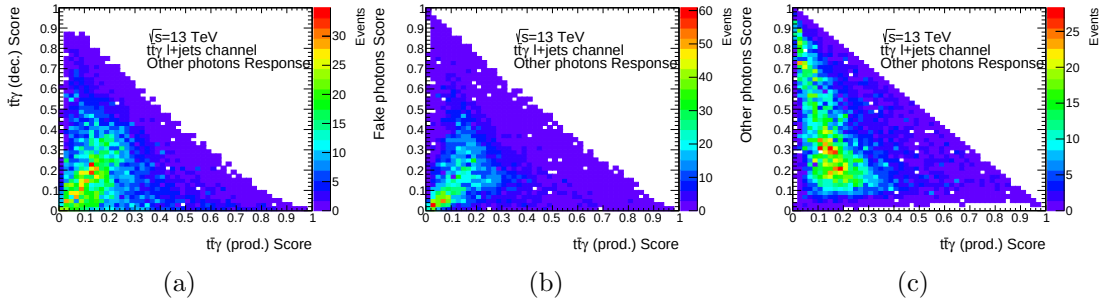


Figure D.59.: The distributions of the $t\bar{t}\gamma$ (prod.) classifier vs. the $t\bar{t}\gamma$ (dec.) classifier (a), of the $t\bar{t}\gamma$ (prod.) classifier vs. the photon fake classifier (b) and of the $t\bar{t}\gamma$ (prod.) classifier vs. the other photon classifier (c) for other photon events.

D.2. Dilepton channel

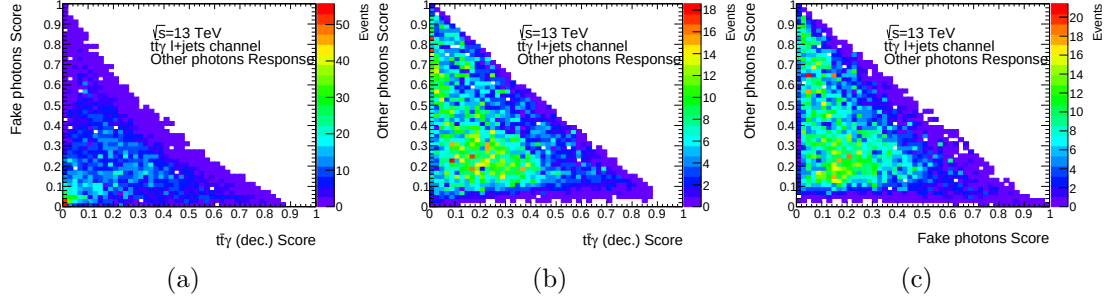


Figure D.60.: The distributions of the $t\bar{t}\gamma$ (dec.) classifier vs. the photon fake classifier (a), of the $t\bar{t}\gamma$ (dec.) classifier vs. the other photon classifier (b) and of the photon fake classifier vs. the other photon classifier (c) for other photon events.

D.2. Dilepton channel

D.2.1. Input distributions

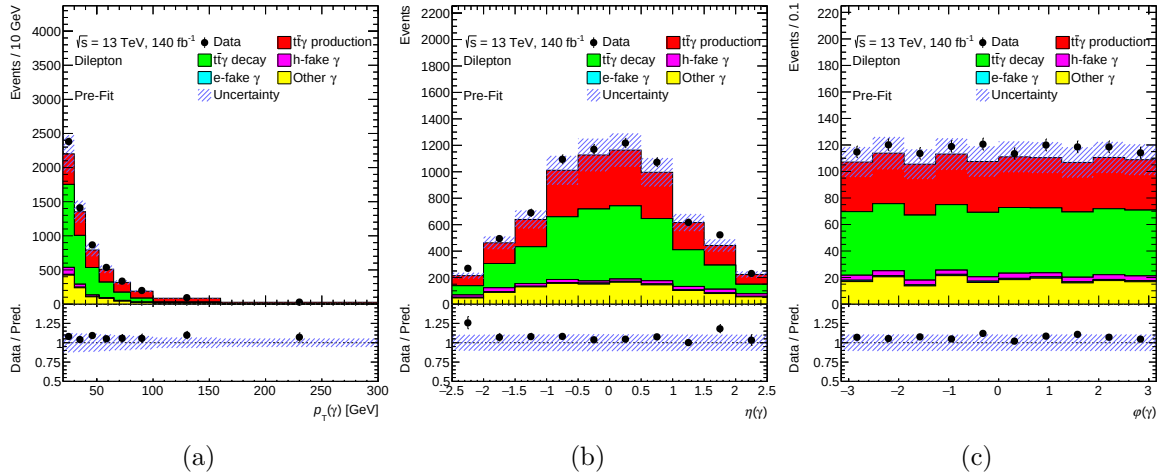


Figure D.61.: The p_T and η and ϕ distributions of the photon (from left to right) inserted into the NN training. The uncertainty band corresponds to the total uncertainty, including stat. and syst. uncertainties.

D. Additional plots on the neural network training

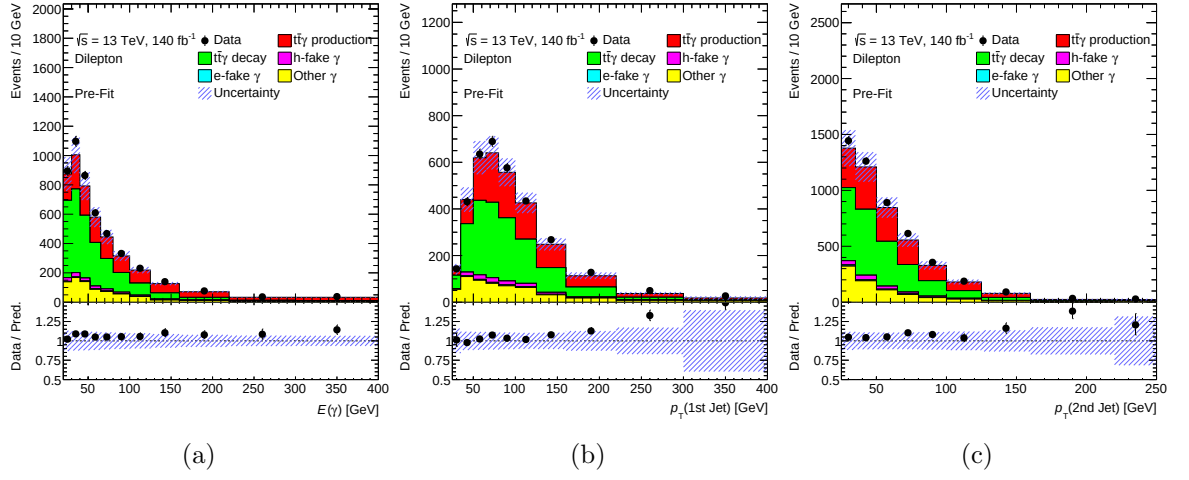


Figure D.62.: The E distribution of the photon and the p_T distributions of the first and second jet (from left to right) inserted into the NN training. The uncertainty band corresponds to the total uncertainty, including stat. and syst. uncertainties.

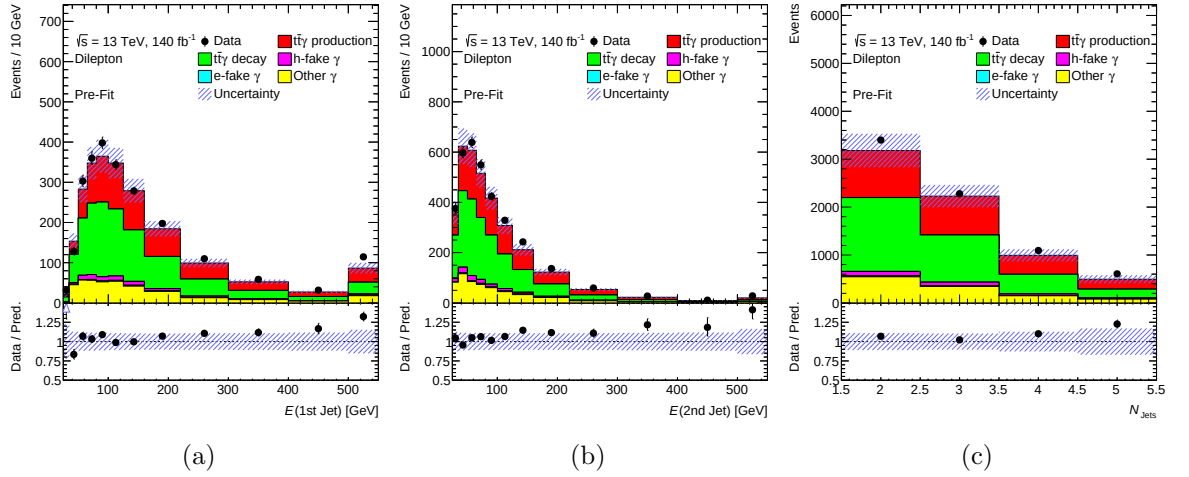


Figure D.63.: The E distributions of the first two jets and distribution of the number of jets per event (from left to right) inserted into the NN training. The uncertainty band corresponds to the total uncertainty, including stat. and syst. uncertainties.

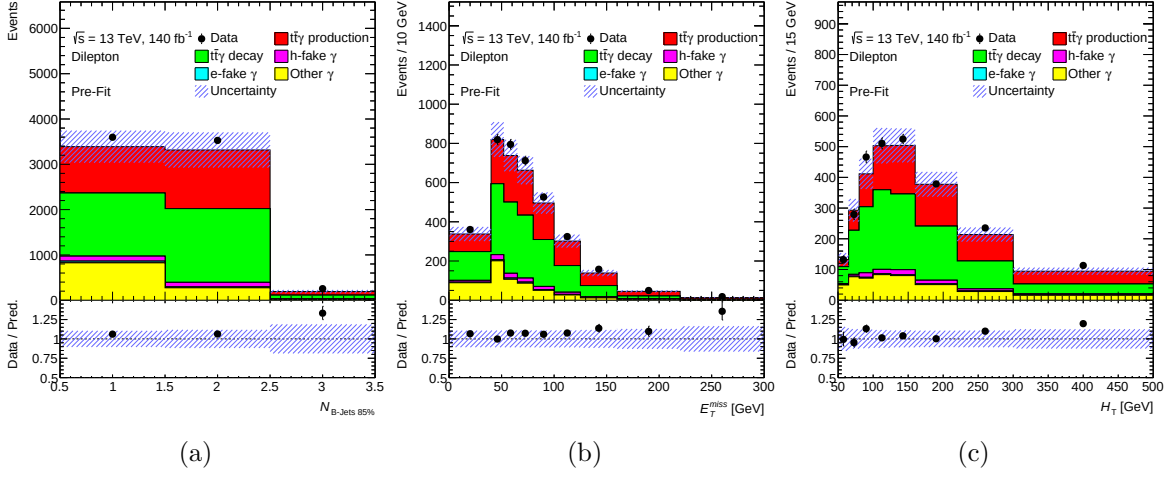


Figure D.64.: The distribution of the number of b -jets per event and the E_T^{miss} and H_T distributions (from left to right) inserted into the NN training. The uncertainty band corresponds to the total uncertainty, including stat. and syst. uncertainties.

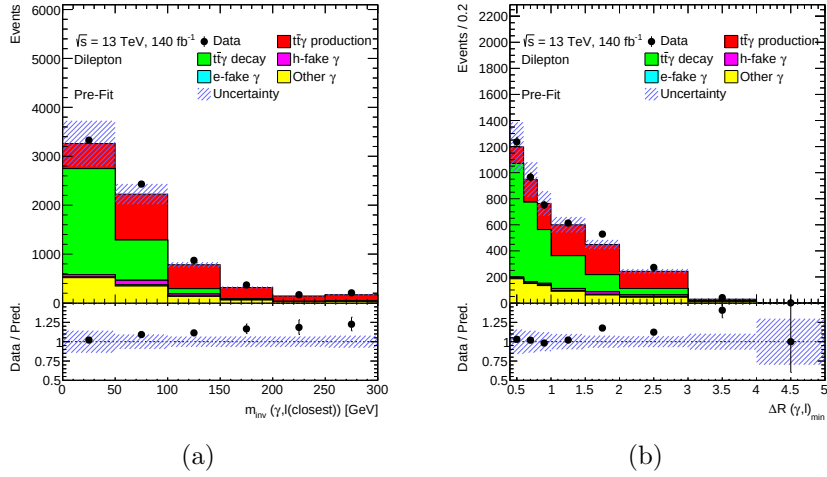


Figure D.65.: The invariant mass and ΔR distributions between the photon and the closest lepton inserted into the NN training. The uncertainty band corresponds to the total uncertainty, including stat. and syst. uncertainties.

D. Additional plots on the neural network training

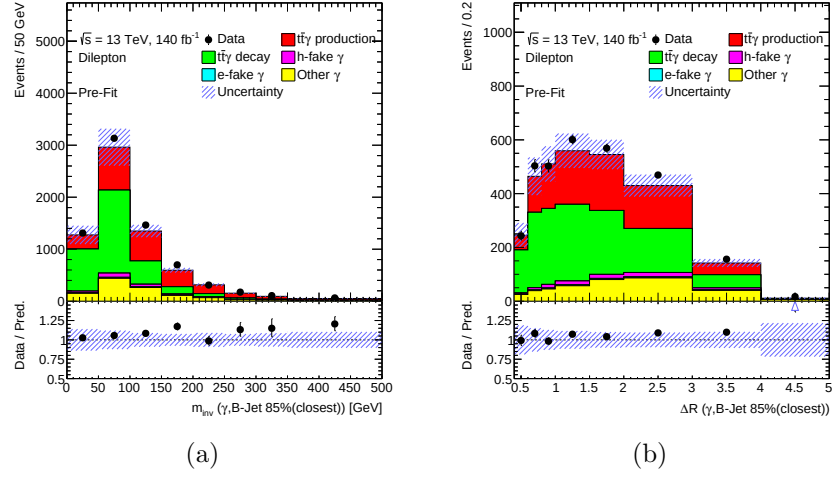


Figure D.66.: The invariant mass and ΔR distributions between the photon and the closest b -jet inserted into the NN training. The uncertainty band corresponds to the total uncertainty, including stat. and syst. uncertainties.

D.2.2. Separation plots

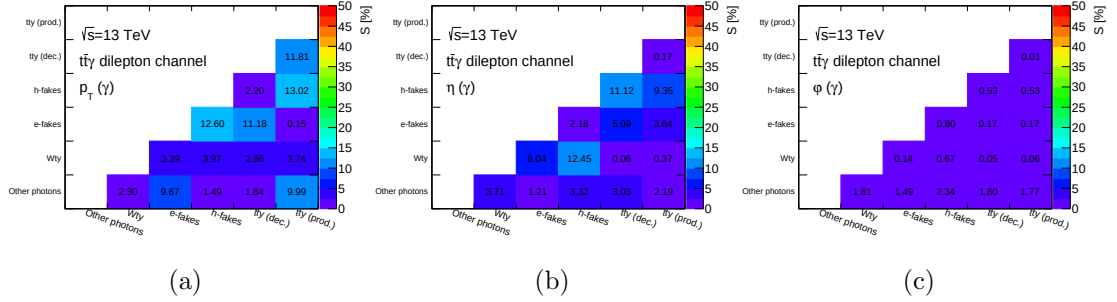


Figure D.67.: The separation values for the p_T and η and ϕ distributions of the photon (from left to right).

D.2. Dilepton channel

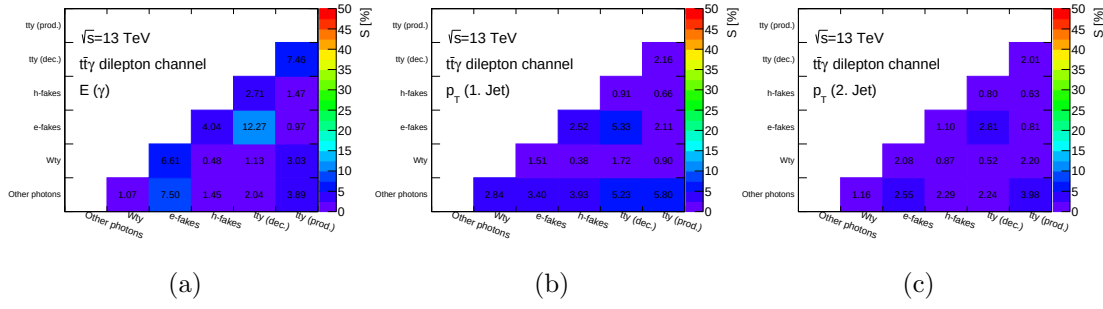


Figure D.68.: The separation values for the E distribution of the photon and for the p_T distributions of the first and second jet (from left to right).

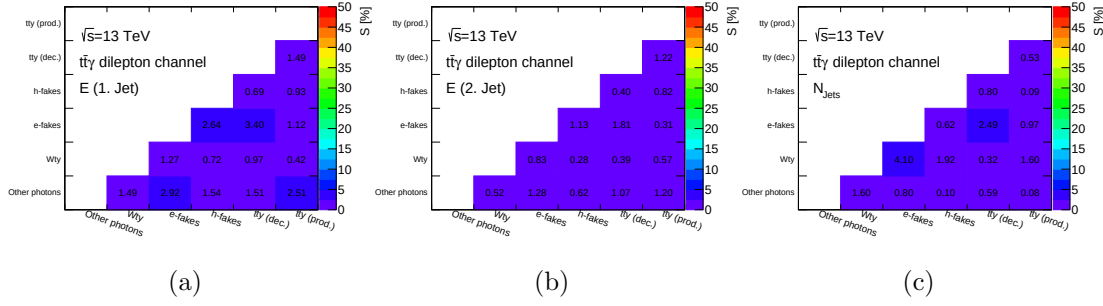


Figure D.69.: The separation values for the E distributions of the first two jets and for the distribution of the number of jets per event (from left to right).

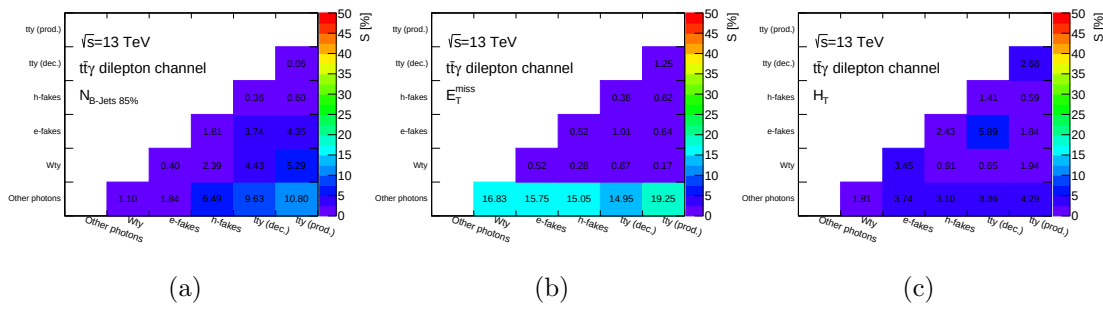


Figure D.70.: The separation values for the distribution of the number of b -jets per event and for the E_T^{miss} and H_T distributions (from left to right).

D. Additional plots on the neural network training

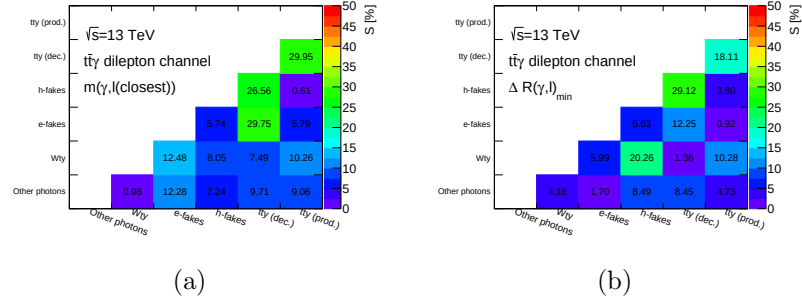


Figure D.71.: The separation values for the invariant mass and ΔR distributions between the photon and the closest lepton (from left to right).

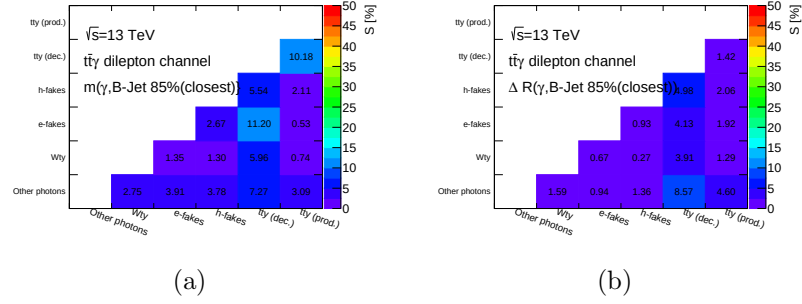


Figure D.72.: The separation values for the invariant mass and ΔR distributions between the photon and the closest b -jet (from left to right).

D.2.3. Negative weight comparisons

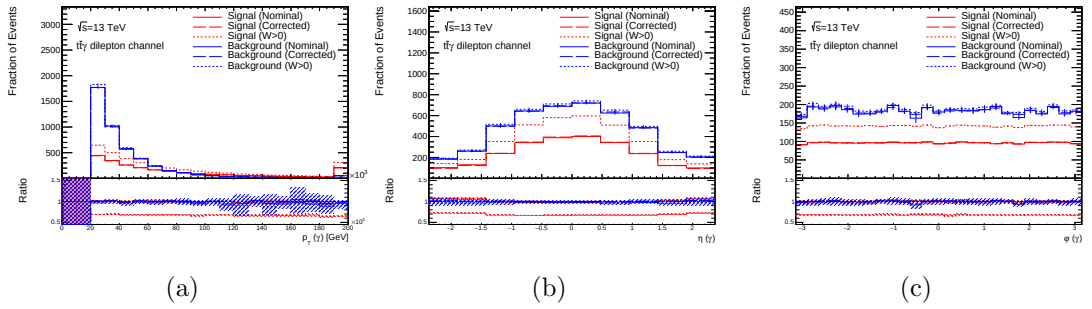


Figure D.73.: The impact of negative weights on the p_T and η and ϕ distributions of the photon (from left to right).

