Enlightening a tiny difference: Measurement of top quark pair charge asymmetry in $\mathbf{t\bar{t}} + \gamma$ production at LHC using the ATLAS detector at $\sqrt{s} = 13\,\text{TeV}$

DISSERTATION
to obtain the degree of
Doctor of Science
(Dr. rer. nat.)

submitted by
Amartya Rej
from Katwa, India

submitted to the
School of Science and Technology
of
University of Siegen
Siegen 2023

Supervisor and first appraiser : Prof. Dr. Ivor Fleck,
University of Siegen

Second appraiser : Prof. Dr. Kevin Kröninger,
TU Dortmund

Additional examiners: Prof. Markus Risse,
University of Siegen

Prof. Wolfgang Kilian
University of Siegen

Date of the oral examination: 17th May 2023

Year of Publication: 2023

Abstract

A first measurement of the top quark pair charge asymmetry (A_C) is presented in top quark pair ($t\bar{t}$) production in association with a photon ($t\bar{t}\gamma$ production), using the data of ATLAS experiment with an integrated luminosity of 139 fb^{-1}collected at proton-proton collisions of $\sqrt{s} = 13$ TeV at the Large Hadron Collider at CERN. The measurement is performed using the semileptonic $t\bar{t}$ decay channel in fiducial phase space at particle level using a profile likelihood unfolding method on the distribution of the difference between the absolute rapidity of the two top quarks. The A_C is measured to be -0.003 ± 0.029, in agreement with the Standard Model prediction. The measurement is expected to be more sensitive to potential new physics than the same measurement in $t\bar{t}$ production. However, currently, it is limited by the available statistics and serves as a stepping stone for future analyses with more data.

Zusammenfassung

Es wird die erste Messung der Ladungsasymmetrie (A_C) des Top-Quark Paares bei der Top-Quark Paarproduktion ($t\bar{t}$) in Verbindung mit einem Photon ($t\bar{t}\gamma$-Produktion) vorgestellt. Dabei werden die Daten des ATLAS Experiments mit einer integrierten Luminosität von 139 fb^{-1}verwendet, die bei Proton-Proton Kollisionen bei $\sqrt{s} = 13$ TeV am Large Hadron Collider am CERN gesammelt wurden. Die Messung erfolgt unter Verwendung des semileptonischen $t\bar{t}$ Zerfallskanals in einem definierten Phasenraum auf Teilchenebene unter Verwendung der Profil Likelihood Entfaltungsmethode auf der Verteilung der Differenz zwischen den Beträgen der Rapiditäten der beiden Top-Quarks. Der gemessene A_C-Wert beträgt -0,003 ± 0,029, was mit der Vorhersage des Standardmodells übereinstimmt. Es wird erwartet, dass diese Messung empfindlicher für potenzielle neue Physik ist als die gleiche Messung bei der $t\bar{t}$ Produktion. Derzeit ist sie jedoch durch die verfügbare Statistik limitiert und dient als ausgangspunkt für künftige Analysen mit mehr Daten.

Acknowledgements

This thesis would not be possible without the help and support of my colleagues and collaborators and friends and family, and all the people I have interacted with throughout my PhD journey, and all the people who inspired me to consider this path.

I would like to thank my supervisor Prof. Dr. Ivor Fleck, for letting me join his working group and providing continuous support and guidance throughout, helping me build a confidence and aptitude for science research.

I want to thank Prof. Dr. Kevin Kröninger for reviewing the thesis as the reviewer of the PhD committee, and also Prof. Dr. Wolfgang Kilian and Prof. Dr. Markus Risse for being the member of the PhD committee.

I am thankful to PD Dr. Carmen Diez Pardos, for the daily supervision and leading the physics analysis of this thesis. I could learn a lot from our long discussions and her experience in this topic. I am also thankful to Dr. Yichen Li, for helping me start my PhD journey and inspiring to work on the topic of the thesis.

I would like to thank all the fellow PhD students and colleagues at Siegen for creating a friendly atmosphere and supporting me with many help and endure numerous discussions.

I would like to thank my colleagues at the ATLAS collaboration for providing support throughout, especially the technical support for the physics analysis.

I want to thank all my friends for keeping patience with me. I want to admit that this thesis could be possible largely because of my parents, for their constant support. I had the freedom to make my own choice from an young age, that led me becoming a free thinker and gave me a curious mind to pursue a research career. I want to thank all my teachers from school and universities, for believing on my abilities and inspiring me for a research career.

Finally, I would like to thank again Prof. Dr. Ivor Fleck and PD Dr. Carmen Diez Pardos for attentively reviewing this thesis multiple times and providing many helpful comments.

Contents

CHAPTER 1

Introduction

What is everything made of? - the question that has always intrigued humankind. The idea of the smallest element brought the concept of the atom and eventually the elementary particles. The theoretical model that can describe elementary particles most accurately today is known as the Standard Model (SM) of particle physics.

Although the SM is one of the most precise scientific theories built, confirming the experimental data, it is still theoretically unsatisfactory as it cannot paint a complete picture. For example, it does not include gravity as a force, does not account for the estimated dark matter in the universe, cannot explain the matter-antimatter asymmetry in the universe and more. Also, it cannot explain the meaning of the quantum numbers (like electric charge, weak isospin, hypercharge and colour). It contains 18 free parameters that are needed to be determined by experiments.