D.2. Dilepton channel

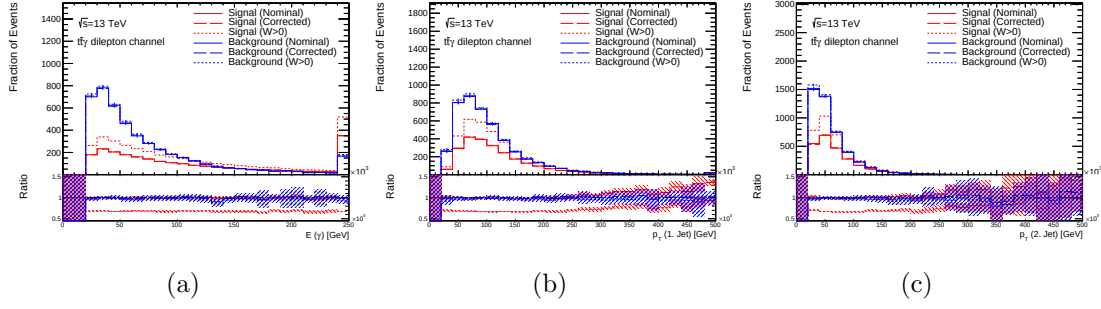


Figure D.74.: The impact of negative weights on the E distribution of the photon and on the p_T distributions of the first and second jet (from left to right).

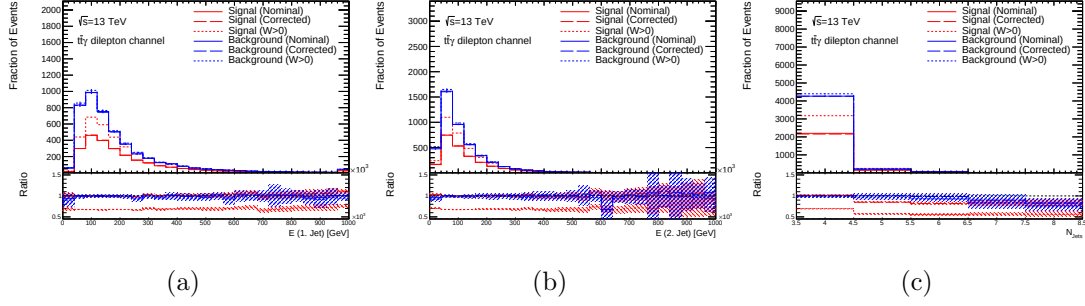


Figure D.75.: The impact of negative weights on the E distributions of the first two jets and on the distribution of the number of jets per event (from left to right).

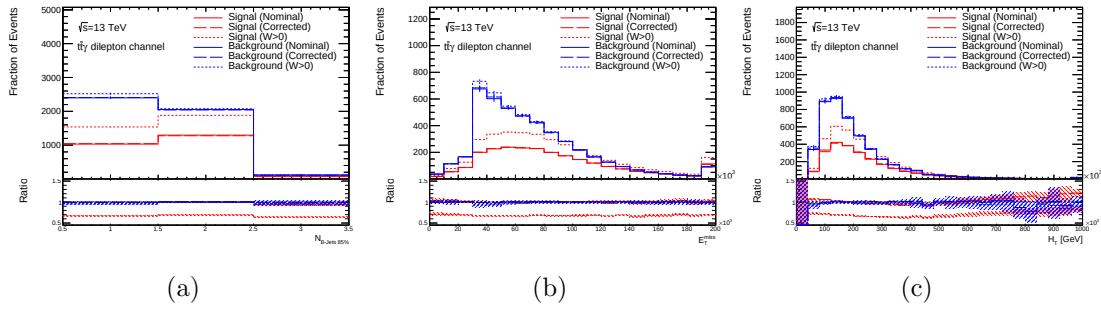


Figure D.76.: The impact of negative weights on the distribution of the number of b -jets per event and on the E_T^{miss} and H_T distributions (from left to right).

D. Additional plots on the neural network training

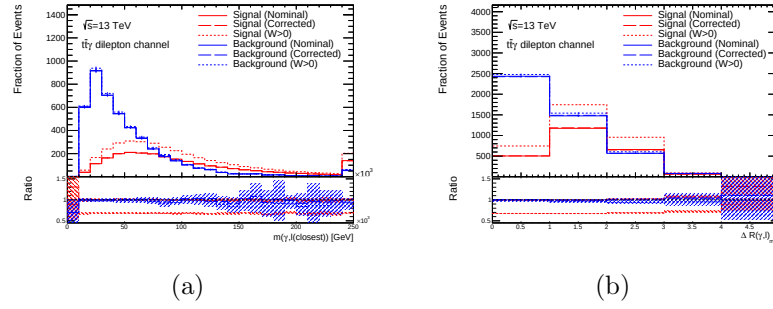


Figure D.77.: The impact of negative weights on the invariant mass and ΔR distributions between the photon and the closest lepton (from left to right).

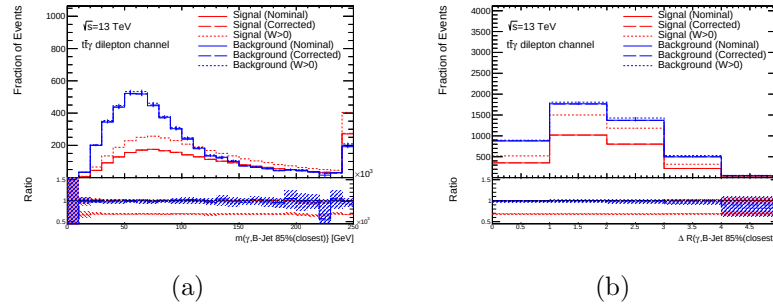


Figure D.78.: The impact of negative weights on the invariant mass and ΔR distributions between the photon and the closest b -jet (from left to right).

D.2.4. Comparison of training and testing output distributions

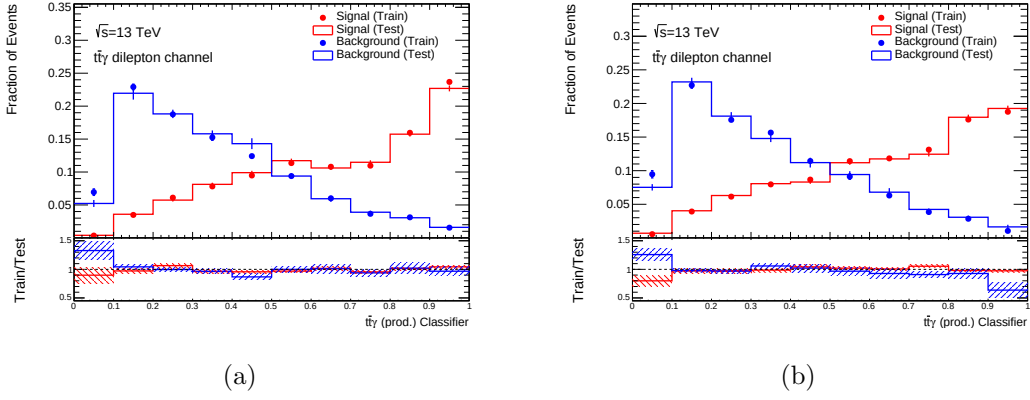


Figure D.79.: Output distributions for background and signal for fold 2 and 3.

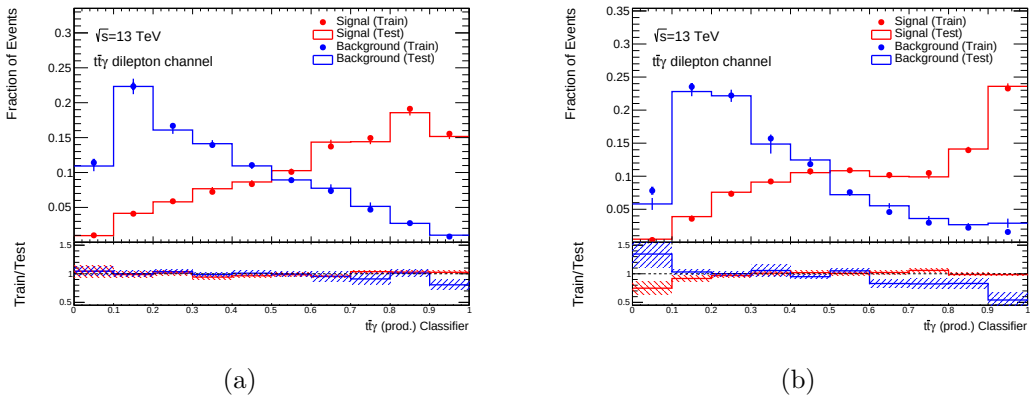


Figure D.80.: Output distributions for background and signal for fold 4 and 5.

Additional material on inclusive cross-section fits

E.1. Additional material on Asimov fits

Correlation matrices for the Asimov fits in the dilepton channel

Signal Strength	100.0	-25.9	-28.5	-13.5	-21.6	16.6	-44.7	-51.2
jet pileup rho topology	-25.9	100.0	-1.6	-0.8	-5.3	9.7	-1.1	4.9
W γ norm. uncert.	-28.5	-1.6	100.0	-2.6	-3.0	2.2	-0.2	-8.8
Z γ norm. uncert.	-13.5	-0.8	-2.6	100.0	-3.4	-12.8	0.4	-26.0
hfake γ SF	-21.6	-5.3	-3.0	-3.4	100.0	7.3	-4.1	21.5
tty (dec) H7	16.6	9.7	2.2	-12.8	7.3	100.0	-0.3	8.3
tty (prod) H7	-44.7	-1.1	-0.2	0.4	-4.1	-0.3	100.0	48.2
tty (dec.) norm. uncert. (k-factor)	-51.2	4.9	-8.8	-26.0	21.5	8.3	48.2	100.0
	Signal Strength	jet pileup rho topology	W γ norm. uncert.	Z γ norm. uncert.	hfake γ SF	tty (dec) H7	tty (prod) H7	tty (dec.) norm. uncert. (k-factor)

(a)

Signal Strength	100.0	-25.7	7.0	-8.8	-16.0	-27.2	-11.2	-22.7	15.6	-46.7	41.1
jet pileup rho topology	-25.7	100.0	-1.3	-1.1	-1.2	-1.6	-1.0	-5.2	9.8	-0.8	-19.9
Pileup	7.0	-1.3	100.0	-1.0	-0.5	-1.0	-2.6	-1.6	-2.0	-0.0	22.4
Photon Identification	-8.8	-1.1	-1.0	100.0	-0.6	-1.0	-2.2	-2.1	-0.3	-0.7	-27.9
Photon Isolation	-16.0	-1.2	-0.5	-0.6	100.0	-0.8	-1.4	-2.0	1.1	-0.8	-27.0
W γ norm. uncert.	-27.2	-1.6	-1.0	-1.0	-0.8	100.0	-2.3	-3.3	2.2	-0.9	-37.3
Z γ norm. uncert.	-11.2	-1.0	-2.6	-2.2	-1.4	-2.3	100.0	-4.2	-13.1	-1.5	-41.2
hfake γ SF	-22.7	-5.2	-1.6	-2.1	-2.0	-3.3	-4.2	100.0	7.6	-2.6	2.3
tty (dec) H7	15.6	9.8	-2.0	-0.3	1.1	2.2	-13.1	7.6	100.0	0.3	25.1
tty (prod) H7	-46.7	-0.8	-0.0	-0.7	-0.8	-0.9	-1.5	-2.6	0.3	100.0	8.6
Norm factor for tty-dec	41.1	-19.9	22.4	-27.9	-27.0	-37.3	-41.2	2.3	25.1	8.6	100.0
	Signal Strength	jet pileup rho topology	Pileup	Photon Identification	Photon Isolation	W γ norm. uncert.	Z γ norm. uncert.	hfake γ SF	tty (dec) H7	tty (prod) H7	Norm factor for tty-dec

(b)

Figure E.1.: The correlation matrix of the Asimov fits with one NF on the left side and with two NF on the right side for the dilepton channel. Only nuisance parameters with at least one correlation of at least $\pm 20\%$ are shown.

E. Additional material on inclusive cross-section fits

Correlation matrices for the Asimov fits in the $l+\text{jets}$ channel

Signal Strength	100.0	15.6	13.1	12.5	-26.3	-20.0	4.7	7.0	14.4	7.0	-2.8	12.6	29.9	-60.3	-53.5
b-tagging eigenvariation B-0	15.6	100.0	-34.4	7.6	-3.1	-3.5	-23.1	-6.4	7.3	0.4	3.9	-4.9	22.0	-2.3	-1.7
b-tagging eigenvariation B-1	13.1	-34.4	100.0	4.2	-3.0	-3.1	-10.2	-1.5	4.5	-1.3	2.0	-0.4	12.5	-0.4	-8.0
jet flavor response	12.5	7.6	4.2	100.0	12.5	2.0	21.2	3.9	8.8	6.5	14.0	-0.8	0.1	-1.4	0.7
jet pileup rho topology	-26.3	-3.1	-3.0	12.5	100.0	-2.6	-9.0	-3.1	-19.5	-4.7	-30.3	6.0	-1.8	3.1	2.7
$W\gamma$ norm. uncert.	-20.0	-3.5	-3.1	2.0	-2.6	100.0	-25.8	-1.2	2.1	2.0	-3.7	0.5	-3.1	-1.5	2.9
$W\gamma$ norm. uncert.	4.7	-23.1	-10.2	21.2	-9.0	-25.8	100.0	-40.4	5.2	-21.0	6.8	-17.5	-2.3	-2.0	-7.9
$Z\gamma$ norm. uncert.	7.0	-6.4	-1.5	3.9	-3.1	-1.2	-40.4	100.0	-1.9	-0.1	-3.6	1.0	5.5	1.8	-6.2
efake γ SF	14.4	7.3	4.5	8.8	-19.5	2.1	5.2	-1.9	100.0	-7.8	-28.6	2.8	-14.0	3.6	4.4
extrapolation uncert. light-jets	7.0	0.4	-1.3	6.5	-4.7	2.0	-21.0	-0.1	-7.8	100.0	9.3	22.4	-11.1	-6.6	17.6
hfake γ SF	-2.8	3.9	2.0	14.0	-30.3	-3.7	6.8	-3.6	-28.6	9.3	100.0	-7.8	-20.1	5.8	15.7
tty (dec) H7	12.6	-4.9	-0.4	-0.8	6.0	0.5	-17.5	1.0	2.8	22.4	-7.8	100.0	3.2	0.6	51.4
tty (prod) H7	29.9	22.0	12.5	0.1	-1.8	-3.1	-2.3	5.5	-14.0	-11.1	-20.1	3.2	100.0	-1.0	-10.0
tty (prod) var3c	-60.3	-2.3	-0.4	-1.4	3.1	-1.5	-2.0	1.8	3.6	-6.6	5.8	0.6	-1.0	100.0	54.2
tty (dec.) norm. uncert. (k-factor)	-53.5	-1.7	-8.0	0.7	2.7	2.9	-7.9	-6.2	4.4	17.6	15.7	51.4	-10.0	54.2	100.0

Figure E.2.: The correlation matrix for the Asimov fit with one NF for the $l+\text{jets}$ channel. Only nuisance parameters with at least one correlation of at least $\pm 20\%$ are shown.

E.1. Additional material on Asimov fits

Signal Strength	100.0	15.4	13.5	12.2	-25.9	-1.4	-5.6	-19.7	5.3	7.4	13.7	5.3	-4.1	-7.5	7.5	30.0	-62.2	31.7
b-tagging eigenvariation B-0	15.4	100.0	-34.4	7.6	-3.1	1.2	0.5	-3.5	-23.1	-6.3	7.3	0.3	3.8	-3.6	-4.9	22.0	-2.4	12.8
b-tagging eigenvariation B-1	13.5	-34.4	100.0	4.1	-3.1	0.6	0.5	-3.1	-10.1	-1.4	4.4	-1.5	1.8	-0.5	-1.1	12.6	-1.1	3.1
jet flavor response	12.2	7.6	4.1	100.0	12.5	2.9	3.1	2.0	21.2	4.0	8.8	6.5	14.0	-1.8	-0.7	0.1	-1.3	12.6
jet pileup rho topology	-25.9	-3.1	-3.1	12.5	100.0	-5.7	-5.9	-2.6	-9.0	-3.2	-19.5	-4.7	-30.1	6.3	6.0	-1.9	3.2	-21.7
Photon Identification	-1.4	1.2	0.6	2.9	-5.7	100.0	-2.0	0.0	-4.2	-1.3	-5.6	0.0	-9.1	-0.4	-2.4	-1.6	1.8	-20.2
Photon Isolation	-5.6	0.5	0.5	3.1	-5.9	-2.0	100.0	1.3	-1.3	-0.8	-6.9	-6.5	-8.0	0.7	-1.2	-2.3	1.7	-20.5
W γ norm. uncert.	-19.7	-3.5	-3.1	2.0	-2.6	0.0	1.3	100.0	-25.9	-1.2	2.1	2.0	-3.6	-0.7	0.7	-3.2	-1.2	-15.5
W γ norm. uncert.	5.3	-23.1	-10.1	21.2	-9.0	-4.2	-1.3	-25.9	100.0	-40.2	5.1	-21.2	6.6	-14.2	-17.7	-2.1	-2.7	-4.7
Z γ norm. uncert.	7.4	-6.3	-1.4	4.0	-3.2	-1.3	-0.8	-1.2	-40.2	100.0	-1.9	-0.3	-3.7	-2.4	0.4	5.6	1.2	-0.5
efake γ SF	13.7	7.3	4.4	8.8	-19.5	-5.6	-6.9	2.1	5.1	-1.9	100.0	-7.6	-28.4	3.5	3.1	-14.0	4.0	18.7
extrapolation uncert. light-jets	5.3	0.3	-1.5	6.5	-4.7	0.0	-6.5	2.0	-21.2	-0.3	-7.6	100.0	9.8	17.0	23.4	-11.3	-4.8	27.0
hfake γ SF	-4.1	3.8	1.8	14.0	-30.1	-9.1	-8.0	-3.6	6.6	-3.7	-28.4	9.8	100.0	-1.8	-6.2	-20.3	7.1	15.5
t $\bar{t}$ bar pt hard	-7.5	-3.6	-0.5	-1.8	6.3	-0.4	0.7	-0.7	-14.2	-2.4	3.5	17.0	-1.8	100.0	-17.9	0.4	2.4	-21.3
t $\bar{t}$ (dec) H7	7.5	-4.9	-1.1	-0.7	6.0	-2.4	-1.2	0.7	-17.7	0.4	3.1	23.4	-6.2	-17.9	100.0	2.3	5.2	69.9
t $\bar{t}$ (prod) H7	30.0	22.0	12.6	0.1	-1.9	-1.6	-2.3	-3.2	-2.1	5.6	-14.0	-11.3	-20.3	0.4	2.3	100.0	-1.8	16.7
t $\bar{t}$ (prod) var3c	-62.2	-2.4	-1.1	-1.3	3.2	1.8	1.7	-1.2	-2.7	1.2	4.0	-4.8	7.1	2.4	5.2	-1.8	100.0	5.7
Norm factor for t $\bar{t}$ -dec	31.7	12.8	3.1	12.6	-21.7	-20.2	-20.5	-15.5	-4.7	-0.5	18.7	27.0	15.5	-21.3	69.9	16.7	5.7	100.0

Figure E.3.: The correlation matrix for the Asimov fit with two NFs for the l+jets channel. Only nuisance parameters with at least one correlation of at least $\pm 20\%$ are shown.

E. Additional material on inclusive cross-section fits

Correlation matrices for the Asimov fits in the combination

Signal Strength	100.0	7.3	7.7	5.0	-20.1	-6.0	-38.3	18.5	-12.1	30.4	8.1	0.7	-11.6	17.4	-14.6	-23.0	-48.2	-42.1
b-tagging eigenvariation B-0	7.3	100.0	-37.3	5.3	0.1	0.7	-7.1	-22.5	-11.1	10.9	0.0	4.0	-4.6	-2.1	16.1	1.1	5.4	9.9
b-tagging eigenvariation B-1	7.7	-37.3	100.0	2.7	-0.9	-0.0	-5.0	-9.1	-4.9	6.3	-1.7	2.2	-0.7	-0.6	7.5	1.5	4.7	-4.0
jet flavor response	5.0	5.3	2.7	100.0	16.8	2.1	-2.2	28.1	-2.1	10.1	4.2	14.8	-0.9	-5.8	-8.5	0.4	4.0	3.8
jet pileup rho topology	-20.1	0.1	-0.9	16.8	100.0	-4.6	2.9	-16.5	5.6	-22.0	-1.5	-31.6	5.3	14.7	9.7	-1.2	-3.8	-0.1
Photon Identification	-6.0	0.7	-0.0	2.1	-4.6	100.0	-2.4	-1.4	-6.0	-5.2	-1.1	-9.1	-0.1	-5.3	-3.8	0.2	4.2	-23.2
W _l norm. uncert.	-38.3	-7.1	-5.0	-2.2	2.9	-2.4	100.0	-17.2	-15.7	4.4	-4.0	-5.3	1.0	-9.2	-10.7	3.0	2.0	-1.4
W _γ norm. uncert.	18.5	-22.5	-9.1	28.1	-16.5	-1.4	-17.2	100.0	-28.2	2.9	-15.6	8.9	-18.8	-5.4	5.0	-1.3	-5.9	-0.5
Z _γ norm. uncert.	-12.1	-11.1	-4.9	-2.1	5.6	-6.0	-15.7	-28.2	100.0	1.2	-9.9	-4.9	2.0	-27.1	-8.2	0.7	10.1	-20.6
efake γ SF	30.4	10.9	6.3	10.1	-22.0	-5.2	4.4	2.9	1.2	100.0	-9.7	-28.6	5.6	-4.6	-8.4	0.7	-2.4	-7.7
extrapolation uncert. light-jets	8.1	0.0	-1.7	4.2	-1.5	-1.1	-4.0	-15.6	-9.9	-9.7	100.0	9.2	20.8	14.7	-14.1	13.6	-8.9	11.2
hfake γ SF	0.7	4.0	2.2	14.8	-31.6	-9.1	-5.3	8.9	-4.9	-28.6	9.2	100.0	-2.0	-8.2	-22.0	-1.0	4.3	18.9
ttbar pt hard	-11.6	-4.6	-0.7	-0.9	5.3	-0.1	1.0	-18.8	2.0	5.6	20.8	-2.0	100.0	-10.9	-0.4	-3.2	5.4	-6.2
tty (dec) H7	17.4	-2.1	-0.6	-5.8	14.7	-5.3	-9.2	-5.4	-27.1	-4.6	14.7	-8.2	-10.9	100.0	6.4	-5.0	-5.8	40.3
tty (prod) H7	-14.6	16.1	7.5	-8.5	9.7	-3.8	-10.7	5.0	-8.2	-8.4	-14.1	-22.0	-0.4	6.4	100.0	11.6	31.4	29.9
tty (prod) mu R	-23.0	1.1	1.5	0.4	-1.2	0.2	3.0	-1.3	0.7	0.7	13.6	-1.0	-3.2	-5.0	11.6	100.0	0.1	14.0
tty (prod) var3c	-48.2	5.4	4.7	4.0	-3.8	4.2	2.0	-5.9	10.1	-2.4	-8.9	4.3	5.4	-5.8	31.4	0.1	100.0	45.4
tty (dec.) norm. uncert. (k-factor)	-42.1	9.9	-4.0	3.8	-0.1	-23.2	-1.4	-0.5	-20.6	-7.7	11.2	18.9	-6.2	40.3	29.9	14.0	45.4	100.0

Figure E.4.: The correlation matrix of the Asimov fit with one NF for the combination of dilepton and l+jets channels. Only nuisance parameters with at least one correlation of at least $\pm 20\%$ are shown.