The possible way forward would be on one hand extending the SM to some other models that could explain those questions and directly verify their existence by experiments; on the other hand, making precise measurements that can test the SM and its parameters and constrain the other models.

The *forward-backward asymmetry* is an observable (calculated from the angular distributions of the final states of a process) that can be precisely measured at particle collider experiments. It had been measured for several important processes at electron-positron colliders like LEP or hadron colliders like Tevatron. From these measurements, the precision of several SM parameters could be improved and some beyond SM theories could be constrained.

At the Tevatron, the CDF and D0 experiments measured the forward-backward asymmetry of the top quarks produced in top quark pair ($\mathrm{t\bar{t}}$) production from the proton anti-proton ($\mathrm{p\bar{p}}$) collisions. The measured asymmetry was found to be in tension with the SM theoretical prediction at that point in time (2011) and for several years afterwards. This brought attention and motivation for the theory community to work for more refined SM prediction. Eventually, predictions with higher orders reduced the disagreement from the CDF and D0 combined experimental result to be within 1.3 standard deviations of the measurement, although the experimentally measured values in all of the measurements considered for the combination [1] are consistently higher.

The same asymmetric effect in $t\bar{t}$ production is further explored at the LHC. However, because of the symmetric nature of the proton-proton collisions (pp) at LHC, the forward-backward asymmetry cannot be measured. A different observable called *charge asymmetry* is measured . The charge asymmetry has been measured by the ATLAS and CMS experiments with data collected at several centre of mass energies (7, 8 and 13 TeV) and found to be consistent with the SM. The latest measurement by the ATLAS experiment could also reduce the uncertainty of the measurement enough so that the value of charge asymmetry is found to be 4.7σ away from the zero value, i.e. not symmetric - evidence (being above 4σ) of the charge asymmetry is found [2].

However, the charge asymmetry in $t\bar{t}$ production comes only from the quark-initiated production and that is only 10% of these events at the LHC. The sensitivity of the measurement to any potential new physics depends on this fraction. Hence, to obtain a better sensitivity of the asymmetry, processes or phase spaces with a higher fraction of quark-initiated production are investigated.

The production of top quark pair ($t\bar{t}$) in association with a photon ($t\bar{t}\gamma$ production) is one such process where the quark-initiated production fraction is much higher ($\sim$26%). Also, the principal asymmetric effect is expected to be present at the leading order prediction, unlike the $t\bar{t}$ production where only higher order corrections introduce the asymmetry. Hence, a measurement of the top quark pair charge asymmetry is performed for the $t\bar{t}\gamma$ production. The measurement is using the data from the ATLAS experiment at the LHC at the centre of mass energy of 13 TeV [3].

This is not just the first such measurement, but also some features are new as well in the analysis. This is the first measurement of the $t\bar{t}\gamma$ process where the top quarks are reconstructed and used as an observable. The analysis is performed only with the $t\bar{t}\gamma$ production process as the signal and is separated from the radiative top decays ($t\bar{t}\gamma$ decay process, explained in Chapter 4.2) which has quite similar signatures. This approach makes the asymmetry measurement more sensitive to the underlying physics and also a potential cross-section measurement more sensitive to the top-photon coupling.

In the following, the theory and the experiment are described in Chapters 2 and 3 respectively. The Monte Carlo based samples and their predictions are described in Chapter 4. The definitions of the reconstructed physics objects are given in Chapter 5. The event selection criteria are described in Chapter 6. All the uncertainties considered for the analysis are described in Chapter 7. The reconstruction of the top quark pair is explained in Chapter 8. The discrimination of the signal from the background as well as all the physics observables are described in Chapter 10. Finally, the analysis and the results can be found in Chapter 11 and concluded in Chapter 12.

CHAPTER 2

The theory of charge asymmetry

In this chapter, the theory and motivation of the measurement of the top quark pair charge asymmetry in the $t\bar{t}\gamma$ production process are described. Firstly, a brief introduction to the Standard Model of particle physics is described in Section 2.1. The concept of asymmetry is introduced in Section 2.3. The asymmetry for $t\bar{t}$ production is discussed in Section 2.4. Finally, the theory behind the charge asymmetry in the $t\bar{t}\gamma$ production process is described in Section 2.5.