E.1. Additional material on Asimov fits

Signal Strength	100.0	6.9	7.8	4.9	-20.0	7.7	-5.1	-12.9	-37.9	18.3	-11.2	30.5	7.7	0.1	-11.3	15.7	-15.5	-23.3	-49.0	42.0
b-tagging eigenvariation B-0	6.9	100.0	-37.3	5.4	0.0	-3.9	0.5	-0.6	-7.1	-22.4	-11.3	10.8	0.1	4.2	-4.7	-1.7	16.3	1.2	5.7	16.4
b-tagging eigenvariation B-1	7.8	-37.3	100.0	2.7	-0.9	-2.7	0.0	-0.3	-5.0	-9.1	-4.8	6.3	-1.8	2.1	-0.7	-0.7	7.4	1.4	4.5	2.5
jet flavor response	4.9	5.4	2.7	100.0	16.8	2.0	2.0	2.0	-2.2	28.1	-2.1	10.1	4.2	14.8	-1.0	-5.6	-8.4	0.5	4.1	8.2
jet pileup rho topology	-20.0	0.0	-0.9	16.8	100.0	-1.5	-4.5	-4.6	2.9	-16.6	5.6	-22.0	-1.5	-31.5	5.3	14.5	9.6	-1.2	-3.8	-17.3
Pileup	7.7	-3.9	-2.7	2.0	-1.5	100.0	-0.8	-0.6	1.6	-1.7	1.1	0.5	-0.3	-0.4	-1.5	-2.3	1.9	0.4	1.0	23.4
Photon Identification	-5.1	0.5	0.0	2.0	-4.5	-0.8	100.0	-2.5	-2.3	-1.4	-5.5	-5.0	-1.3	-9.5	0.0	-6.1	-4.4	-0.1	3.2	-28.8
Photon Isolation	-12.9	-0.6	-0.3	2.0	-4.6	-0.6	-2.5	100.0	-0.4	0.7	-3.8	-6.2	-7.5	-8.1	0.9	-2.8	-6.7	0.7	4.4	-27.7
Wty norm. uncert.	-37.9	-7.1	-5.0	-2.2	2.9	1.6	-2.3	-0.4	100.0	-17.2	-15.7	4.4	-4.1	-5.3	1.0	-9.2	-10.7	3.0	1.9	-34.1
Wy norm. uncert.	18.3	-22.4	-9.1	28.1	-16.6	-1.7	-1.4	0.7	-17.2	100.0	-28.1	3.0	-15.6	8.9	-18.8	-5.4	4.9	-1.3	-5.8	15.3
Zy norm. uncert.	-11.2	-11.3	-4.8	-2.1	5.6	1.1	-5.5	-3.8	-15.7	-28.1	100.0	1.3	-10.1	-5.3	2.1	-27.6	-8.7	0.5	9.2	-31.3
efake γ SF	30.5	10.8	6.3	10.1	-22.0	0.5	-5.0	-6.2	4.4	3.0	1.3	100.0	-9.7	-28.6	5.6	-4.9	-8.5	0.6	-2.7	18.1
extrapolation uncert. light-jets	7.7	0.1	-1.8	4.2	-1.5	-0.3	-1.3	-7.5	-4.1	-15.6	-10.1	-9.7	100.0	9.4	20.7	15.0	-13.7	13.7	-8.3	18.4
hfake γ SF	0.1	4.2	2.1	14.8	-31.5	-0.4	-9.5	-8.1	-5.3	8.9	-5.3	-28.6	9.4	100.0	-2.1	-7.4	-21.4	-0.8	5.0	20.0
ttbar pt hard	-11.3	-4.7	-0.7	-1.0	5.3	-1.5	0.0	0.9	1.0	-18.8	2.1	5.6	20.7	-2.1	100.0	-11.1	-0.5	-3.3	5.1	-16.2
tty (dec) H7	15.7	-1.7	-0.7	-5.6	14.5	-2.3	-6.1	-2.8	-9.2	-5.4	-27.6	-4.9	15.0	-7.4	-11.1	100.0	7.4	-4.5	-4.1	55.5
tty (prod) H7	-15.5	16.3	7.4	-8.4	9.6	1.9	-4.4	-6.7	-10.7	4.9	-8.7	-8.5	-13.7	-21.4	-0.5	7.4	100.0	11.9	32.2	18.0
tty (prod) mu R	-23.3	1.2	1.4	0.5	-1.2	0.4	-0.1	0.7	3.0	-1.3	0.5	0.6	13.7	-0.8	-3.3	-4.5	11.9	100.0	0.6	-5.3
tty (prod) var3c	-49.0	5.7	4.5	4.1	-3.8	1.0	3.2	4.4	1.9	-5.8	9.2	-2.7	-8.3	5.0	5.1	-4.1	32.2	0.6	100.0	5.2
Norm factor for tty-dec	42.0	16.4	2.5	8.2	-17.3	23.4	-28.8	-27.7	-34.1	15.3	-31.3	18.1	18.4	20.0	-16.2	55.5	18.0	-5.3	5.2	100.0
Signal Strength																				
b-tagging eigenvariation B-0																				
b-tagging eigenvariation B-1																				
jet flavor response																				
jet pileup rho topology																				
Pileup																				
Photon Identification																				
Photon Isolation																				
Wty norm. uncert.																				
Wy norm. uncert.																				
Zy norm. uncert.																				
efake γ SF																				
extrapolation uncert. light-jets																				
hfake γ SF																				
ttbar pt hard																				
tty (dec) H7																				
tty (prod) H7																				
tty (prod) mu R																				
tty (prod) var3c																				
Norm factor for tty-dec																				

Figure E.5.: The correlation matrix of the Asimov fit with two NFs for the combination of dilepton and l+jets channels. Only nuisance parameters with at least one correlation of at least $\pm 20\%$ are shown.

E.2. Additional material on unblinded fits

Pull-plots for the unblinded fits with 1 and 2 NF in the $l+jets$ channel

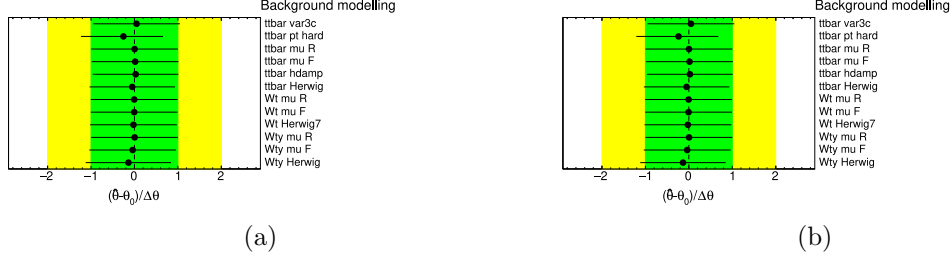


Figure E.6.: The left plot shows the pulls and constraints of those nuisance parameters grouped together as background modelling uncertainties in the fit with one NF, while the right plot shows the same nuisance parameters for the fit with 2 NF for the $l+jets$ channel.

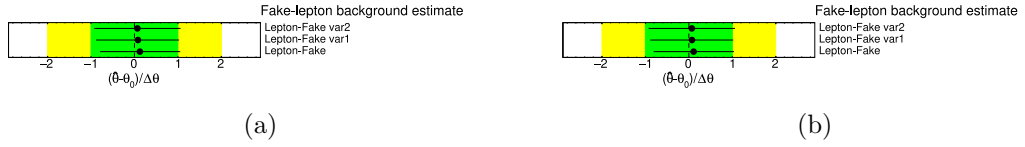


Figure E.7.: The left plot shows the pulls and constraints of those nuisance parameters grouped together as fake lepton background estimate uncertainties in the fit with one NF, while the right plot shows the same nuisance parameters for the fit with 2 NF for the $l+jets$ channel.

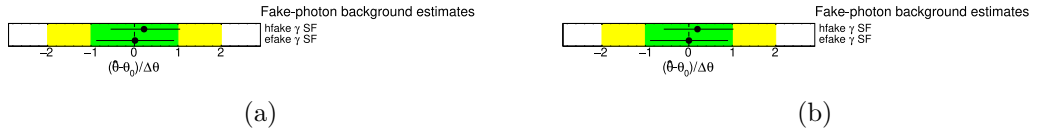


Figure E.8.: The left plot shows the pulls and constraints of those nuisance parameters grouped together as fake photon background estimate uncertainties in the fit with one NF, while the right plot shows the same nuisance parameters for the fit with 2 NF for the $l+jets$ channel.

E.2. Additional material on unblinded fits



Figure E.9.: The left plot shows the pull and constraint of the luminosity uncertainty in the fit with one NF, while the right plot shows the same nuisance parameter for the fit with 2 NF for the l+jets channel.



Figure E.10.: The left plot shows the pulls and constraints of those nuisance parameters grouped together as E_T^{miss} related uncertainties in the fit with one NF, while the right plot shows the same nuisance parameters for the fit with 2 NF for the l+jets channel.



Figure E.11.: The left plot shows the pull and constraint of the parton shower uncertainty for the $t\bar{t}\gamma$ from production sample in the fit with one NF, while the right plot shows the same nuisance parameters for the fit with 2 NF for the l+jets channel.

E. Additional material on inclusive cross-section fits

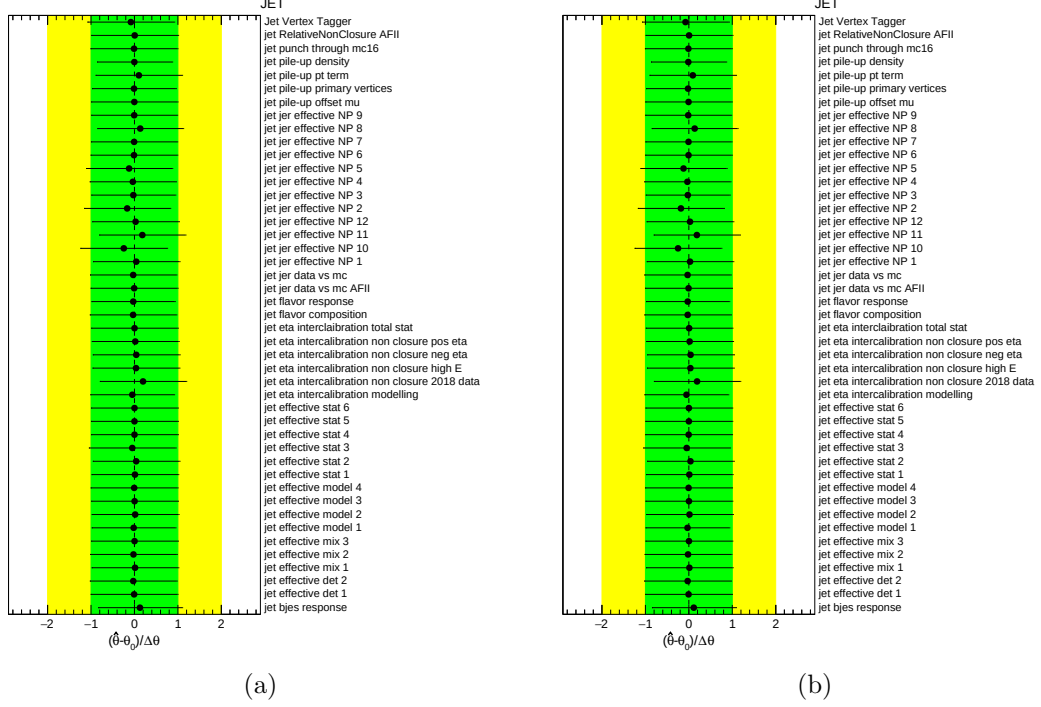


Figure E.12.: The left plot shows the pulls and constraints of those nuisance parameters grouped together as jet related uncertainties in the fit with one NF, while the right plot shows the same nuisance parameters for the fit with 2 NF for the l+jets channel.



Figure E.13.: Pull and constraint of the normalisation uncertainty for the $t\bar{t}\gamma$ from decay sample in the fit with one NF for the l+jets channel.

E.2. Additional material on unblinded fits

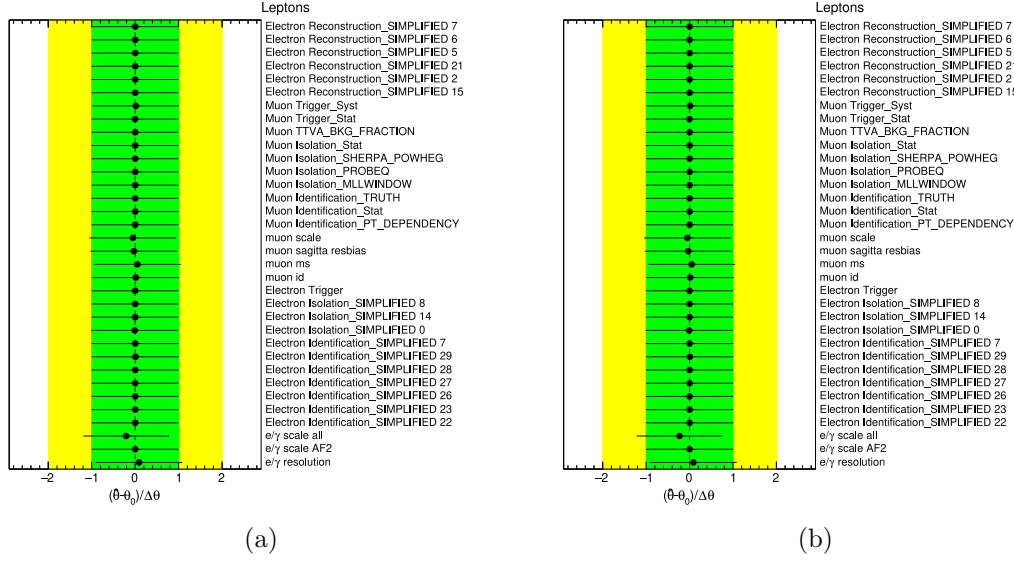


Figure E.14.: The left plot shows the pulls and constraints of those nuisance parameters grouped together as lepton related uncertainties in the fit with one NF, while the right plot shows the same nuisance parameters for the fit with 2 NF for the $l+jets$ channel.

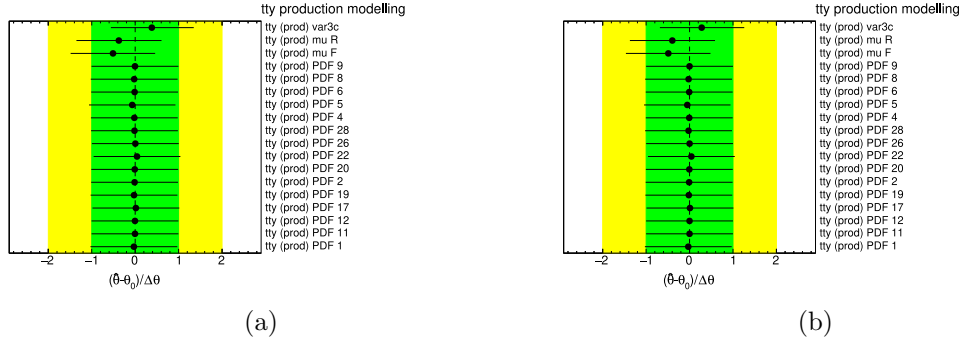


Figure E.15.: The left plot shows the pulls and constraints of those nuisance parameters related to the modelling of the $t\bar{t}\gamma$ from production process (excluding the uncertainty on the parton shower) in the fit with one NF, while the right plot shows the same nuisance parameters for the fit with 2 NF for the $l+jets$ channel.

E. Additional material on inclusive cross-section fits



Figure E.16.: The left plot shows the pulls and constraints of those nuisance parameters grouped together as photon related uncertainties in the fit with one NF, while the right plot shows the same nuisance parameters for the fit with 2 NF for the l+jets channel.



Figure E.17.: The left plot shows the pull and constraint of the pile-up reweighting uncertainty in the fit with one NF, while the right plot shows the same nuisance parameter for the fit with 2 NF for the l+jets channel.

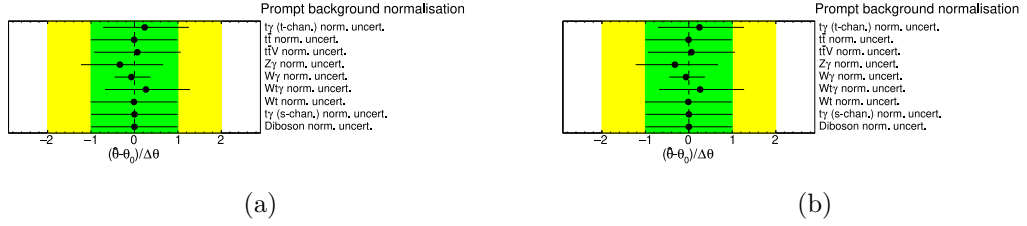


Figure E.18.: The left plot shows the pulls and constraints of those nuisance parameters related to prompt background normalisation uncertainties in the fit with one NF, while the right plot shows the same nuisance parameters for the fit with 2 NF for the l+jets channel.

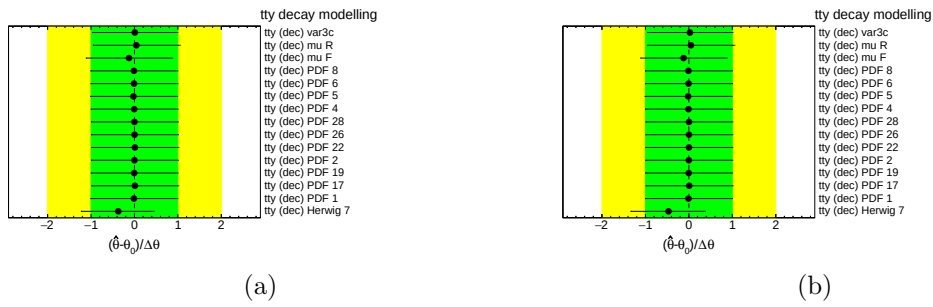


Figure E.19.: The left plot shows the pulls and constraints of those nuisance parameters related to the modelling of the $t\bar{t}\gamma$ from decay process in the fit with one NF, while the right plot shows the same nuisance parameters for the fit with 2 NF for the l+jets channel.