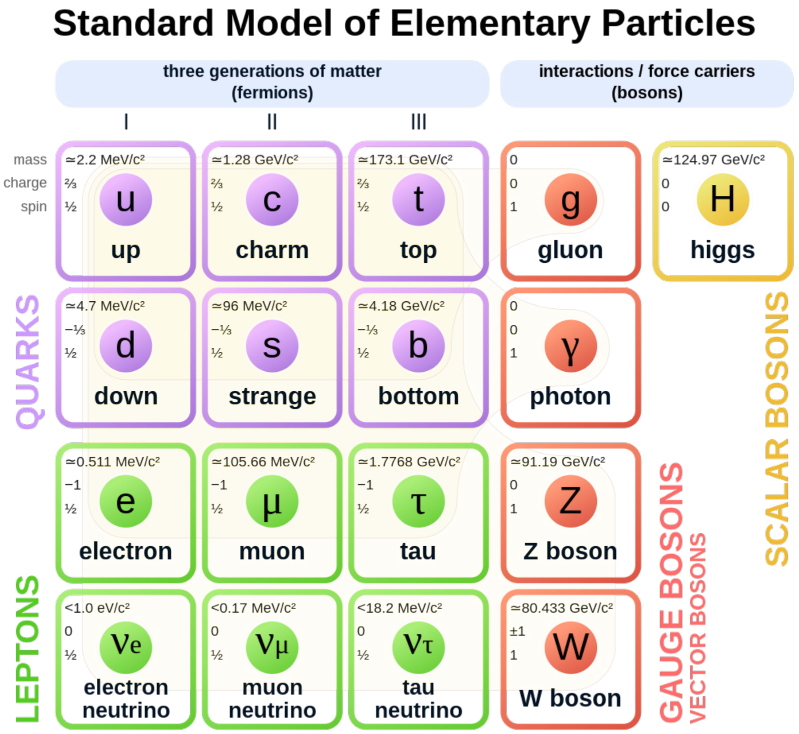
2.1 The Standard Model

The Standard Model (SM) of particle physics [6] is a theoretical model that describes the interactions between elementary particles. It consists of two types of particles called *fermions* and *bosons* with different quantum numbers. The fermions are of two further categories - leptons and quarks all with spin = $\frac{1}{2}$. All fermions have their anti-particles that have the same properties but opposite electric charge (or other additive quantum numbers). Both leptons and quarks have three *generations* of particles, where higher generations are more massive.

For quarks, each generation has an up-type quark with charge = $+\frac{2}{3}e$ (e is the absolute value of the charge of the electron) and a down-type quark with charge = $-\frac{1}{3}e$. They are called up (u) and down(d) quarks from the first generation, charm (c) and strange (s) quarks from the second generation, and top (t) and bottom (b) quarks from the third generation.

For leptons, each generation consists of an electrically charged lepton with charge -e and a corresponding electrically neutral particle called *neutrino*: electron (e) and electron neutrino (ν_e) from the first generation, muon (μ) and ν_μ from the second generation, and tau (τ) and ν_τ from the third generation.

The interactions between fermions take place via *gauge bosons* of spin 1: photon



(a)

Figure 2.1: The Standard Model of the elementary particles [4]. The mass values in the figure are not updated, the latest values can be found in [5].

(γ) for the *electromagnetic* force, $W^{\pm}$ and Z boson for the *weak* force and gluons (g) for the *strong* force. All electrically charged particles interact via photons, all fermions interact via weak bosons and only quarks interact via gluons. The particles get their mass through the spontaneous symmetry breaking of the electroweak gauge symmetry [7–10] via the Higgs mechanism with the existence of a scalar particle called Higgs boson of spin 0.

Among all of these elementary particles, the top quark has the highest mass. The neutrinos are massless in the SM. Although from experimental evidence, neutrinos should have non-zero mass. However, it can be negligible in the context of the subject of this thesis. Among the bosons, the W and Z bosons are massive, while the photon and gluons are massless. The Higgs boson also gets its mass via self-interaction.

A representation of all known elementary particles is shown in Figure 2.1. The mass, electric charge and spin quantum number values are also shown in the figure. The mass values in the figure are not updated with the latest values that can be found in [5].

2.2 Physics of the top quark

The top quark is the heaviest particle in the SM with a mass of 172.69 ± 0.30 GeV [5]. Because of its heavy mass, it may couple preferentially with particles predicted in the physics models beyond the Standard Model (BSM). Also, it has a very short decay time ($\sim 10^{-25}$ second) and unlike other quarks, it decays before its hadronisation (explained in Chapter 4.1.1). Hence, it is the only quark that allows the studies of a bare quark.

Top quarks were first discovered in 1995 from proton-antiproton collisions in the Tevatron [11] collider. They are produced in abundance in the Large Hadron Collider (LHC) [12] from proton-proton collisions and so it is called a *top quark factory*. At these hadron colliders, the top quark production mainly happens as the pair production of the top quark with its anti-particle anti-top quark, i.e. $\mathrm{t\bar{t}}$ production. This process is primarily produced via gluons, mainly from gluon fusion (at LHC) or quark pair annihilation (at Tevatron) as their Feynman diagrams are shown in Figure 2.2 respectively. The fraction of the production mode changes with the energy at the collision. The EW production $q\bar{q} \to \gamma^*/Z \to t\bar{t}$ is negligible at the hadron colliders.

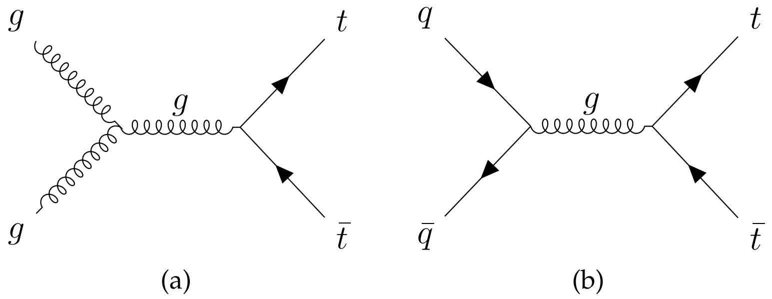
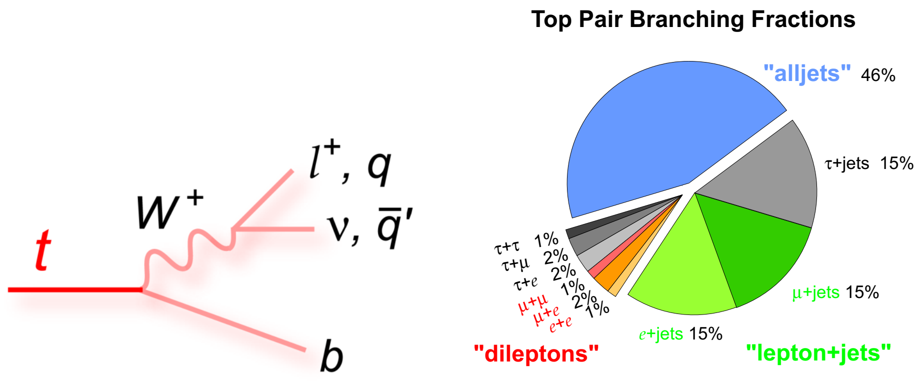