E.2. Additional material on unblinded fits

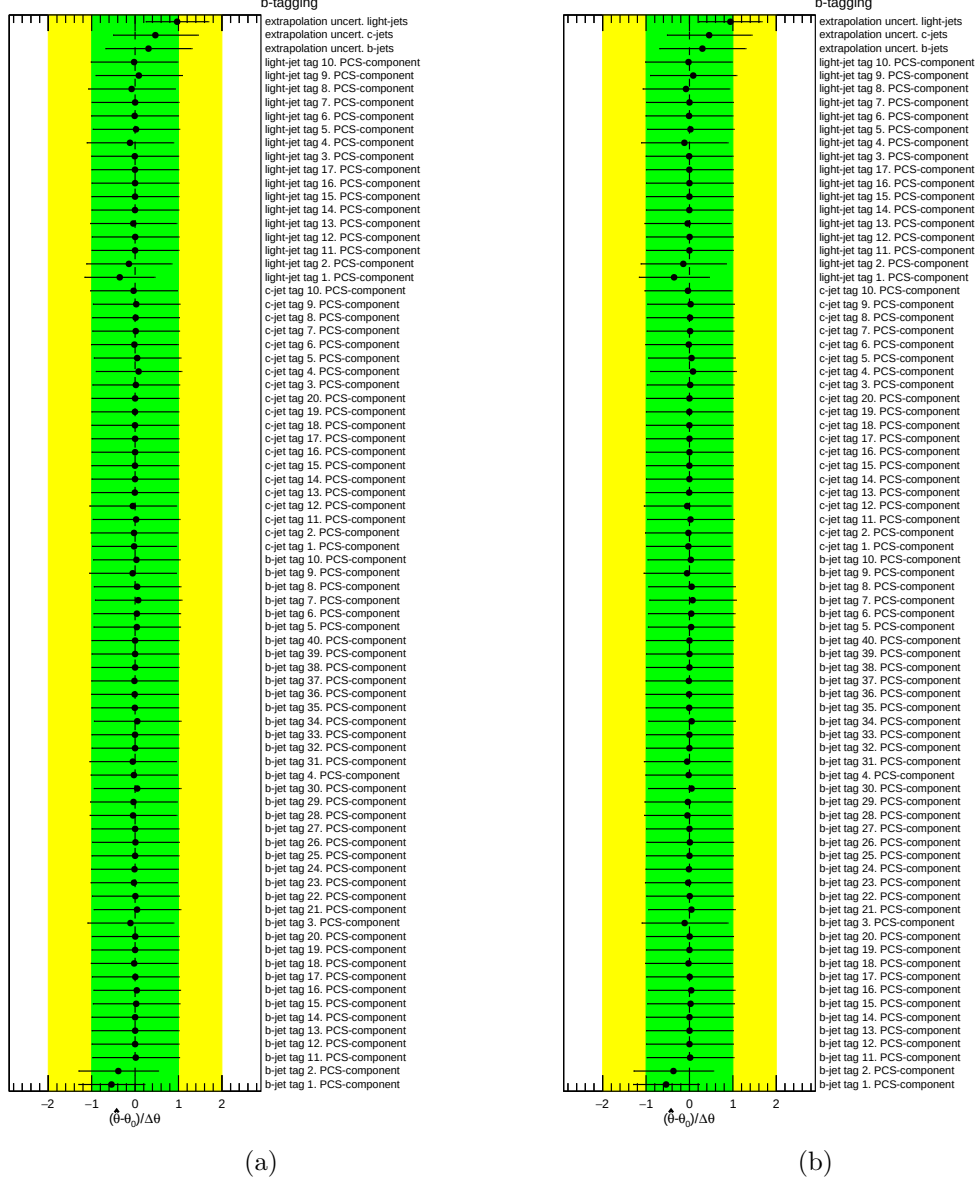


Figure E.20.: The left plot shows the pulls and constraints of those nuisance parameters grouped together as b-tagging uncertainties in the fit with one NF, while the right plot shows the same nuisance parameters for the fit with 2 NF for the l+jets channel.

E. Additional material on inclusive cross-section fits

Correlation matrices for the unblinded fits with 1 and 2 NF in the l+jets channel

Signal Strength	100.0	17.0	14.6	13.6	-28.2	-17.6	4.2	8.2	11.2	3.8	-3.0	7.9	31.7	-59.8	-54.6
b-tagging eigenvariation B-0	17.0	100.0	-34.5	7.5	-2.1	-2.8	-24.6	-4.2	4.2	-0.9	1.4	-4.0	25.2	-2.7	-2.9
b-tagging eigenvariation B-1	14.6	-34.5	100.0	4.6	-3.0	-2.4	-9.9	-1.2	3.1	-2.4	0.2	-0.1	15.1	-0.8	-9.0
jet flavor response	13.6	7.5	4.6	100.0	14.0	0.9	21.6	2.5	8.0	6.3	12.6	-1.7	0.4	-1.8	-1.1
jet pileup rho topology	-28.2	-2.1	-3.0	14.0	100.0	-1.7	-6.6	-6.5	-18.7	-4.2	-26.9	5.8	-4.2	3.3	5.7
W γ norm. uncert.	-17.6	-2.8	-2.4	0.9	-1.7	100.0	-23.0	-0.8	0.9	1.8	-3.6	1.0	-2.9	-1.0	1.6
W γ norm. uncert.	4.2	-24.6	-9.9	21.6	-6.6	-23.0	100.0	-45.5	6.6	-18.4	3.7	-16.1	-2.9	-2.7	-7.5
Z γ norm. uncert.	8.2	-4.2	-1.2	2.5	-6.5	-0.8	-45.5	100.0	-1.3	-0.4	-2.2	-0.1	6.4	2.0	-6.7
efake γ SF	11.2	4.2	3.1	8.0	-18.7	0.9	6.6	-1.3	100.0	-8.7	-27.0	1.9	-15.7	4.8	5.5
extrapolation uncert. light-jets	3.8	-0.9	-2.4	6.3	-4.2	1.8	-18.4	-0.4	-8.7	100.0	11.0	21.2	-14.1	-6.6	18.0
hfake γ SF	-3.0	1.4	0.2	12.6	-26.9	-3.6	3.7	-2.2	-27.0	11.0	100.0	-6.2	-20.7	6.3	13.4
t $\bar{t}\gamma$ (dec) H7	7.9	-4.0	-0.1	-1.7	5.8	1.0	-16.1	-0.1	1.9	21.2	-6.2	100.0	2.0	2.2	55.0
t $\bar{t}\gamma$ (prod) H7	31.7	25.2	15.1	0.4	-4.2	-2.9	-2.9	6.4	-15.7	-14.1	-20.7	2.0	100.0	-0.7	-9.7
t $\bar{t}\gamma$ (prod) var3c	-59.8	-2.7	-0.8	-1.8	3.3	-1.0	-2.7	2.0	4.8	-6.6	6.3	2.2	-0.7	100.0	53.8
t $\bar{t}\gamma$ (dec.) norm. uncert. (k-factor)	-54.6	-2.9	-9.0	-1.1	5.7	1.6	-7.5	-6.7	5.5	18.0	13.4	55.0	-9.7	53.8	100.0

Figure E.21.: The correlation matrix of the fit with one NF for the l+jets channel. Only nuisance parameters with at least one correlation of at least $\pm 20\%$ are shown.

E.2. Additional material on unblinded fits

Signal Strength	100.0	16.8	15.0	13.4	-28.1	-6.1	-16.8	4.4	8.5	10.3	2.7	-3.6	4.1	31.4	-20.0	-61.2	31.1
b-tagging eigenvariation B-0	16.8	100.0	-34.3	8.0	-2.2	0.7	-2.5	-23.8	-4.6	4.4	-1.1	1.2	-4.1	25.4	0.4	-2.7	12.8
b-tagging eigenvariation B-1	15.0	-34.3	100.0	4.9	-2.9	0.6	-2.3	-10.2	-1.1	2.4	-2.2	0.6	-0.4	15.2	0.8	-1.4	3.6
jet flavor response	13.4	8.0	4.9	100.0	13.9	3.5	1.3	21.6	2.2	7.5	6.5	13.8	-1.6	0.5	-0.3	-1.8	11.8
jet pileup rho topology	-28.1	-2.2	-2.9	13.9	100.0	-6.5	-2.1	-6.4	-6.5	-18.6	-4.1	-26.2	5.8	-4.2	-0.0	3.7	-20.2
Photon Isolation	-6.1	0.7	0.6	3.5	-6.5	100.0	1.2	-1.3	-0.7	-7.7	-6.7	-7.4	-1.0	-3.1	0.4	2.3	-20.4
Wty norm. uncert.	-16.8	-2.5	-2.3	1.3	-2.1	1.2	100.0	-23.1	-0.7	0.9	2.1	-3.1	1.5	-2.4	0.6	-0.9	-14.1
Wy norm. uncert.	4.4	-23.8	-10.2	21.6	-6.4	-1.3	-23.1	100.0	-45.7	6.0	-18.4	4.4	-16.6	-3.1	0.8	-3.0	-4.9
Zy norm. uncert.	8.5	-4.6	-1.1	2.2	-6.5	-0.7	-0.7	-45.7	100.0	-0.9	-0.8	-3.4	-0.5	6.6	-2.7	1.6	-0.2
efake γ SF	10.3	4.4	2.4	7.5	-18.6	-7.7	0.9	6.0	-0.9	100.0	-8.6	-27.2	2.2	-16.8	1.0	5.0	16.8
extrapolation uncert. light-jets	2.7	-1.1	-2.2	6.5	-4.1	-6.7	2.1	-18.4	-0.8	-8.6	100.0	11.3	22.1	-14.1	15.2	-5.5	24.4
hfake γ SF	-3.6	1.2	0.6	13.8	-26.2	-7.4	-3.1	4.4	-3.4	-27.2	11.3	100.0	-4.8	-20.0	-0.7	6.9	12.8
tty (dec) H7	4.1	-4.1	-0.4	-1.6	5.8	-1.0	1.5	-16.6	-0.5	2.2	22.1	-4.8	100.0	1.3	-5.2	5.6	70.3
tty (prod) H7	31.4	25.4	15.2	0.5	-4.2	-3.1	-2.4	-3.1	6.6	-16.8	-14.1	-20.0	1.3	100.0	9.4	-1.1	18.8
tty (prod) mu R	-20.0	0.4	0.8	-0.3	-0.0	0.4	0.6	0.8	-2.7	1.0	15.2	-0.7	-5.2	9.4	100.0	1.8	-6.6
tty (prod) var3c	-61.2	-2.7	-1.4	-1.8	3.7	2.3	-0.9	-3.0	1.6	5.0	-5.5	6.9	5.6	-1.1	1.8	100.0	5.3
Norm factor for tty-dec	31.1	12.8	3.6	11.8	-20.2	-20.4	-14.1	-4.9	-0.2	16.8	24.4	12.8	70.3	18.8	-6.6	5.3	100.0
	Signal Strength	b-tagging eigenvariation B-0	b-tagging eigenvariation B-1	jet flavor response	jet pileup rho topology	Photon Isolation	Wty norm. uncert.	Wy norm. uncert.	Zy norm. uncert.	efake γ SF	extrapolation uncert. light-jets	hfake γ SF	tty (dec) H7	tty (prod) H7	tty (prod) mu R	tty (prod) var3c	Norm factor for tty-dec

Figure E.22.: The correlation matrix of the fit with two NFs for the 1+jets channel. Only nuisance parameters with at least one correlation of at least $\pm 20\%$ are shown.

Pull-plots for the unblinded fits with 1 and 2 NF in the dilepton channel

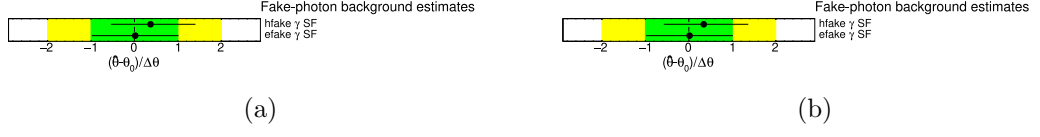


Figure E.23.: The left plot shows the pulls and constraints of those nuisance parameters grouped together as fake photon background estimate uncertainties in the fit with one NF, while the right plot shows the same nuisance parameters for the fit with 2 NF in the dilepton channel.



Figure E.24.: The left plot shows the pull and constraint of the luminosity uncertainty in the fit with one NF, while the right plot shows the same nuisance parameter for the fit with 2 NF in the dilepton channel.



Figure E.25.: The left plot shows the pulls and constraints of those nuisance parameters grouped together as E_T^{miss} related uncertainties in the fit with one NF, while the right plot shows the same nuisance parameters for the fit with 2 NF in the dilepton channel.

E.2. Additional material on unblinded fits

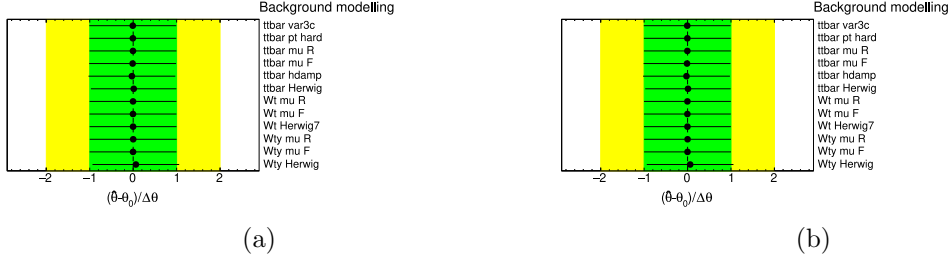


Figure E.26.: The left plot shows the pulls and constraints of those nuisance parameters grouped together as background modelling uncertainties in the fit with one NF, while the right plot shows the same nuisance parameters for the fit with 2 NF in the dilepton channel.

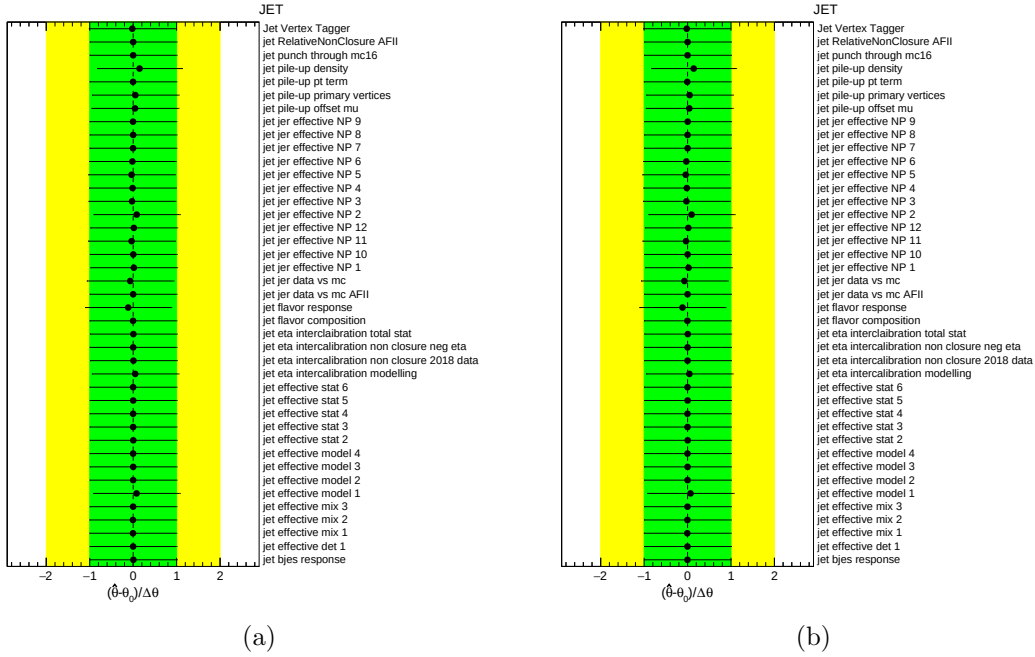


Figure E.27.: The left plot shows the pulls and constraints of those nuisance parameters grouped together as jet related uncertainties in the fit with one NF, while the right plot shows the same nuisance parameters for the fit with 2 NF in the dilepton channel.

E. Additional material on inclusive cross-section fits

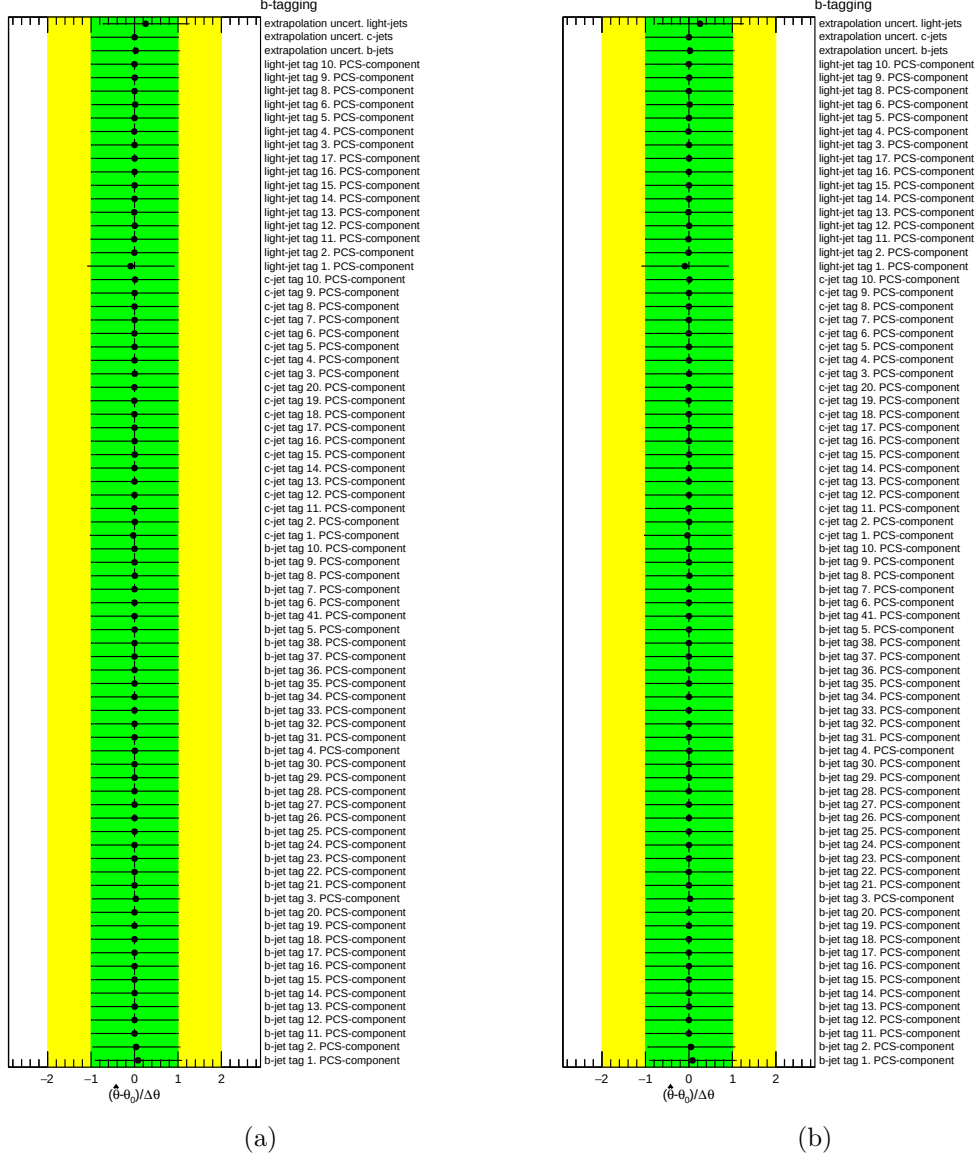


Figure E.28.: The left plot shows the pulls and constraints of those nuisance parameters grouped together as b-tagging uncertainties in the fit with one NF, while the right plot shows the same nuisance parameters for the fit with 2 NF in the dilepton channel.

E.2. Additional material on unblinded fits

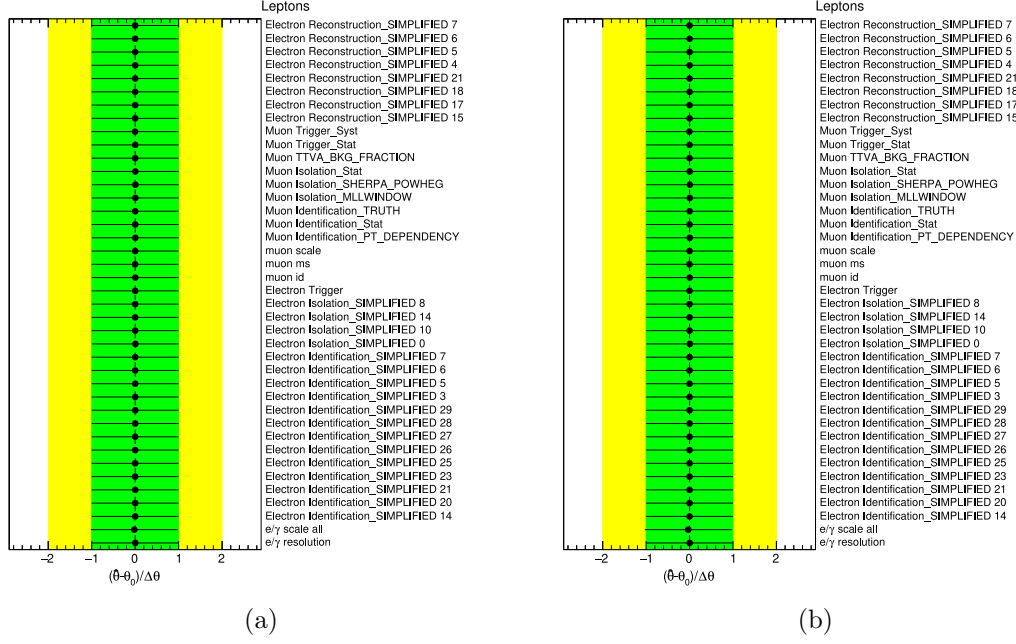


Figure E.29.: The left plot shows the pulls and constraints of those nuisance parameters grouped together as lepton related uncertainties in the fit with one NF, while the right plot shows the same nuisance parameters for the fit with 2 NF in the dilepton channel.



Figure E.30.: The left plot shows the pulls and constraints of those nuisance parameters grouped together as photon related uncertainties in the fit with one NF, while the right plot shows the same nuisance parameters for the fit with 2 NF in the dilepton channel.



Figure E.31.: The left plot shows the pull and constraint of the pile-up reweighting uncertainty in the fit with one NF, while the right plot shows the same nuisance parameter for the fit with 2 NF in the dilepton channel.

E. Additional material on inclusive cross-section fits

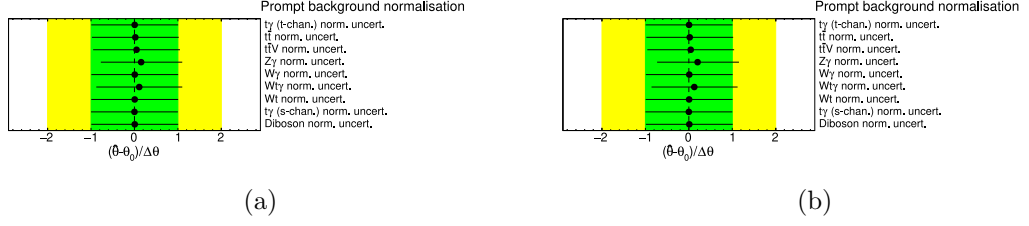


Figure E.32.: The left plot shows the pulls and constraints of those nuisance parameters related to prompt background normalisation uncertainties in the fit with one NF, while the right plot shows the same nuisance parameters for the fit with 2 NF in the dilepton channel.

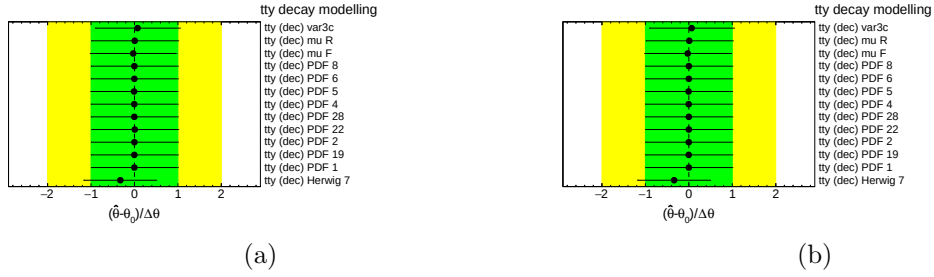


Figure E.33.: The left plot shows the pulls and constraints of those nuisance parameters related to the modelling of the $t\bar{t}\gamma$ from decay process in the fit with one NF, while the right plot shows the same nuisance parameters for the fit with 2 NF in the dilepton channel.

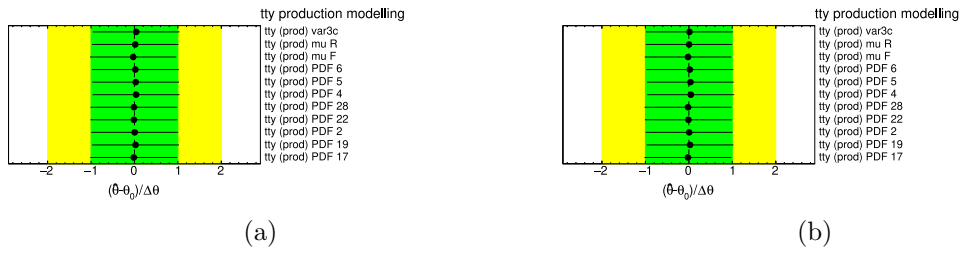


Figure E.34.: The left plot shows the pulls and constraints of those nuisance parameters related to the modelling of the $t\bar{t}\gamma$ from production process (excluding the uncertainty on the parton shower) in the fit with one NF, while the right plot shows the same nuisance parameters for the fit with 2 NF in the dilepton channel.



Figure E.35.: The left plot shows the pull and constraint of the parton shower uncertainty for the $t\bar{t}\gamma$ from production sample in the fit with one NF, while the right plot shows the same nuisance parameters for the fit with 2 NF in the dilepton channel.

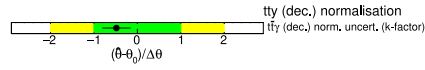


Figure E.36.: Pull and constraint of the normalisation uncertainty for the $t\bar{t}\gamma$ from decay sample in the fit with one NF in the dilepton channel.

E. Additional material on inclusive cross-section fits

Correlation matrices for the unblinded fits with 1 and 2 NF in the dilepton channel

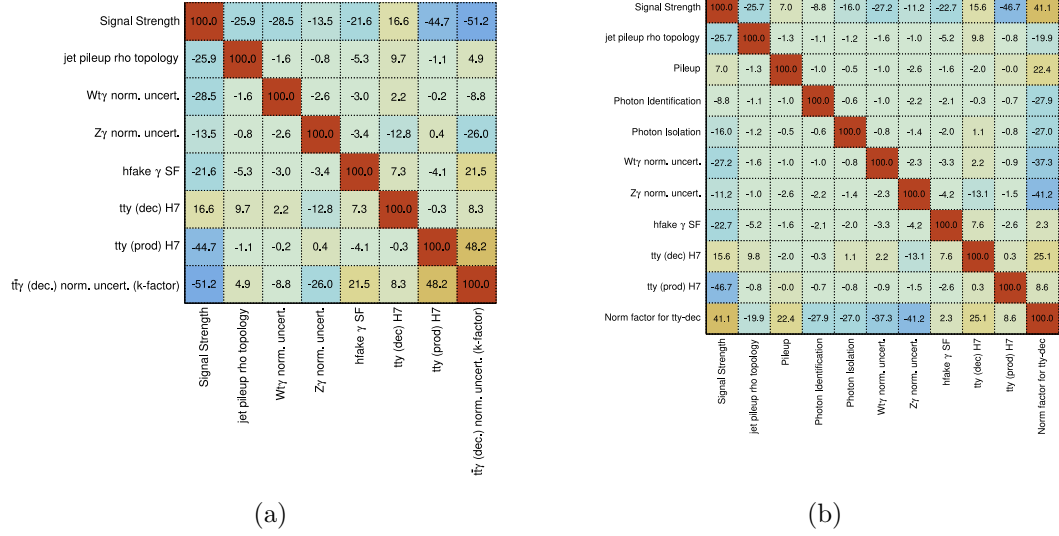


Figure E.37.: The correlation matrix of the fits in the dilepton channel with one NF on the left side and with two NF on the right side. Only nuisance parameters with at least one correlation of at least $\pm 20\%$ are shown.

Correlation matrices for the unblinded fits with 1 and 2 NF in the combination of l+jets and dilepton channels

Signal Strength	100.0	5.4	7.2	5.5	-19.8	-6.4	-32.8	18.2	-13.8	30.2	1.9	16.1	-17.2	-24.7	-47.7	-39.4
b-tagging eigenvariation B-0	5.4	100.0	-38.5	5.0	2.1	0.9	-5.8	-23.3	-10.3	8.6	1.8	-0.4	17.7	1.1	6.5	11.0
b-tagging eigenvariation B-1	7.2	-38.5	100.0	2.6	0.1	0.1	-4.3	-9.1	-5.5	5.3	0.7	0.6	8.5	1.5	5.4	-3.3
jet flavor response	5.5	5.0	2.6	100.0	18.4	2.6	-2.0	28.9	-5.3	10.1	14.8	-4.6	-8.7	0.4	4.3	4.4
jet pileup rho topology	-19.8	2.1	0.1	18.4	100.0	-5.1	3.2	-14.3	3.6	-22.6	-28.3	12.9	9.1	-1.3	-4.8	1.1
Photon Identification	-6.4	0.9	0.1	2.6	-5.1	100.0	-2.3	-1.2	-7.1	-5.8	-9.2	-4.7	-4.5	-0.0	4.5	-22.3
W γ norm. uncert.	-32.8	-5.8	-4.3	-2.0	3.2	-2.3	100.0	-14.8	-15.9	3.1	-5.1	-7.0	-8.1	2.2	0.6	-3.4
W γ norm. uncert.	18.2	-23.3	-9.1	28.9	-14.3	-1.2	-14.8	100.0	-33.9	3.6	7.3	-2.3	4.0	-1.1	-6.2	1.6
Z γ norm. uncert.	-13.8	-10.3	-5.5	-5.3	3.6	-7.1	-15.9	-33.9	100.0	3.6	-5.0	-31.3	-9.1	0.6	10.1	-24.0
efake γ SF	30.2	8.6	5.3	10.1	-22.6	-5.8	3.1	3.6	3.6	100.0	-26.9	-4.8	-9.3	0.6	-2.3	-6.3
hfake γ SF	1.9	1.8	0.7	14.8	-28.3	-9.2	-5.1	7.3	-5.0	-26.9	100.0	-5.3	-21.4	-1.2	4.8	15.7
t $\bar{t}\gamma$ (dec) H7	16.1	-0.4	0.6	-4.6	12.9	-4.7	-7.0	-2.3	-31.3	-4.8	-5.3	100.0	7.0	-5.0	-5.6	45.4
t $\bar{t}\gamma$ (prod) H7	-17.2	17.7	8.5	-8.7	9.1	-4.5	-8.1	4.0	-9.1	-9.3	-21.4	7.0	100.0	13.4	33.3	33.2
t $\bar{t}\gamma$ (prod) mu R	-24.7	1.1	1.5	0.4	-1.3	-0.0	2.2	-1.1	0.6	0.6	-1.2	-5.0	13.4	100.0	0.2	14.4
t $\bar{t}\gamma$ (prod) var3c	-47.7	6.5	5.4	4.3	-4.8	4.5	0.6	-6.2	10.1	-2.3	4.8	-5.6	33.3	0.2	100.0	44.0
t $\bar{t}\gamma$ (dec.) norm. uncert. (k-factor)	-39.4	11.0	-3.3	4.4	1.1	-22.3	-3.4	1.6	-24.0	-6.3	15.7	45.4	33.2	14.4	44.0	100.0
Signal Strength																
b-tagging eigenvariation B-0																
b-tagging eigenvariation B-1																
jet flavor response																
jet pileup rho topology																
Photon Identification																
W γ norm. uncert.																
W γ norm. uncert.																
Z γ norm. uncert.																
efake γ SF																
hfake γ SF																
t $\bar{t}\gamma$ (dec) H7																
t $\bar{t}\gamma$ (prod) H7																
t $\bar{t}\gamma$ (prod) mu R																
t $\bar{t}\gamma$ (prod) var3c																
t $\bar{t}\gamma$ (dec.) norm. uncert. (k-factor)																

Figure E.38.: The correlation matrix of the fit combining l+jets and dilepton channels with one NF. Only nuisance parameters with at least one correlation of at least $\pm 20\%$ are shown.

E. Additional material on inclusive cross-section fits

Signal Strength	100.0	5.2	7.2	5.4	-19.8	10.1	-5.9	-14.2	-32.6	18.1	-13.2	30.3	1.6	15.0	-17.8	-24.9	-48.1	40.7
b-tagging eigenvariation B-0	5.2	100.0	-38.5	5.0	2.1	-2.6	0.7	-0.5	-5.8	-23.4	-10.4	8.6	2.0	-0.0	17.9	1.2	6.8	15.5
b-tagging eigenvariation B-1	7.2	-38.5	100.0	2.6	0.1	-2.2	0.2	-0.2	-4.3	-9.1	-5.5	5.3	0.7	0.5	8.4	1.4	5.3	2.5
jet flavor response	5.4	5.0	2.6	100.0	18.4	2.5	2.5	2.5	-2.0	29.0	-5.4	10.1	14.7	-4.4	-8.5	0.5	4.4	8.9
jet pileup rho topology	-19.8	2.1	0.1	18.4	100.0	-1.9	-5.1	-5.1	3.2	-14.3	3.6	-22.7	-28.2	12.8	9.1	-1.3	-4.7	-14.8
Pileup	10.1	-2.6	-2.2	2.5	-1.9	100.0	-0.4	-0.4	2.6	-0.8	-3.1	1.4	-1.1	-0.5	3.1	0.3	0.6	25.5
Photon Identification	-5.9	0.7	0.2	2.5	-5.1	-0.4	100.0	-2.9	-2.2	-1.2	-6.7	-5.8	-9.3	-5.3	-5.0	-0.2	3.8	-27.6
Photon Isolation	-14.2	-0.5	-0.2	2.5	-5.1	-0.4	-2.9	100.0	-0.3	1.0	-4.3	-7.0	-7.6	-2.3	-7.4	0.8	4.9	-27.0
W γ norm. uncert.	-32.6	-5.8	-4.3	-2.0	3.2	2.6	-2.2	-0.3	100.0	-14.9	-15.7	3.1	-5.1	-7.2	-8.2	2.2	0.5	-30.0
W γ norm. uncert.	18.1	-23.4	-9.1	29.0	-14.3	-0.8	-1.2	1.0	-14.9	100.0	-33.7	3.6	7.2	-2.2	4.0	-1.1	-6.1	16.3
Z γ norm. uncert.	-13.2	-10.4	-5.5	-5.4	3.6	-3.1	-6.7	-4.3	-15.7	-33.7	100.0	3.6	-5.2	-31.9	-9.6	0.5	9.4	-35.5
efake γ SF	30.3	8.6	5.3	10.1	-22.7	1.4	-5.8	-7.0	3.1	3.6	3.6	100.0	-26.8	-4.9	-9.5	0.5	-2.4	18.0
hfake γ SF	1.6	2.0	0.7	14.7	-28.2	-1.1	-9.3	-7.6	-5.1	7.2	-5.2	-26.8	100.0	-4.7	-20.9	-1.1	5.2	17.4
tty (dec) H7	15.0	-0.0	0.5	-4.4	12.8	-0.5	-5.3	-2.3	-7.2	-2.2	-31.9	-4.9	-4.7	100.0	7.9	-4.7	-4.3	58.8
tty (prod) H7	-17.8	17.9	8.4	-8.5	9.1	3.1	-5.0	-7.4	-8.2	4.0	-9.6	-9.5	-20.9	7.9	100.0	13.6	33.8	19.7
tty (prod) mu R	-24.9	1.2	1.4	0.5	-1.3	0.3	-0.2	0.8	2.2	-1.1	0.5	0.5	-1.1	-4.7	13.6	100.0	0.6	-5.5
tty (prod) var3c	-48.1	6.8	5.3	4.4	-4.7	0.6	3.8	4.9	0.5	-6.1	9.4	-2.4	5.2	-4.3	33.8	0.6	100.0	6.0
Norm factor for tty-dec	40.7	15.5	2.5	8.9	-14.8	25.5	-27.6	-27.0	-30.0	16.3	-35.5	18.0	17.4	58.8	19.7	-5.5	6.0	100.0
Signal Strength																		
b-tagging eigenvariation B-0																		
b-tagging eigenvariation B-1																		
jet flavor response																		
jet pileup rho topology																		
Pileup																		
Photon Identification																		
Photon Isolation																		
W γ norm. uncert.																		
W γ norm. uncert.																		
Z γ norm. uncert.																		
efake γ SF																		
hfake γ SF																		
tty (dec) H7																		
tty (prod) H7																		
tty (prod) mu R																		
tty (prod) var3c																		
Norm factor for tty-dec																		

Figure E.39.: The correlation matrix of the fit combining l+jets and dilepton channels with two NFs. Only nuisance parameters with at least one correlation of at least $\pm 20\%$ are shown.

E.3. Post-fit plots

E.3.1. $t\bar{t}\gamma$ production: Post-fit plots for the fit in the lepton+jets channel with two NF

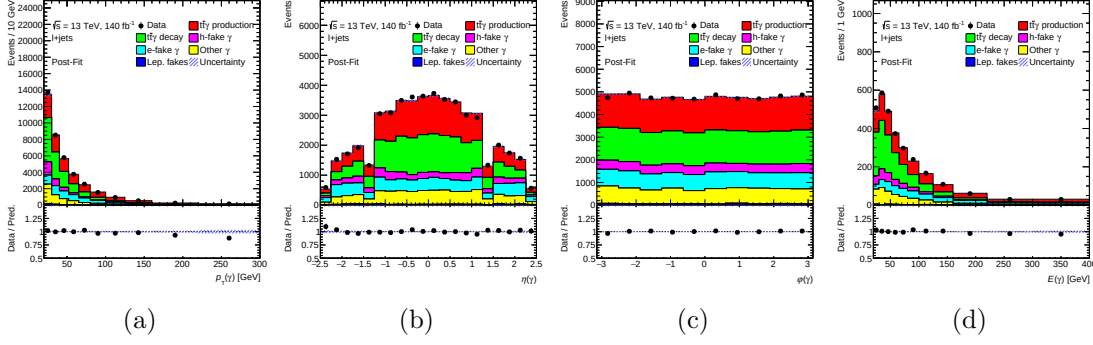


Figure E.40.: Post-fit plots for the p_T , η , ϕ and E distributions of the photon (from left to right). The uncertainty band corresponds to the total uncertainty, including stat. and syst. uncertainties.

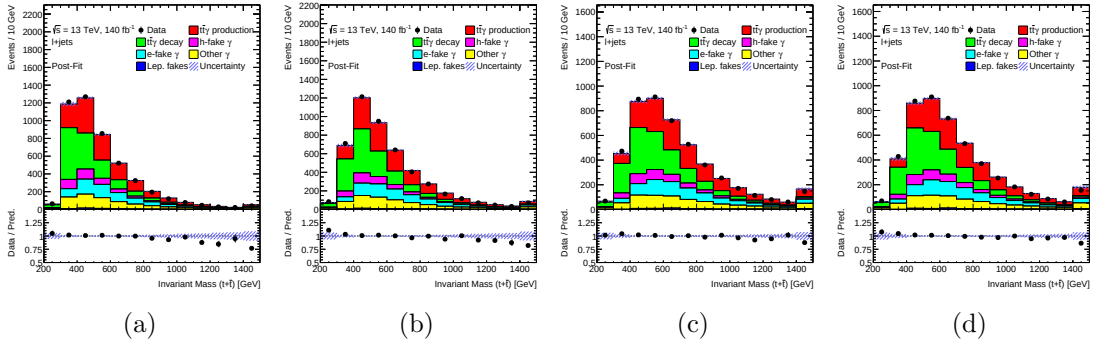


Figure E.41.: Post-fit plots for the distributions of the sum of the reconstructed $t+\bar{t}$ mass for all four permutations to reconstruct them (from left (best permutation) to right (worst permutation)). The uncertainty band corresponds to the total uncertainty, including stat. and syst. uncertainties.

E. Additional material on inclusive cross-section fits

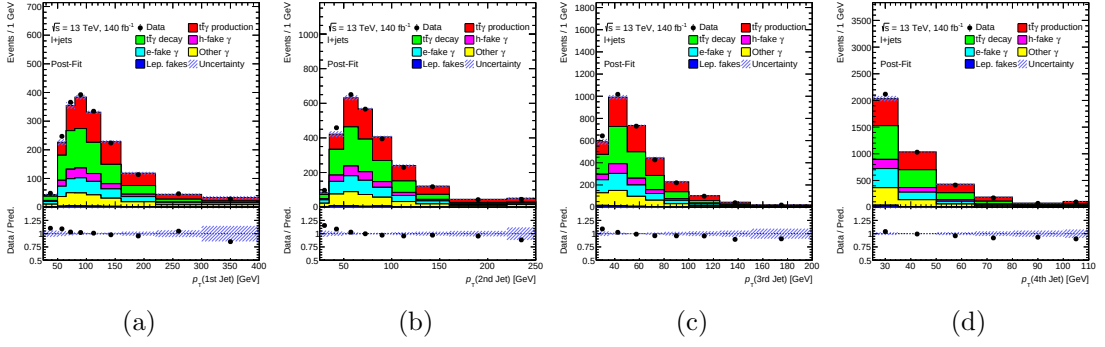


Figure E.42.: Post-fit plots for the p_T distributions of the first four jets (from left to right). The uncertainty band corresponds to the total uncertainty, including stat. and syst. uncertainties.

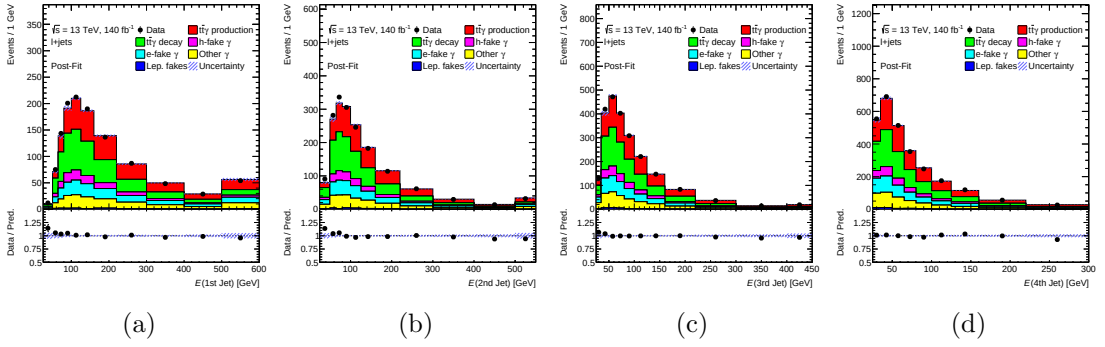


Figure E.43.: Post-fit plots for the E distributions of the first four jets (from left to right). The uncertainty band corresponds to the total uncertainty, including stat. and syst. uncertainties.

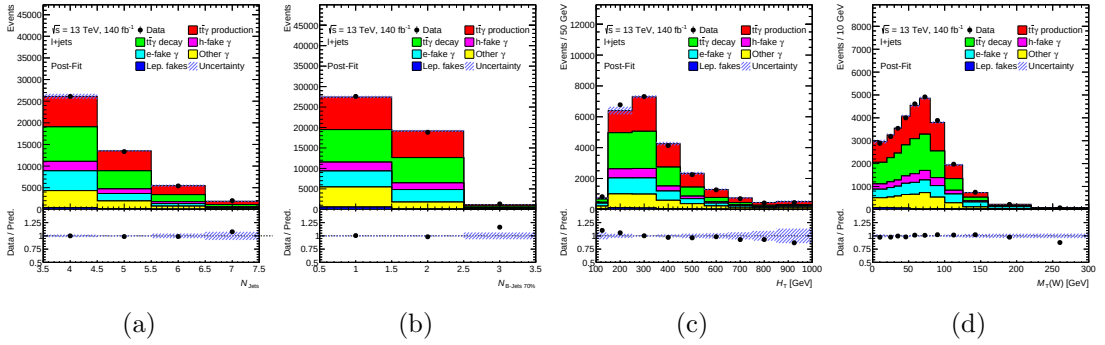


Figure E.44.: Post-fit plots for the distributions of the number of jets and b -jets per event, as well as the H_T and $m_T(W)$ distributions (from left to right). The uncertainty band corresponds to the total uncertainty, including stat. and syst. uncertainties.

E.3. Post-fit plots

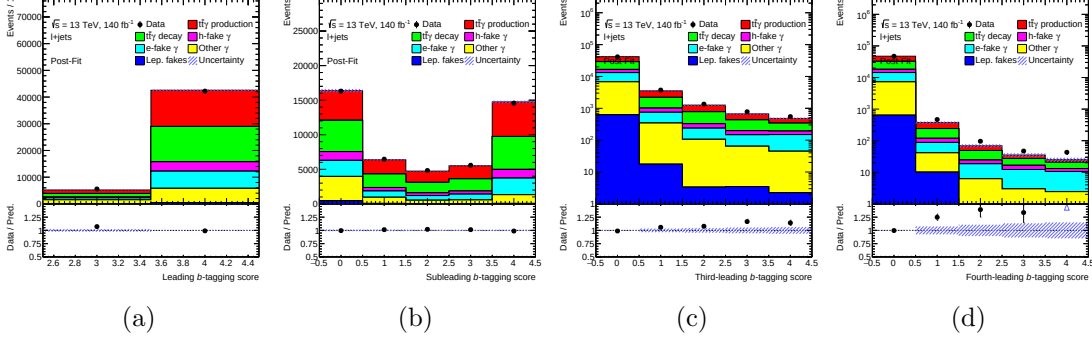


Figure E.45.: Post-fit plots for the PCBT distributions of the four leading PCBT jets (from left to right). The uncertainty band corresponds to the total uncertainty, including stat. and syst. uncertainties.