Figure 2.2: Example Feynman diagrams of the $\mathrm{t\bar{t}}$ production at the LO with (a) gluon-initiated production (b) quark-initiated production.

Top quarks decay through weak interaction almost always to a W boson and a bottom quark according to the SM. The W boson can decay either to a charged lepton and neutrino, called *leptonic* decay; or to a pair of quarks (light quarks u/d/c/s are most favourable), called *hadronic* decay. This leads to several decay channels for $\mathrm{t\bar{t}}$ production. The branching fraction of these channels depends on the decay modes of both the W bosons coming from the top and anti-top quark. The quarks as the decay products get hadronised and form *jets*. The branching fractions in terms of the combination of the lepton flavour (e, μ or τ) and jets are shown in Figure 2.3(b).



(a) (b)

Figure 2.3: (a) The decay of a top quark with different possibilities of decay products (b) The branching fraction of the decay channels of the $t\bar{t}$ production process [13].

2.3 Asymmetry in particle colliders

In the context of this thesis, *asymmetry* means the asymmetry in angular distributions of the final states in hadron colliders. However, the asymmetry can be defined starting with an electron-positron collider experiment where for example the production of muon pairs ($\mu^+\mu^-$) can be studied with great precision while the corresponding theoretical calculations are simple.

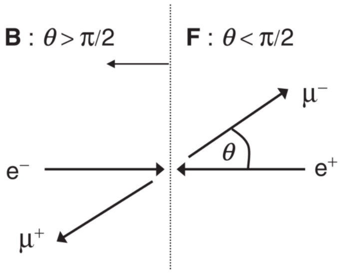


Figure 2.4: The forward (F) and backward (B) regions defined in the context of the production of muon (μ) pairs in electron-positron collisions (Taken from [14] Figure 16.7)

A schematic description of such a collision can be found in Figure 2.4, where a collision of electron (e^-) and positron (e^+) as initial states are shown in the horizontal axis. The final states μ^+ and μ^- are emitted opposite to each other, with μ^- emitted with an angle θ with respect to the direction of incident electron (e^-). The forward and backward regions are marked as **F** and **B**, using the ranges $0 < \theta < \frac{\pi}{2}$ and $\frac{\pi}{2} < \theta < \pi$ respectively.

For the production of $\mu^+\mu^-$ via a Z boson (i.e. $e^+e^- \rightarrow Z \rightarrow \mu^+\mu^-$) the differential cross-section (with respect to the solid angle Ω of μ^-) at the leading order (LO) approximation of quantum electrodynamics (QED) can be expressed in terms of the angle θ as (From [14] Equation 16.25):

$$\frac{d\sigma}{d\Omega} \propto a(1+\cos^2\theta) + 2b\cos\theta \tag{2.1}$$

where the coefficients a and b are related to the couplings of the Z boson to *left-handed* and *right-handed* leptons. The handedness is defined in terms of the spin and direction of motion of the particle [15].

The above expression (Equation 2.1) can be extended to the pair production of any fermion pair instead of the muon pair production. It can be generalised that the coefficient b would be zero if the couplings of the Z boson with left- and right-handed fermions are equal.

For a better understanding of its physical relevance, the cross-sections in the forward and backward regions can be calculated by integrating the differential equation with their respective ranges:

$$\sigma_{\mathrm{F}} \propto \int_0^{\frac{\pi}{2}} \frac{d\sigma}{d\Omega}\mathrm{d}\theta = \frac{4}{3}a + b \tag{2.2}$$

$$\sigma_{\mathrm{B}} \propto \int_{\frac{\pi}{2}}^{\pi} \frac{d\sigma}{d\Omega}\mathrm{d}\theta = \frac{4}{3}a - b \tag{2.3}$$

This difference in the forward and backward region is calculated in terms of the **forward-backward asymmetry**,

$$A_{\mathrm{FB}} = \frac{\sigma_{\mathrm{F}} - \sigma_{\mathrm{B}}}{\sigma_{\mathrm{F}} + \sigma_{\mathrm{B}}} = \frac{3b}{4a} = \frac{\mathrm{N_F} - \mathrm{N_B}}{\mathrm{N_F} + \mathrm{N_B}} \tag{2.4}$$

From Equation 2.4, it can be understood that A_{FB} is proportional to the coefficient b. Therefore, in the case the couplings of the Z boson to the left- and right-handed fermions are equal, b as well as the asymmetry A_{FB} will be zero. This will produce a symmetric production of the fermions in the forward and backward regions. However, the differences in the couplings will lead to a non-zero asymmetry. If b has a positive value, the negatively charged fermion will be produced more in the forward region than the backward region, and preferentially more in the direction of the incident electron (e^-) because of the $\cos\theta$ dependence of the differential cross-section.

2.3.1 Asymmetry at hadron colliders

The forward-backward asymmetry can also be defined similarly for the annihilation of quarks and anti-quarks of the same flavour) at hadron colliders producing fermions via the exchange of either a gluon, photon or Z boson. The exchange via a Z boson is the Drell-Yan process and this is well-studied for charged lepton pair final states at different particle collider experiments. However, the asymmetry for the production of quark pairs produced (mainly via gluons) is difficult to study as the quarks hadronise to form jets and loss the information of the charges of their parent particles, which are needed to be distinguished. There had been efforts in the past [16–18] to study the asymmetry for bottom quark pair production, by measuring the charge of the jets. However, the easier and best available option is the production of tt̄ with at least one of the top quark decaying with a charged lepton whose charge corresponds to the charge of the top quark as shown in Figure 2.3(a).

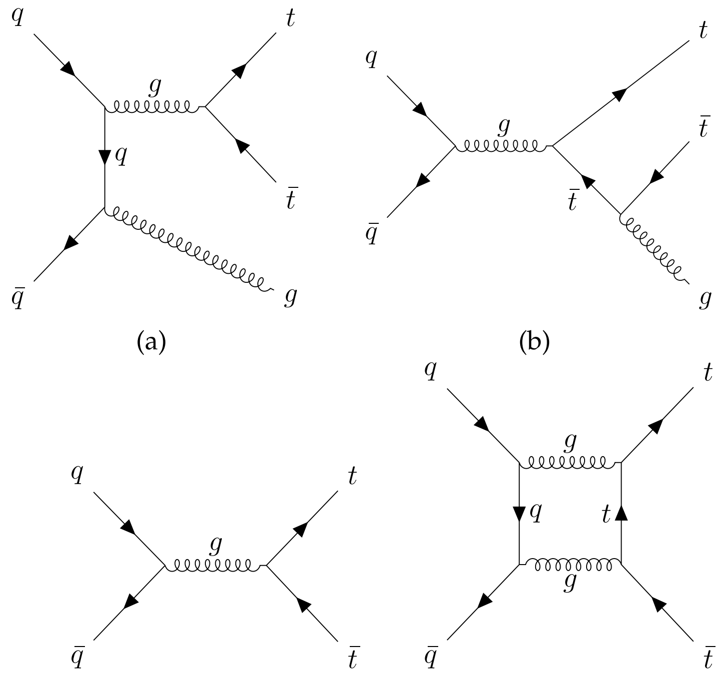
2.3.2 Sensitivity to new physics phenomena

The asymmetry as a precisely measurable observable serves the purpose of testing the validity of the SM and also BSM, as well as precisely determining fundamental parameters e.g. the weak mixing angle θ_W, being related to the asymmetry in $e^+e^- \to Z \to \mu^+\mu^-$. The same is also true for the tt̄ production process, by studying which state-of-the-art theory predictions can be precisely tested in the experiments and any deviations may indicate a possible existence of new physics.

The measurement of asymmetry in tt̄ production can be sensitive to the existence of the hypothetical models with *axigluon* particles [19]. The exchange via an axigluon instead of a gluon can give different couplings to the left- and right-handed quarks (as well as between light quarks and top quarks) resulting in a different asymmetry value than expected from the SM. A small fraction of events with axigluon exchange along with the SM production may change the measured asymmetry, depending on the particular model's parameters. Also, interference of such models with the SM can modify the asymmetry. There are several other BSM examples [1] that are also relevant for the asymmetry measurements.

2.4 Asymmetry in tt̄ production

The dominant tt̄ production at the hadron colliders via gluon exchange does not have the forward-backward asymmetry, it results in a symmetric angular distribution of the top quarks. However, this prediction assumes a simplified calculation considering a LO approximation of QCD (quantum chromodynamics). Once the higher-order corrections are taken into account, some asymmetric contributions lead to an overall non-zero asymmetry. For example, at the NLO QCD approximation, the interference of initial- and final-state radiation, Figure 2.5 (a) and (b), as well as the interference of the box diagrams with the born diagrams, Figure 2.5 (c) and (d), contribute to negative and positive values of asymmetry respectively.



(a) (b)

(c) (d)

Figure 2.5: Example Feynman diagrams of $t\bar{t}$ production that contributes to the charge asymmetry

The forward-backward asymmetry is measured with the CDF and D0 experiments at the Tevatron collider in $p\bar{p}$ collisions predominantly with quark-initiated production and the direction of initial quarks being largely along the direction of initial protons and anti-quarks for anti-protons. The asymmetry for the top and anti-top quarks is defined in terms of the rapidity, y (defined in Chapter 3.2.1) distribution of the top quarks of them in the laboratory frame as:

$$A_{\mathrm{FB}}^{p\bar{p}} = \frac{N(y(t) > 0) - N(y(t) < 0)}{N(y(t) > 0) + N(y(t) < 0)}, \tag{2.5}$$

where N is the number of events in the corresponding region with the condition on the rapidity. However, a better suitable observable is calculated in the $t\bar{t}$ rest frame as:

$$A_{\mathrm{FB}} = \frac{N(\Delta y > 0) - N(\Delta y < 0)}{N(\Delta y > 0) + N(\Delta y < 0)}, \tag{2.6}$$

where $\Delta y = y(t) - y(\bar{t})$. The A_{FB} is predicted from theory [20] at NNLO QCD+EW accuracy to be 0.095 ± 0.007, and is consistent with its combined measurement at Tevatron [21] 0.128 ± 0.025.