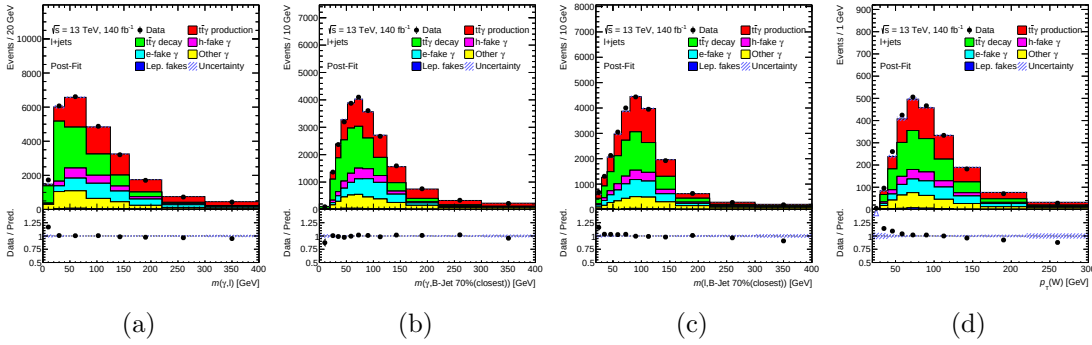


Figure E.46.: Post-fit plots for the invariant mass distributions of the photon and the lepton, of the photon and the closest b -jet and of the lepton and the closest b -jet, as well as the W -boson p_T distribution (from left to right). The uncertainty band corresponds to the total uncertainty, including stat. and syst. uncertainties.

E. Additional material on inclusive cross-section fits

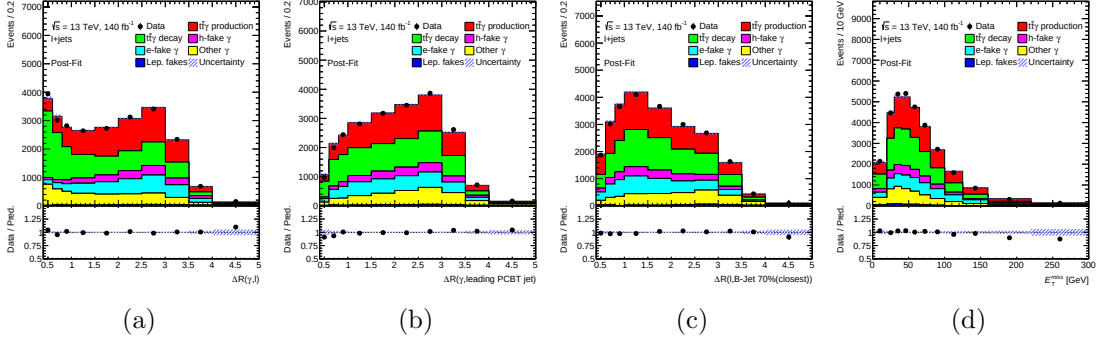


Figure E.47.: Post-fit plots for the ΔR distributions of the photon and the lepton, of the photon and the leading PCBT jet and of the lepton and the closest b -jet, as well as the E_T^{miss} distribution (from left to right). The uncertainty band corresponds to the total uncertainty, including stat. and syst. uncertainties.

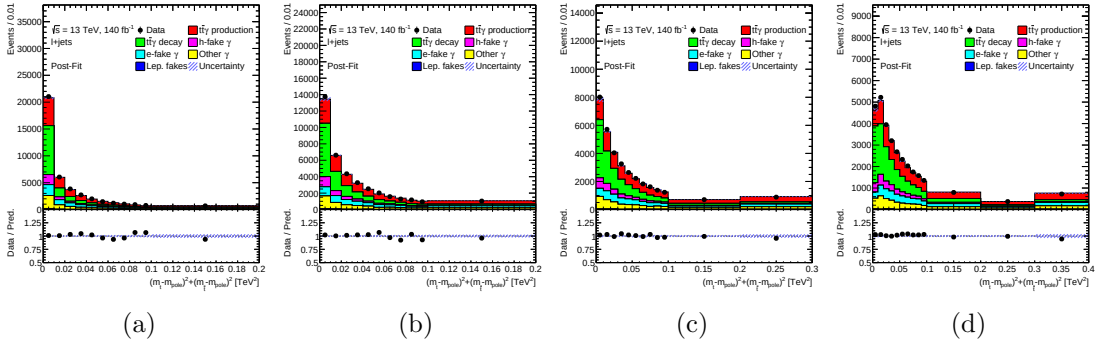


Figure E.48.: Post-fit plots for the distributions showing the squared differences of the reconstructed t and $\bar{t}$ masses concerning the top quark pole-mass for all four permutations to reconstruct them (from left (best permutation) to right (worst permutation)). The uncertainty band corresponds to the total uncertainty, including stat. and syst. uncertainties.

E.3. Post-fit plots

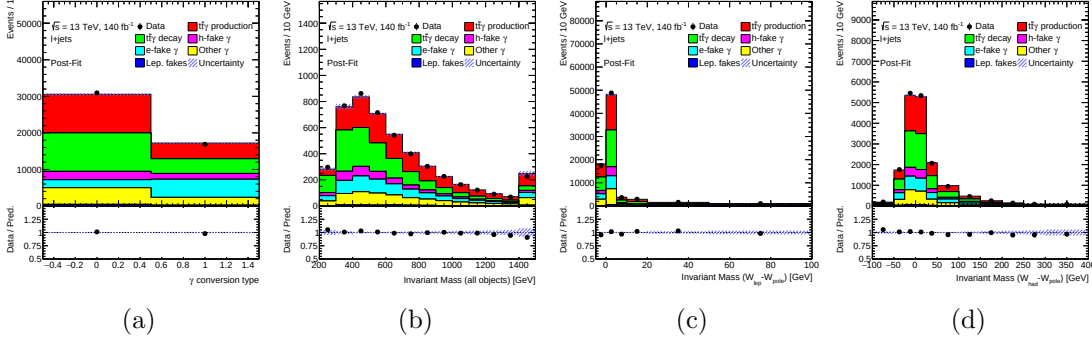


Figure E.49.: Post-fit plots for the photon conversion type distribution, for the invariant mass distribution of all reconstructed objects and for the distributions of the difference between the reconstructed W -boson mass (leptonic and hadronic decay) and the W -boson pole mass (from left to right). The uncertainty band corresponds to the total uncertainty, including stat. and syst. uncertainties.

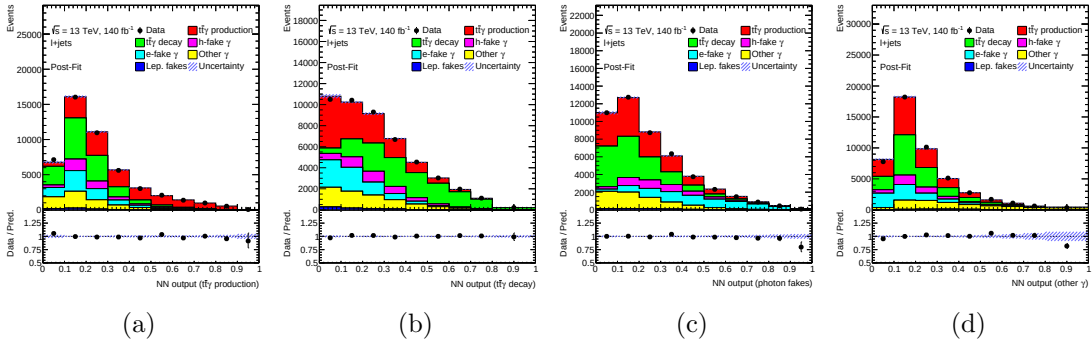


Figure E.50.: Post-fit plots for the four NN output distributions. The uncertainty band corresponds to the total uncertainty, including stat. and syst. uncertainties.

E.3.2. Inclusive $t\bar{t}\gamma$: Post-fit plots for the fit in the lepton+jets channel with one NF

E. Additional material on inclusive cross-section fits

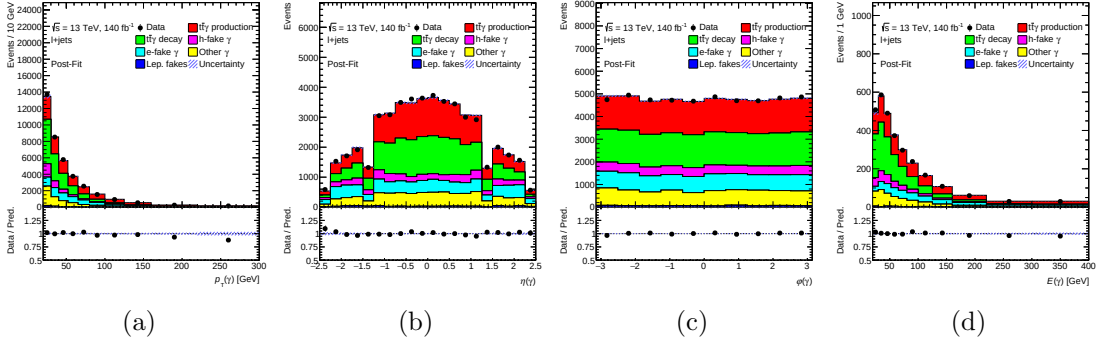


Figure E.51.: Post-fit plots for the p_T , η , ϕ and E distributions of the photon (from left to right). The uncertainty band corresponds to the total uncertainty, including stat. and syst. uncertainties.

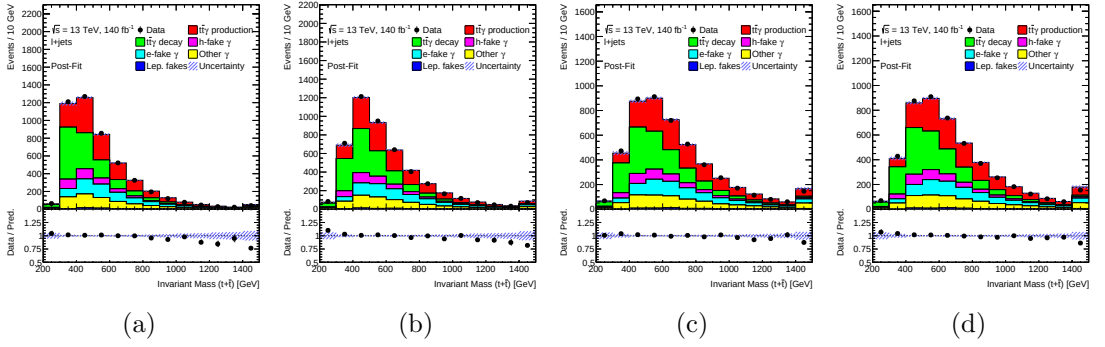


Figure E.52.: Post-fit plots for the distributions of the sum of the reconstructed $t + \bar{t}$ mass for all four permutations to reconstruct them (from left (best permutation) to right (worst permutation)). The uncertainty band corresponds to the total uncertainty, including stat. and syst. uncertainties.

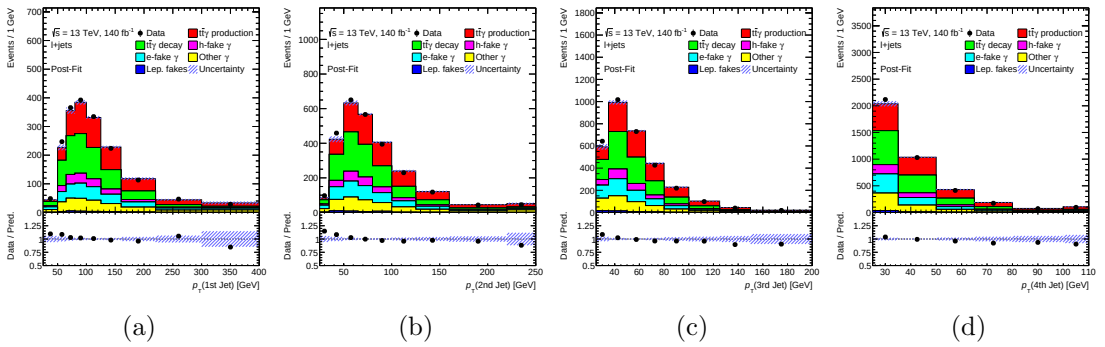


Figure E.53.: Post-fit plots for the p_T distributions of the first four jets (from left to right). The uncertainty band corresponds to the total uncertainty, including stat. and syst. uncertainties.

E.3. Post-fit plots

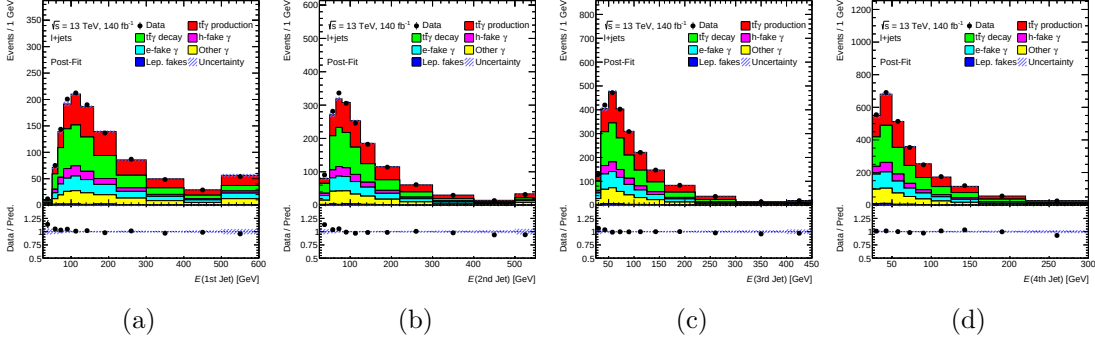


Figure E.54.: Post-fit plots for the E distributions of the first four jets (from left to right). The uncertainty band corresponds to the total uncertainty, including stat. and syst. uncertainties.

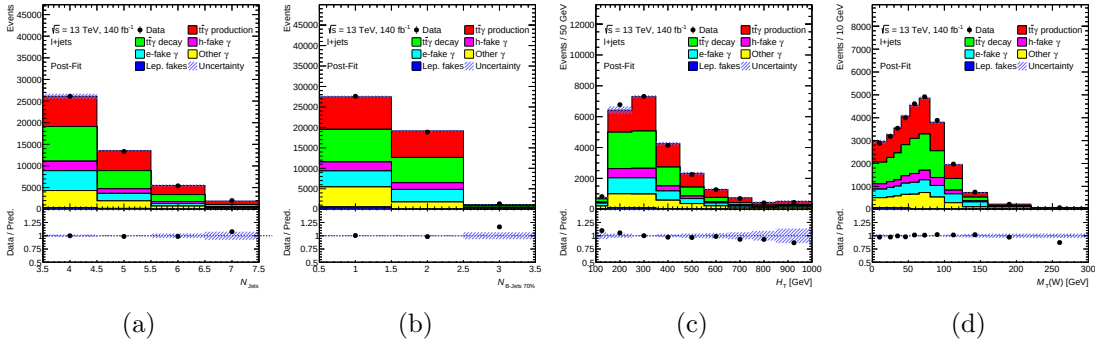


Figure E.55.: Post-fit plots for the distributions of the number of jets and b -jets per event, as well as the H_T and $m_T(W)$ distributions (from left to right). The uncertainty band corresponds to the total uncertainty, including stat. and syst. uncertainties.

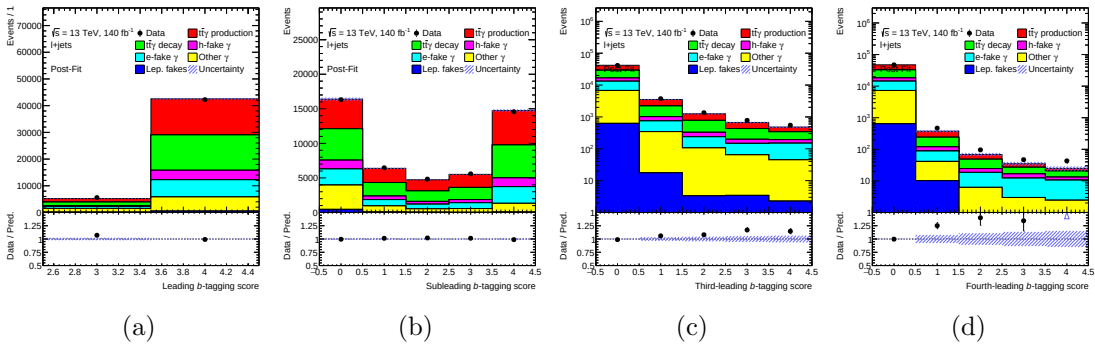


Figure E.56.: Post-fit plots for the PCBT distributions of the four leading PCBT jets (from left to right). The uncertainty band corresponds to the total uncertainty, including stat. and syst. uncertainties.

E. Additional material on inclusive cross-section fits

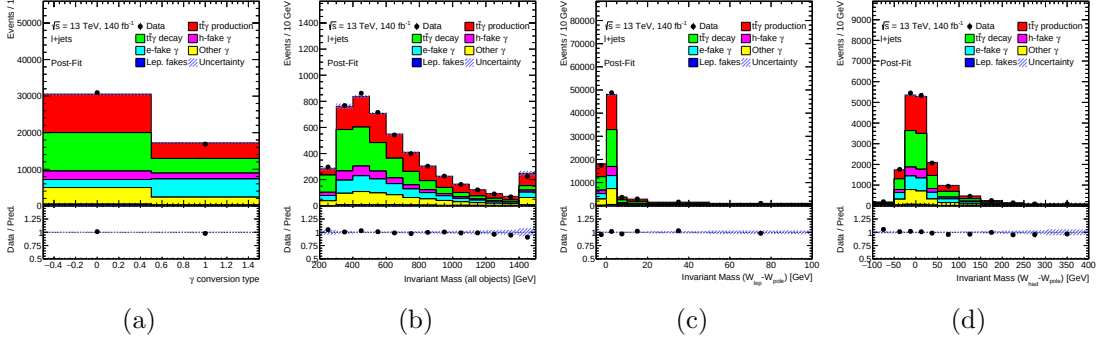


Figure E.57.: Post-fit plots for the photon conversion type distribution, for the invariant mass distribution of all reconstructed objects and for the distributions of the difference between the reconstructed W -boson mass (leptonic and hadronic decay) and the W -boson pole mass (from left to right). The uncertainty band corresponds to the total uncertainty, including stat. and syst. uncertainties.

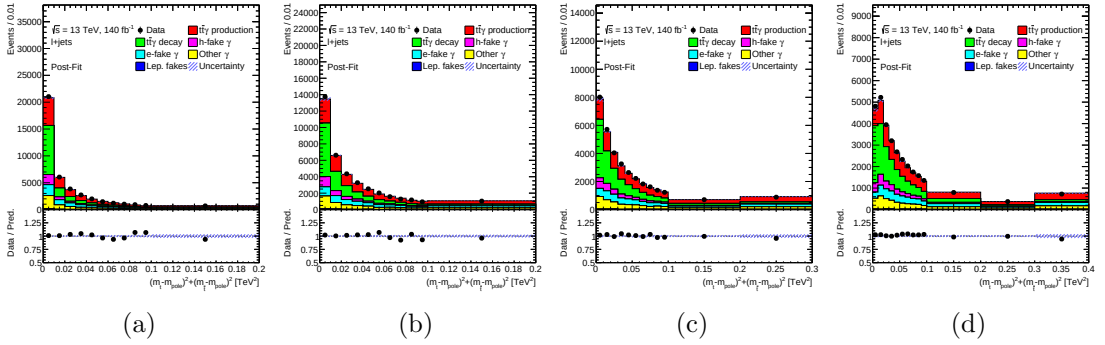


Figure E.58.: Post-fit plots for the distributions of the squared differences of the reconstructed t and $\bar{t}$ masses concerning the top quark pole-mass for all four permutations to reconstruct them (from left (best permutation) to right (worst permutation)). The uncertainty band corresponds to the total uncertainty, including stat. and syst. uncertainties.

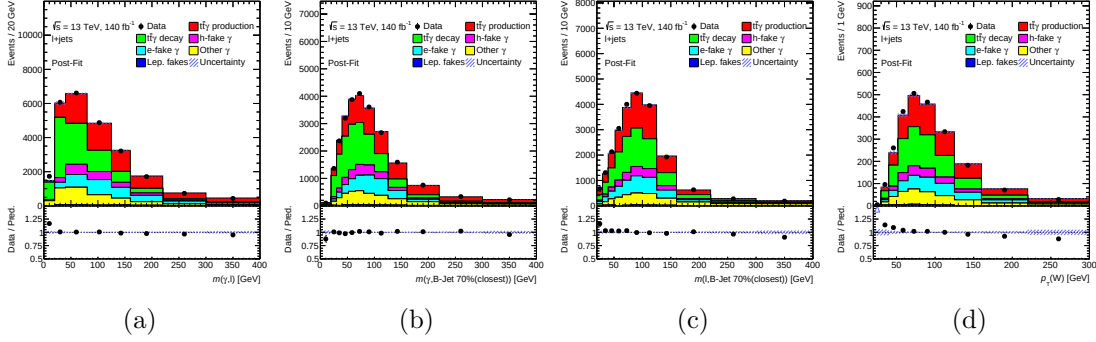


Figure E.59.: Post-fit plots for the invariant mass distributions of the photon and the lepton, of the photon and the closest b -jet and of the lepton and the closest b -jet, as well as the W -boson p_T distribution (from left to right). The uncertainty band corresponds to the total uncertainty, including stat. and syst. uncertainties.