Similar asymmetric effects are also measured with the ATLAS and the CMS experiments at the LHC with proton-proton collisions. Since the $\mathrm{t\bar{t}}$ production at the LHC is dominated by gluon-initiated production, the asymmetric effects are suppressed as only quark-initiated production contributes to the asymmetry and hence gives an overall less value as a result. Besides this, at the LHC the initial quarks are mainly valence quarks and the anti-quarks are the sea quarks of the colliding protons. Since valence quarks have on average a larger momentum fraction than sea quarks, this leads to a larger longitudinal momentum for the initial quark than the anti-quark on average. Due to the absence of a fixed direction of the initial quark, in the presence of a positive valued forward-backward asymmetry, the top quarks would be produced more in the forward and backward regions than the central region, while the anti-top would be more centrally produced. This gets reversed for a negatively valued asymmetry. Because of this reason, a **central-forward asymmetry** is measured at the LHC and it is simply referred to as **charge asymmetry** A_C and defined as:

$$A_C = \frac{N(\Delta|y| > 0) - N(\Delta|y| < 0)}{N(\Delta|y| > 0) + N(\Delta|y| < 0)} \tag{2.7}$$

where $\Delta|y| = |y(t)| - |y(\bar{t})|$.

The latest measurement of A_C by the ATLAS experiment [2] for $\mathrm{t\bar{t}}$ production is 0.0068 ± 0.0015. As explained earlier, the asymmetric effects are present only via quark-initiated production and the fraction of such events is small in $\mathrm{t\bar{t}}$ production (∼10%). The sensitivity of the asymmetry measurements to potential new physics (e.g. axigluons) contributions could be improved by increasing this fraction. There are two approaches to achieving this.

- By looking at the phase space with higher values of the invariant mass of the $\mathrm{t\bar{t}}$ system $m_{t\bar{t}}$ or a longitudinally boosted $\mathrm{t\bar{t}}$ system increases this fraction.

- The associated production of $\mathrm{t\bar{t}}$ with either of γ [22, 23] or $\mathrm{W}^{\pm}$ [24] have a higher fraction of quark-initiated production and also the dominant asymmetric contributions are expected to be present from their leading order approximations.

2.5 Charge asymmetry in $t\bar{t}\gamma$ production

The $t\bar{t}\gamma$ production is a relatively rare process in the SM compared to the $t\bar{t}$ production. This is the dominant production mode (compared to single top production) of top quarks along with a photon,i.e. with a top-photon coupling. The evidence of the $t\bar{t}\gamma$ process was found by the CDF experiment [25] in 2011 and was discovered by ATLAS [26] in 2015. Since then the cross-section of the events containing $t\bar{t}\gamma$ production has been measured by the ATLAS and the CMS experiments [27–30] at $\sqrt{s} = 7, 8$ and 13 TeV.

This process gives the window to directly study the electric charge of the top quark. By measuring its cross-section, any deviation from the SM would indicate a different value of the top quark charge or the existence of anomalous dipole moments. Some example Feynman diagrams of the quark-initiated $t\bar{t}\gamma$ production are shown in Figure 2.6. The cross-section measurements were performed for the $t\bar{t}\gamma$ production including the radiative top decays ($t\bar{t}\gamma$ decay process, explained in Chapter 4.2). However, for this thesis, the analysis is performed only for the $t\bar{t}\gamma$ production to make the measurement more sensitive to the underlying physics.

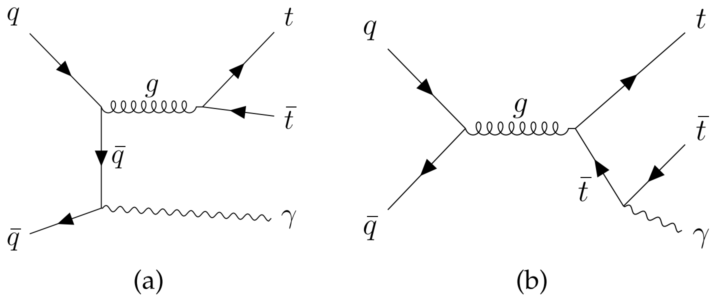


Figure 2.6: Example Feynman diagrams of $t\bar{t}\gamma$ production contributing to the charge asymmetry.

Effects of higher fraction of quark-initiated production $\sim 26\%$ in $t\bar{t}\gamma$ production, compared to $\sim 10\%$ in $t\bar{t}$ production [5], enhances the asymmetric contributions. Similar to $t\bar{t}$ production the asymmetric interference contributions considering the NLO QCD approximation are also present in $t\bar{t}\gamma$ production with a positive asymmetry value. Additionally, the interference of the diagrams of initial and final state photons as shown in Figure 2.6(a) Figure 2.6(b) respectively contributes to a larger (in absolute value) negative asymmetry. The interference arising from higher-order EW corrections is also expected to contribute to the asymmetry [31]. By considering all of the SM contributions, an overall negative asymmetry is expected for $t\bar{t}\gamma$ production.

Therefore, a first measurement of the A_C is performed for the $t\bar{t}\gamma$ production process. For this measurement, the single-lepton decay channel of $t\bar{t}$ pair (i.e. the e+jets and μ+jets channels, as in Figure 2.3(b)) are used. This is the most suitable channel for this analysis considering the requirement of charged leptons for easily determining the sign of the top quark charge in the final state and also has the largest branching fraction.

The charge asymmetry (A_C) measurement [3] is performed by first reconstructing the two top quarks together with reconstructed objects from the detector. After following the chain of the analysis, the distributions of the difference of the absolute rapidity of the reconstructed top quarks ($|y(t)| - |y(\bar{t})|$) are unfolded in a particle level *fiducial* region (defined in Chapter 4.1.1 and Chapter 4.5) to extract the charge asymmetry using a profile likelihood unfolding method.

CHAPTER 3

The LHC and the ATLAS Experiment

The $t\bar{t}\gamma$ production events are generated by colliding protons at a very high centre of mass energy. Such an experiment obviously cannot be done on a tabletop, but rather needs an accelerator with high-energy proton beams and the particles emitted from the collision are needed to be captured and recorded in an efficient manner to be analysed later. This is done with the ATLAS (A Toroidal LHC ApparatuS) experiment [32] at the Large Hadron Collider (LHC) [33] located at CERN near Geneva, Switzerland.

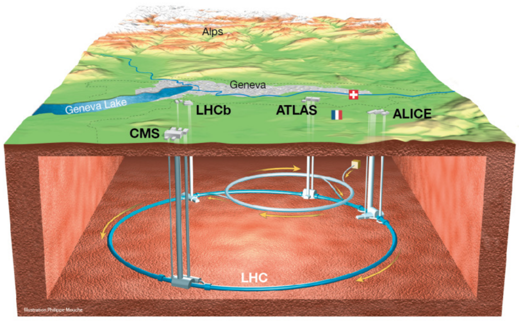


Figure 3.1: Overall view of the LHC. View of the 4 LHC detectors: ALICE, ATLAS, CMS and LHCb near Geneva, Switzerland [34]

The LHC is a particle accelerator that collides hadrons, i.e. not only protons but also heavy ions. It is installed in an underground ring with a circumference of around 27 km below the earth's surface. There are two beams in the ring that carries protons (for proton-proton collisions) and collide them at four of their intersection points where four different experiments are located to capture these collision events. ATLAS is one of these experiments, among the others: ALICE, CMS and LHCb. They are shown in a schematic view in Figure 3.1.

The LHC is designed to collide protons up to a centre of mass energy ($\sqrt{s}$) of 14 TeV. It has been operated at $\sqrt{s}$ = 7, 8 TeV as Run 1 during 2010-2012 and at 13 TeV as Run 2 during 2015-2018. Currently, it is being run at 13.6 TeV as Run 3. In the future, it is planned to run with much higher luminosity with some upgrade, referred to as *HL-LHC* run. The analysis in this thesis has been performed only with the Run 2 data.

3.1 Luminosity and pileup

The *instantaneous luminosity* from the proton beams determine the rate of the collisions of the physics processes and the number of total events depends on the luminosity integrated over the time of the operations, i.e. *integrated luminosity*. The primary luminosity measurement is performed using the LUCID-2 detector located around the LHC beam pipe on either side of the ATLAS detector.

During the Run 2 operation, an integrated luminosity of 156 $\text{fb}^{-1}$was delivered by LHC[1], of which ATLAS recorded 147 fb^{-1}. The data used in this analysis is considered only from those data-taking periods, where the ATLAS detector was fully operational and the recorded data satisfy quality criteria. This corresponds to an integrated luminosity of 138.97 fb^{-1} for the full Run 2 period. The uncertainty in the combined 2015–2018 integrated luminosity is ±1.7 % [35]. The collected data luminosity by each year is summarised in Table 3.1. The cumulative luminosity versus the time of operation is shown in Figure 3.2(a).

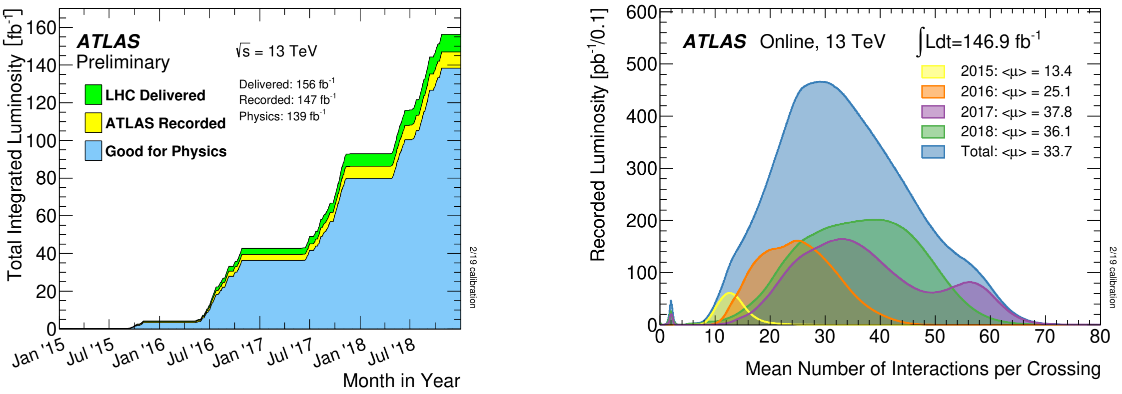
Table 3.1: Summary of data luminosity in the years between 2015 and 2018