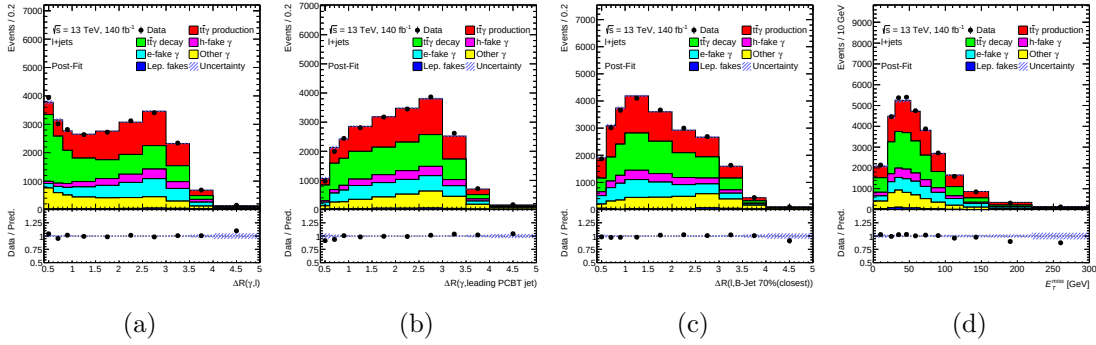


Figure E.60.: Post-fit plots for the ΔR distributions of the photon and the lepton, of the photon and the leading PCBT jet and of the lepton and the closest b -jet, as well as the E_T^{miss} distribution (from left to right). The uncertainty band corresponds to the total uncertainty, including stat. and syst. uncertainties.

E. Additional material on inclusive cross-section fits

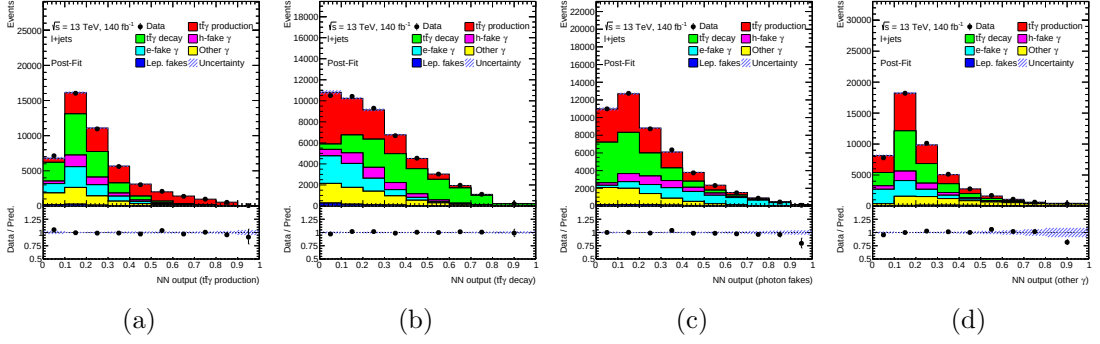


Figure E.61.: Post-fit plots for the four NN output distributions. The uncertainty band corresponds to the total uncertainty, including stat. and syst. uncertainties.

E.3.3. $t\bar{t}\gamma$ production: Post-fit plots for the fit in the dilepton channel with two NF

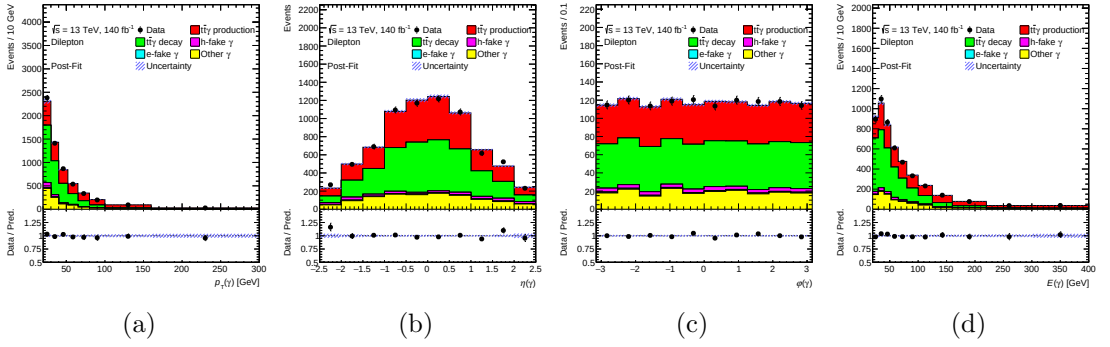


Figure E.62.: Post-fit plots for the p_T , η , ϕ and E distributions of the photon (from left to right). The uncertainty band corresponds to the total uncertainty, including stat. and syst. uncertainties.

E.3. Post-fit plots

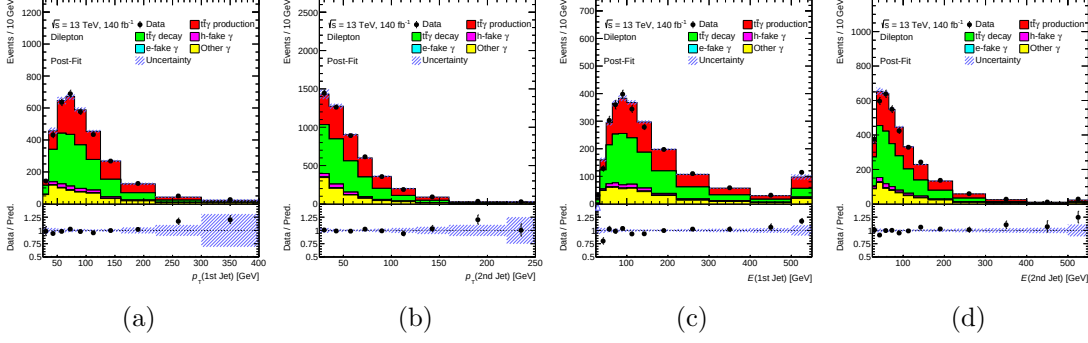


Figure E.63.: Post-fit plots for the p_T , E distributions of the first two jets (from left to right). The uncertainty band corresponds to the total uncertainty, including stat. and syst. uncertainties.

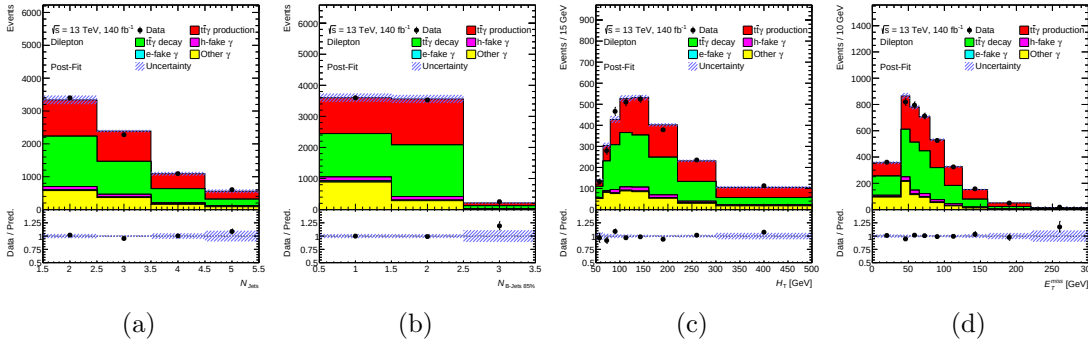


Figure E.64.: Post-fit plots for the distributions of the number of jets and b -jets per event and for the H_T and E_T^{miss} distributions (from left to right). The uncertainty band corresponds to the total uncertainty, including stat. and syst. uncertainties.

E. Additional material on inclusive cross-section fits

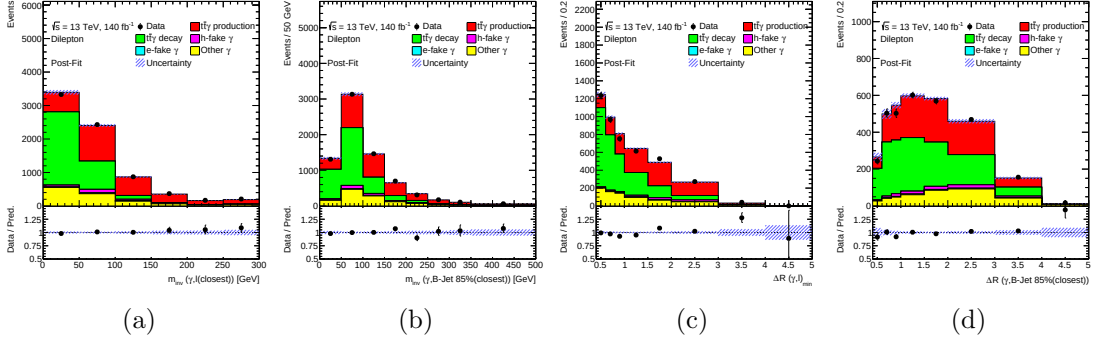


Figure E.65.: Post-fit plots for invariant mass distributions of the photon and the closest lepton and of the photon and the closest b -jet, as well as the ΔR distributions for the same two configurations (from left to right). The uncertainty band corresponds to the total uncertainty, including stat. and syst. uncertainties.

E.3.4. Inclusive $t\bar{t}\gamma$: Post-fit plots for the fit in the dilepton channel with one NF

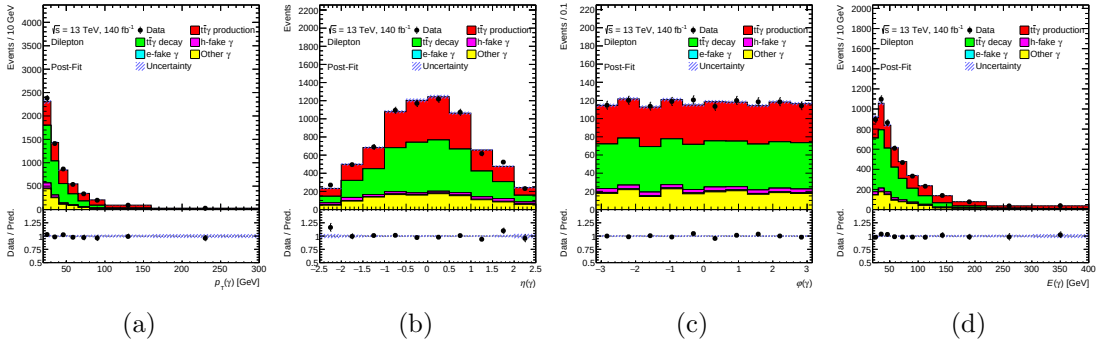


Figure E.66.: Post-fit plots for the p_T , η , ϕ and E distributions of the photon (from left to right). The uncertainty band corresponds to the total uncertainty, including stat. and syst. uncertainties.

E.3. Post-fit plots

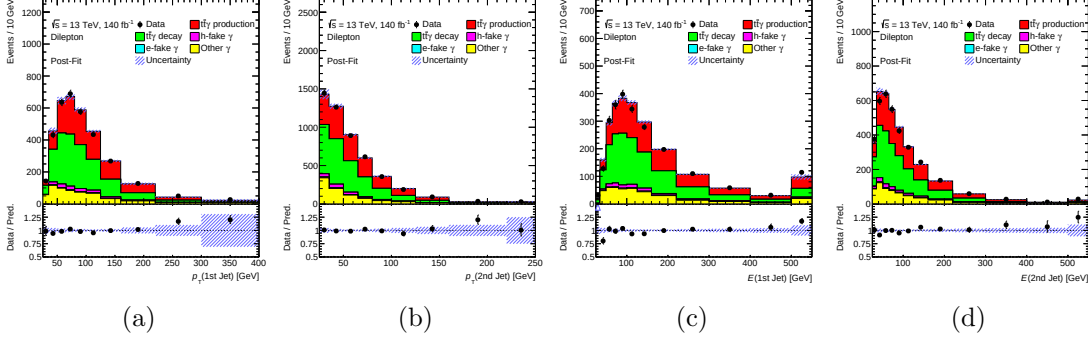


Figure E.67.: Post-fit plots for the p_T , E distributions of the first two jets (from left to right). The uncertainty band corresponds to the total uncertainty, including stat. and syst. uncertainties.

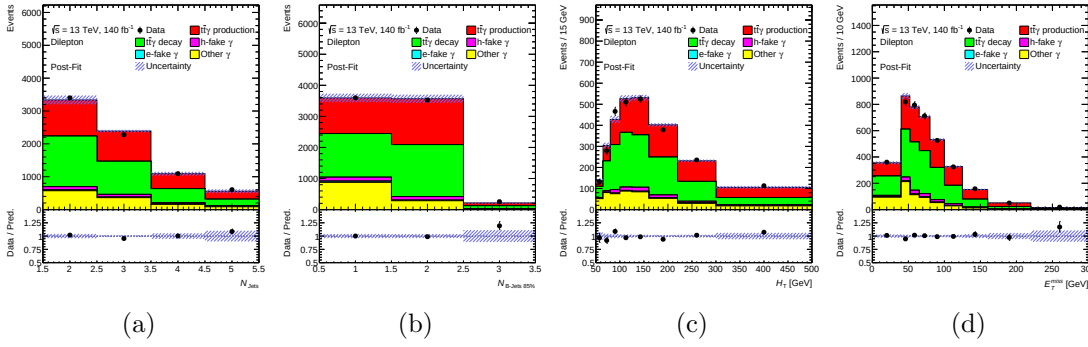


Figure E.68.: Post-fit plots for the distributions of the number of jets and b -jets per event and for the H_T and E_T^{miss} distributions (from left to right). The uncertainty band corresponds to the total uncertainty, including stat. and syst. uncertainties.

E. Additional material on inclusive cross-section fits

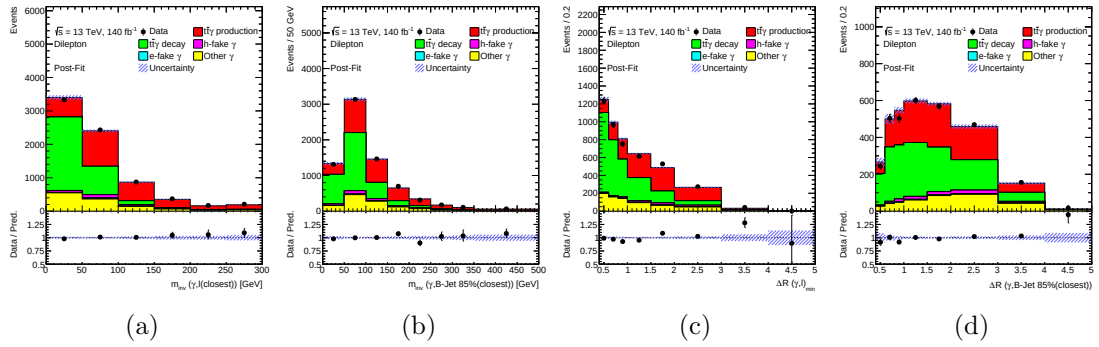


Figure E.69.: Post-fit plots for invariant mass distributions of the photon and the closest lepton and of the photon and the closest b -jets, as well as the ΔR distributions for the same two configurations (from left to right). The uncertainty band corresponds to the total uncertainty, including stat. and syst. uncertainties.

Results on differential cross-sections

F.1. Absolute cross-sections

F.1.1. Lepton+jets channel

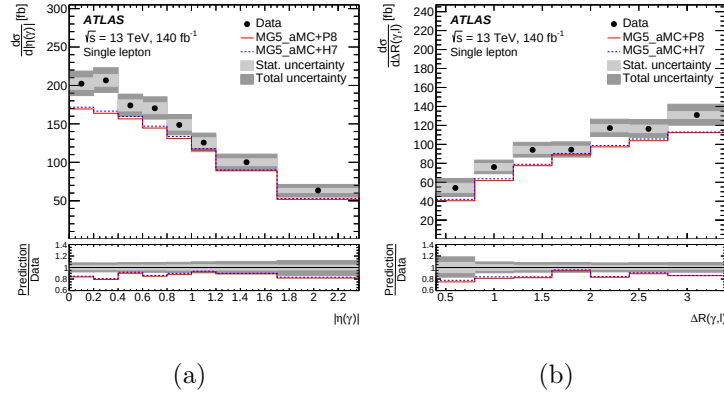


Figure F.1.: Particle-level unfolded distributions in the lepton+jets¹ channel for $t\bar{t}\gamma$ from production for $|\eta(\gamma)|$ (a) and $\Delta R(\gamma, l)$ (b). The error bars represent both statistical and systematic uncertainties. Overflow events are included in the last bin of each distribution. Distributions were taken from [3].

¹In this Appendix the term single lepton is used on the figures.

F. Results on differential cross-sections

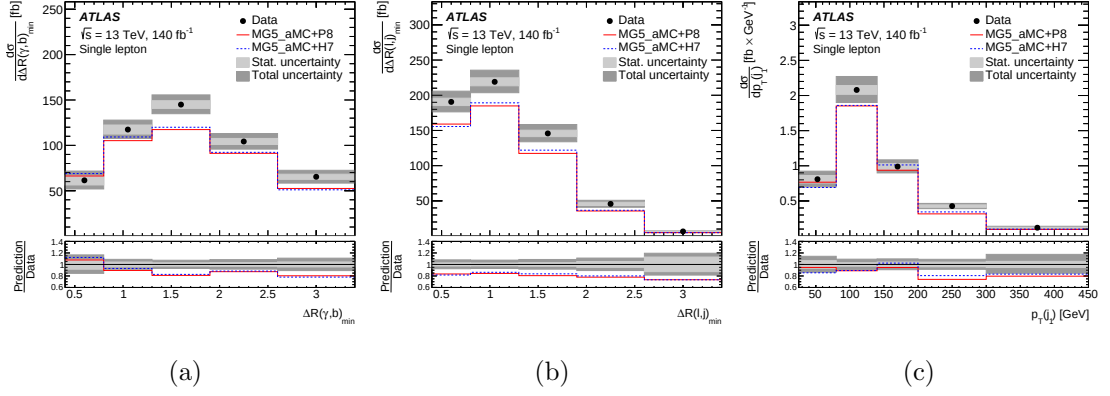


Figure F.2.: Particle-level unfolded distributions in the lepton+jets channel for $t\bar{t}\gamma$ from production for $\Delta R(\gamma, b)_{\min}$ (a), $\Delta R(\gamma, j)_{\min}$ (b) and $p_T(\text{1st jet})$ (c). The error bars represent both statistical and systematic uncertainties. Overflow events are included in the last bin of each distribution. Distributions were taken from [3].

F.1.2. Dilepton channel

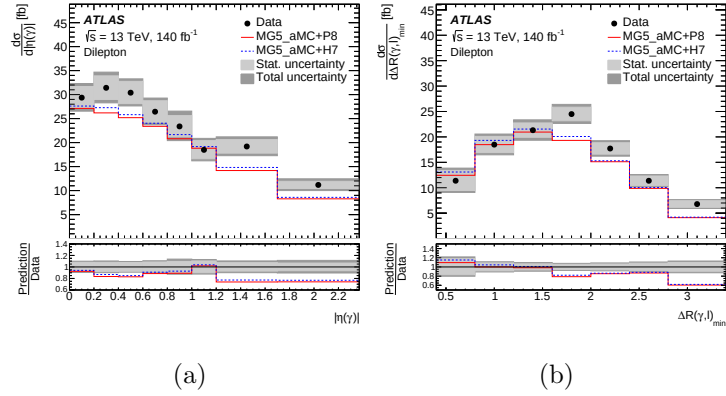


Figure F.3.: Particle-level unfolded distributions in the dilepton channel for $t\bar{t}\gamma$ from production for $|\eta(\gamma)|$ (a) and $\Delta R(\gamma, l)_{\min}$ (b). The error bars represent both statistical and systematic uncertainties. Overflow events are included in the last bin of each distribution. Distributions were taken from [3].

F.2. Normalized cross-sections

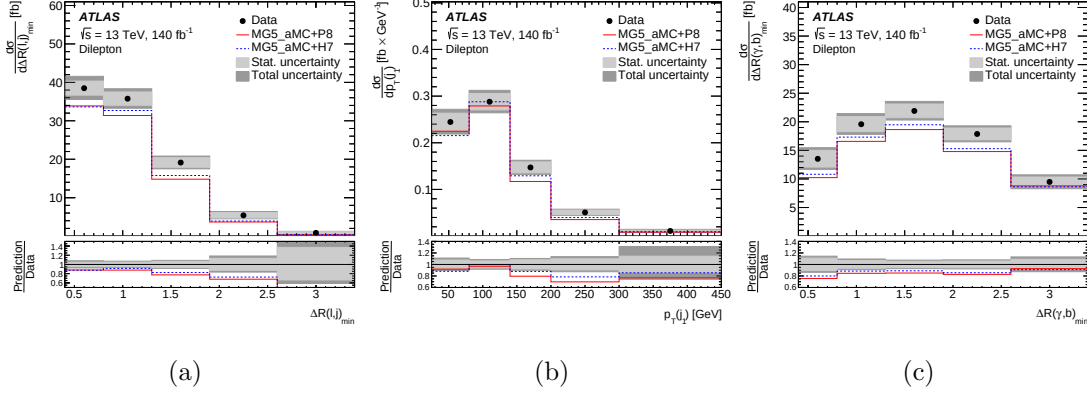


Figure F.4.: Particle-level unfolded distributions in the dilepton channel for $t\bar{t}\gamma$ from production for $\Delta R(l, j)_{\min}$ (a), $p_T(j)$ (b) and $\Delta R(\gamma, b)_{\min}$ (c). The error bars represent both statistical and systematic uncertainties. Overflow events are included in the last bin of each distribution. Distributions were taken from [3].

F.2. Normalized cross-sections

F.2.1. Lepton+jets channel

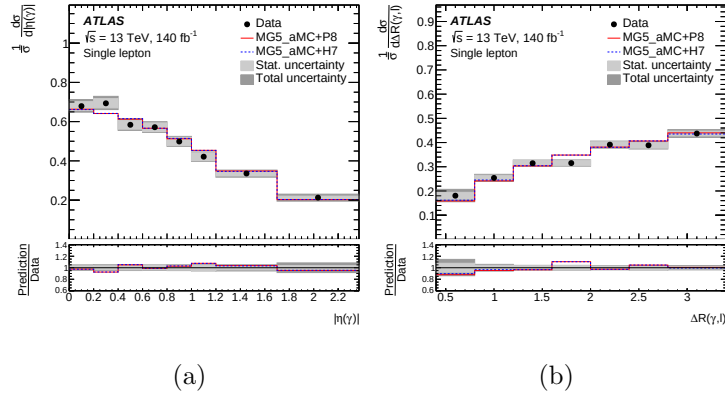


Figure F.5.: Particle-level unfolded distributions in the lepton+jets channel for $t\bar{t}\gamma$ from production for $|\eta(\gamma)|$ (a) and $\Delta R(\gamma, l)$ (b). The error bars represent both statistical and systematic uncertainties. Overflow events are included in the last bin of each distribution. Distributions were taken from [3].

F. Results on differential cross-sections

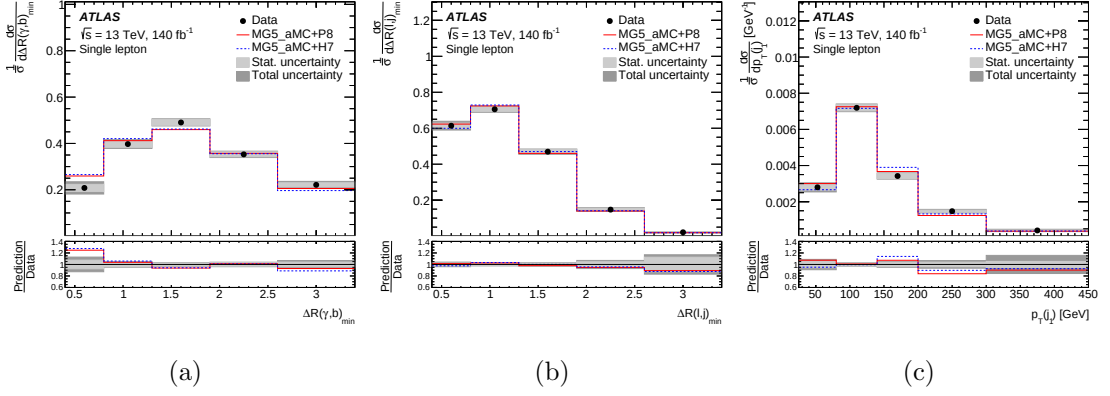


Figure F.6.: Particle-level unfolded distributions in the lepton+jets channel for $t\bar{t}\gamma$ from production for $\Delta R(\gamma, b)_{\min}$ (a), $\Delta R(\gamma, j)_{\min}$ (b) and $p_T(\text{1st jet})$ (c). The error bars represent both statistical and systematic uncertainties. Overflow events are included in the last bin of each distribution. Distributions were taken from [3].

F.2.2. Dilepton channel

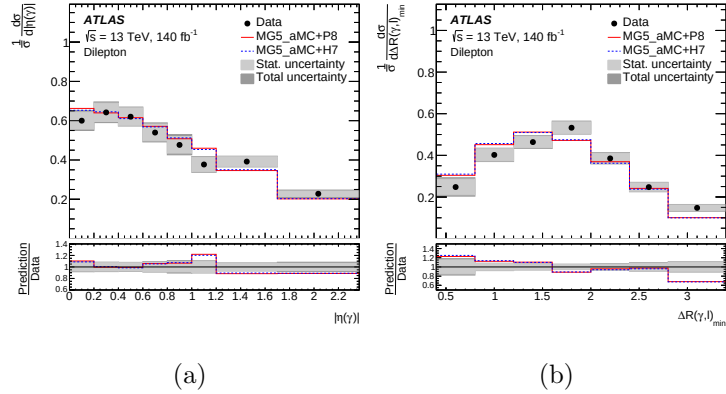


Figure F.7.: Particle-level unfolded distributions in the dilepton channel for $t\bar{t}\gamma$ from production for $|\eta(\gamma)|$ (a) and $\Delta R(\gamma, l)_{\min}$ (b). The error bars represent both statistical and systematic uncertainties. Overflow events are included in the last bin of each distribution. Distributions were taken from [3].

F.3. χ^2 and p -values

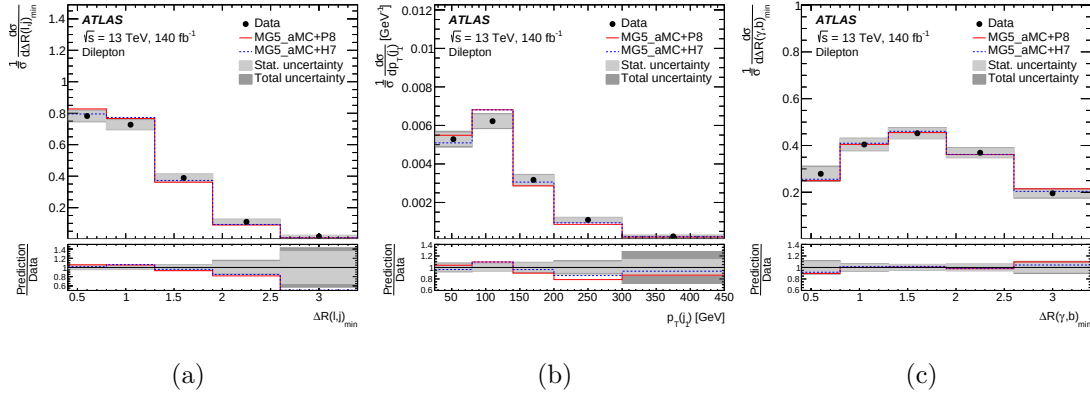


Figure F.8.: Particle-level unfolded distributions in the dilepton channel for $t\bar{t}\gamma$ from production for $\Delta R(l, j)_{min}$ (a), $p_T(j)$ (b) and $\Delta R(\gamma, b)_{min}$ (c). The error bars represent both statistical and systematic uncertainties. Overflow events are included in the last bin of each distribution. Distributions were taken from [3].

F.3. χ^2 and p -values

Table F.1.: χ^2/ndf and p -values for the measured absolute and normalised cross-sections for $t\bar{t}\gamma$ from production using the NLO MadGraph5_aMC@NLO simulations interfaced with Pythia 8 and Herwig 7. Table was taken from [3].

Variables	Absolute cross-sections				Normalised cross-sections			
	MG5_aMC@NLO+Pythia 8 χ^2/ndf	p -value	MG5_aMC@NLO+Herwig 7 χ^2/ndf	p -value	MG5_aMC@NLO+Pythia 8 χ^2/ndf	p -value	MG5_aMC@NLO+Herwig 7 χ^2/ndf	p -value
Single-lepton and dilepton combined								
$p_T(\gamma)$	10.7/10	0.38	9.3/10	0.50	11.6/9	0.23	8.6/9	0.47
$ \eta (\gamma)$	15.8/8	0.04	14.2/8	0.08	10.2/7	0.18	10.0/7	0.19
Single-lepton channel								
$p_T(\gamma)$	11.7/10	0.31	10.4/10	0.40	43.9/9	< 0.01	32.1/9	< 0.01
$ \eta (\gamma)$	11.8/8	0.16	11.1/8	0.20	8.1/7	0.33	8.1/7	0.32
$\Delta R(\gamma, l)$	10.5/7	0.16	9.9/7	0.19	8.8/6	0.19	8.8/6	0.19
$\Delta R(\gamma, b)_{min}$	12.5/5	0.03	12.3/5	0.03	7.6/4	0.11	9.0/4	0.06
$\Delta R(l, j)_{min}$	6.3/5	0.28	6.6/5	0.25	1.5/4	0.83	2.5/4	0.65
$p_T(j_1)$	12.6/5	0.03	10.8/5	0.06	8.2/4	0.08	9.7/4	0.05
Dilepton channel								
$p_T(\gamma)$	8.6/6	0.20	7.2/6	0.31	6.5/5	0.26	5.9/5	0.32
$ \eta (\gamma)$	12.1/8	0.15	9.9/8	0.27	9.2/7	0.24	7.9/7	0.34
$\Delta R(\gamma, l)_{min}$	17.2/7	0.02	16.5/7	0.02	14.2/6	0.03	14.4/6	0.03
$\Delta R(\gamma, b)_{min}$	7.8/5	0.17	5.0/5	0.42	1.4/4	0.84	0.8/4	0.93
$\Delta R(l, j)_{min}$	9.3/5	0.10	6.4/5	0.27	5.4/4	0.25	3.7/4	0.45
$p_T(j_1)$	10.5/5	0.06	5.0/5	0.42	7.8/4	0.10	3.6/4	0.46

APPENDIX G

EFT results and combination with $t\bar{t}Z$

G.1. EFT

Table G.1.: Confidence intervals at 68% and 95% credibility (expected and observed) for the c_{tW} and c_{tB} Wilson coefficients obtained from the $t\bar{t}\gamma$ measurement. Results are shown for the values obtained from marginalised distributions and from the individual fits. Best-fit values (global mode) are shown as well. Table was taken from [3].

Wilson coefficient		68% CI (exp.)	95% CI (exp.)	68% CI (obs.)	95% CI (obs.)	Best-fit
$\Re[C_{tW}]$	$\mathcal{O}(\Lambda^{-4})$ (marg.)	$[-0.35, 1.9]$	$[-1.0, 2.7]$	$[-0.55, 1.9]$	$[-1.2, 2.8]$	-0.21
	$\mathcal{O}(\Lambda^{-4})$ (indp.)	$[-0.52, 0.60]$	$[-0.90, 0.98]$	$[-0.44, 0.46]$	$[-0.78, 0.84]$	-0.01
$\Im[C_{tW}]$	$\mathcal{O}(\Lambda^{-4})$ (marg.)	$[-1.1, 1.1]$	$[-1.9, 2.0]$	$[-0.95, 0.95]$	$[-1.8, 1.8]$	0.01
	$\mathcal{O}(\Lambda^{-4})$ (indp.)	$[-0.58, 0.56]$	$[-0.96, 0.94]$	$[-0.48, 0.44]$	$[-0.84, 0.80]$	-0.01
$\Re[C_{tB}]$	$\mathcal{O}(\Lambda^{-4})$ (marg.)	$[-0.20, 1.0]$	$[-0.58, 1.5]$	$[-0.30, 1.0]$	$[-0.66, 1.5]$	-0.13
	$\mathcal{O}(\Lambda^{-4})$ (indp.)	$[-0.31, 0.31]$	$[-0.51, 0.52]$	$[-0.26, 0.23]$	$[-0.44, 0.44]$	-0.03
$\Im[C_{tB}]$	$\mathcal{O}(\Lambda^{-4})$ (marg.)	$[-0.58, 0.62]$	$[-1.1, 1.1]$	$[-0.50, 0.54]$	$[-0.96, 0.98]$	-0.01
	$\mathcal{O}(\Lambda^{-4})$ (indp.)	$[-0.32, 0.31]$	$[-0.53, 0.51]$	$[-0.26, 0.24]$	$[-0.45, 0.44]$	-0.01

G.2. Combination with $t\bar{t}Z$

Table G.2.: Confidence intervals at 68% and 95% credibility (expected and observed) for the c_{tW} and c_{tB} Wilson coefficients obtained from the combined $t\bar{t}\gamma$ and $t\bar{t}Z$ measurement. Results are shown for the values obtained from marginalised distributions and from the individual fits. Best-fit values (global mode) are shown as well. Table was taken from [3].

Wilson coefficient		68% CI (exp.)	95% CI (exp.)	68% CI (obs.)	95% CI (obs.)	Best-fit
$\Re[C_{tW}]$	$\mathcal{O}(\Lambda^{-4})$ (marg.)	$[-0.50, 1.2]$	$[-1.0, 2.2]$	$[-0.35, 1.8]$	$[-1.1, 2.5]$	1.61
	$\mathcal{O}(\Lambda^{-4})$ (indp.)	$[-0.44, 0.54]$	$[-0.78, 0.86]$	$[-0.28, 0.32]$	$[-0.54, 0.60]$	0.01
$\Im[C_{tW}]$	$\mathcal{O}(\Lambda^{-4})$ (marg.)	$[-1.1, 0.45]$	$[-1.7, 1.2]$	$[-1.3, 0.15]$	$[-1.8, 1.1]$	-0.89
	$\mathcal{O}(\Lambda^{-4})$ (indp.)	$[-0.48, 0.50]$	$[-0.82, 0.82]$	$[-0.32, 0.30]$	$[-0.58, 0.58]$	-0.01
$\Re[C_{tB}]$	$\mathcal{O}(\Lambda^{-4})$ (marg.)	$[-0.28, 0.70]$	$[-0.60, 1.2]$	$[-0.34, 1.2]$	$[-0.68, 1.4]$	0.94
	$\mathcal{O}(\Lambda^{-4})$ (indp.)	$[-0.28, 0.32]$	$[-0.49, 0.52]$	$[-0.24, 0.19]$	$[-0.41, 0.39]$	-0.03
$\Im[C_{tB}]$	$\mathcal{O}(\Lambda^{-4})$ (marg.)	$[-0.34, 0.62]$	$[-0.78, 1.0]$	$[-0.18, 0.72]$	$[-0.64, 1.0]$	0.46
	$\mathcal{O}(\Lambda^{-4})$ (indp.)	$[-0.31, 0.30]$	$[-0.51, 0.51]$	$[-0.23, 0.21]$	$[-0.40, 0.40]$	-0.01

Danksagung

Ich denke, es ist kein Geheimnis zu sagen, dass eine solche Arbeit, wie die hier vorliegende, nicht aus dem Nichts entsteht. Natürlich habe ich diese Arbeit selbst verfasst und die dem zugrunde liegende Forschungsarbeit selbst geleistet. Damit ich diese Forschungsarbeit aber so überhaupt leisten konnte, mussten auch die äußeren Bedingungen stimmen. Dafür möchte ich gerne einigen Menschen danken:

Danken möchte ich Arnulf für die nunmehr fast siebenjährige Begleitung, zuerst während der Bachelorarbeit, dann während der Masterarbeit und nun auch während der Promotion. Gerade die Themenähnlichkeit von Masterarbeit und Promotion sowie die kontinuierliche Betreuung hat vieles einfacher gemacht. Des weiteren gilt mein Dank den Kolleg*innen im II. Physikalischen Institut, die immer für eine gute Arbeitsatmosphäre gesorgt haben, vor allem die tägliche Kaffeepause war gerade an anstrengenden Tagen der Lichtblick, den es benötigte, um sich danach wieder mit voller Konzentration der Arbeit widmen zu können. Leider habe ich es nicht geschafft, mich auch in die Aktivitäten außerhalb der Arbeit so einzubringen, wie ich es gerne getan hätte.

Für die enge und gute Zusammenarbeit möchte ich an dieser Stelle vor allem Steffen und Lisa danken, die ich jederzeit mit Fragen behelligen durfte und die auch die Arbeit Korrektur gelesen haben. Fürs Korrekturlesen möchte ich auch Baptiste danken. Ebenfalls möchte ich Stephen für die gute Atmosphäre im Büro danken. Unsere Diskussionen haben mir immer viel Spaß gemacht.

Für die ebenfalls gute Zusammenarbeit innerhalb des Analyse-Teams möchte ich den ATLAS-Kolleg*innen aus Siegen, Dortmund und Göttingen danken. Dies sind Carmen (dir auch dafür, dass du als Co-Referentin für die Bewertung dieser Arbeit zur Verfügung standest), Amartya, Buddha, Arpan, Jan, Tomas und Baptiste.

Aber auch außerhalb der Arbeit gab und gibt es viele Menschen, die Teil der Atmosphäre waren, die es mir erlaubt hat, diese Arbeit zu verfassen. Zu nennen wären hier meine Freunde aus dem Studium, allen voran Jonas, Belal und Nils. Es ist schön, dass der Kontakt auch nach dem Studium noch so eng geblieben ist. Ebenso möchte ich den Menschen der Grünen Jugend und von Ende Gelände Göttingen danken, die meinen politischen Werdegang begleitet und geprägt haben. Ohne euch wäre ich heute sicherlich nicht der Mensch, der ich bin.

Genauso wichtig sind aber auch die Menschen, die mit mir über die Jahre hinweg Basket-

Danksagung

ball gespielt haben. Das Training war auch immer eine wichtige Quelle der Regeneration. Vor allem Johann möchte ich für die entstandene Freundschaft danken.

Ein weiterer Dank geht an alle Menschen, die mich irgendwie durch mein Leben bisher begleitet haben, wie Lehrer*innen, Mitschüler*innen und sonstige Freund*innen. Wie es nun mal so ist, besteht nicht mehr zu allen Kontakt. Umso schöner und wichtiger sind die Kontakte, die geblieben sind. Daher auch ein ganz besonderer Dank an Lukas, Matthias, Tobias, Fabian und Tim.

Nun kommt der wichtigste Teil dieser Danksagung. Der gehört meiner Familie und damit sind hier meine schon verstorbenen Großeltern, meine Tanten, Onkels, Cous*inen und alle anderen Menschen gemeint, die sich diesem verrückten Haufen zugehörig fühlen. Der Zusammenhalt ist beeindruckend und ich freue mich jedes Mal euch zu sehen, sei es bei Festen, an Weihnachten oder sonstigen Anlässen. Vor allem meiner Oma Paula hätte ich es gegönnt, den Abschluss dieser Arbeit noch erleben zu dürfen. Leider ist sie dafür im Dezember leider etwas zu früh von uns gegangen.

Der größte Dank geht natürlich an meinen Vater. Mit Worten kann ich nicht beschreiben, wie dankbar ich dir für alles bin! Egal was war, du hast immer hinter mir gestanden. Ein ebenso wichtiger Dank geht an meinen verstorbenen Bruder und meine verstorbenen Mutter. Ich vermisse euch.

Zwei Personen, die mein Leben auf unbeschreibliche Weise bereichern, habe ich bisher noch nicht erwähnt. Meine Liebe Aless, ich kann dir gar nicht genug danken für die letzten 5 Jahre. Du hast mir an so vielen Stellen den Rücken freigehalten, sodass ich mich auf die Arbeit konzentrieren konnte und hast immer wieder zurückgesteckt. Ich hoffe dir dies eines Tages ebenfalls zurückgeben zu können. Danke auch, dass du mich so nimmst, wie ich bin, mit allen meinen Ecken, Kanten und Marotten. Und dir Camilo gilt ein ganz großer Dank. Es ist etwas ganz Besonderes dich zu haben und zu sehen, wie du langsam immer größer und selbstständiger wirst. Lass dich nicht vom Leben unterkriegen!

Curriculum Vitae

Personal information

Name:	Andreas Achim Kirchhoff
Date of birth:	13th March 1993
Place of birth:	Wipperfürth
Nationality:	German
Address:	Am Kirschberge 61
Postal code, city:	37085 Göttingen
Civil status:	unmarried
Children:	Camilo Alexander Kirchhoff Padilla (born on 7th March 2013). Since July 2018 he lives in Germany. Before he lived in Nicaragua.

Education

August '97 - July '99	Catholic kindergarten St. Michael in Oldenburg-Kreyenbrück
August '99 - July '02	Catholic primary school Kreyenbrück
August '02 - June '11	Liebfrauenschule Oldenburg (catholic grammar school) Qualification: Abitur (university-entrance diploma; grade: 1.0)
October '09 -September '10	Attending courses in physics at the Carl von Ossietzky University Oldenburg Experimental physics I and II approved with grade 2.0

University studies

October '12 - August '15	Bachelor studies in physics at the Georg-August-University Göttingen Bachelor thesis: „Implementation of a 2-dimensional fit to measure the top-quark width“
January '13 -October '18	Bachelor and Master studies supported by Studienstiftung des deutschen Volkes
September '15 - January '16	Erasmus semester in Madrid at the Universidad Autónoma de Madrid (UAM)

March '16	Receiving bachelor degree (grade: 1.8)
October '16 - October '18	Master studies in physics (main focus: elementary particle physics) at the Georg-August-University Göttingen (grade: 1.2) Master thesis: „Studies to classify prompt photons in the $t\bar{t}\gamma$ -process.“
November '18- February '24	PhD studies at the II. Physical Institute at the University of Göttingen Topic: „Measurement of the $t\bar{t}\gamma$ production cross section with the ATLAS experiment at $\sqrt{s} = 13$ TeV using the full run 2 data set“
September '19 -November '23	Doctoral studies supported by the Studienstiftung des deutschen Volkes
October '19 - September '23	Member of the system administration team at the II. Physical Institute at the University of Göttingen

University teaching

Winter semester '17/'18	Tutor for the introductory lecture into particle and nuclear physics
Summer semester '19	Tutor for the lecture on quantum mechanics 1
Winter semester '19/'20	Tutor for the introductory lecture into particle and nuclear physics
Summer semester '20	Tutor for the lecture on experimental physics IV on atomic physics
Summer semester '21	Tutor for the lecture on experimental physics IV on atomic physics
Winter semester '23/'24	Tutor for the introductory lab course

Miscellaneous

July '11 - June '12	Voluntary service with „weltwärts“in Nicaragua (program of BMZ; sending organisation Nicaragua-Verein Oldenburg e.V.)
December '12 - December '17	Member of the extended steering committee of the Nicaragua-Verein Oldenburg e.V.

March '16 - September '16	Long-term stay with my son in San Francisco Libre (NIC)
December '17 - now	Member of the steering committee of the Nicaragua-Verein Oldenburg e.V.; main duty: finances
July '18 - January '19	Taking care of my son as a single father

Languages

German (mother language)
English (fluent)
Spanish (fluent)
French (basic knowledge)

IT/programming

C++, Python
Root, Keras (neuronal networks)
Latex, Gnuplot, MS Office
Linux