Year	Integrated Luminosity [fb^{-1}]
2015	3.2
2016	33.0
2017	44.3
2018	58.5
Total	139 ± 1.7%

The protons are accelerated in bunches. Collisions between bunches occur every 25 ns. The high luminosity of data produced by LHC also has the drawback of having a large number of proton-proton interactions per bunch crossing. The mean number of interactions per crossing μ is shown in Figure 3.2(b) during Run 2. In the case of multiple interactions, the interactions coincident with the interest of physics interest are referred to as *in-time pileup*. In order to isolate the physics-relevant interaction,

---

[1] fb is femto-barn = $10^{-43}m^2$



(a) (b)

Figure 3.2: (a) Total Integrated Luminosity and Data Quality in 2015-2018: Cumulative luminosity versus time delivered to ATLAS (green), recorded by ATLAS (yellow), and certified to be good quality data (blue) during stable beams for pp collisions at 13 TeV centre-of-mass energy in 2015-2018. (b) Number of Interactions per Crossing: Shown is the luminosity-weighted distribution of the mean number of interactions per crossing for 2015 – 2018 pp collision data at 13 TeV centre-of-mass energy [36].

the interaction vertices are determined along with their associated particles. There can also be overlapping detector signals from interactions from adjacent bunches, this is called *out-of-time pileup*.

3.2 The ATLAS experiment

ATLAS is one of the two (the other one being CMS) general-purpose detectors at LHC. It is designed to study a large range of physics related to the Standard Model and beyond. A schematic cut-away view of the detector is shown in Figure 3.3.

The ATLAS detector is made with a concentric cylindrical geometry (*barrel*) that consists of several sub-detectors that surround the beam pipe symmetrically in concentric layers with the interaction point at the centre. The two outer ends (*endcap*) also consist of layers of sub-detectors. This covers almost the full solid angle of 4π from the interaction point.

3.2.1 Co-ordinate system

Naturally, a cylindrical coordinate system is used with a nominal interaction point as the origin and the beam direction as the z-axis. The x-y plane is transverse to the beam direction (z), with the corresponding azimuthal angle ϕ. The polar angle θ is the angle from the z-axis.

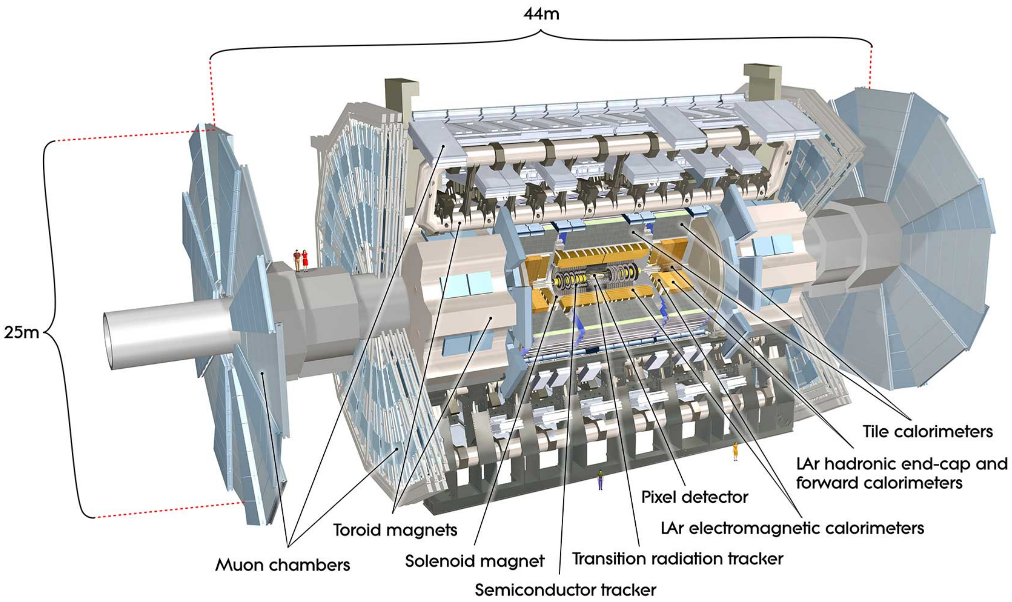


Figure 3.3: Cut-away view of the ATLAS detector [34]

A Lorentz invariant angular observable called *rapidity* (y) of a particle with energy E and momentum $\vec{p}$ (p_z along the z-axis) is defined as:

$$y = \frac{1}{2}\ln(\frac{E+p_z}{E-p_z}) \tag{3.1}$$

Another closely related (can be derived from rapidity in the limit of E$\approx|\vec{p}|$) observable called pseudorapidity (η) of a particle can be expressed in terms of its θ as:

$$\eta = -\ln[\tan(\theta/2)] \tag{3.2}$$

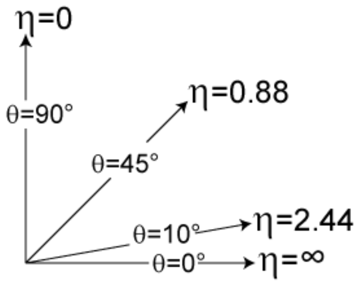


Figure 3.4: The relation between the pseudorapidity η and the polar angle θ [37]

This relation can be visualised in Figure 3.4. The angular distance between particles (i and j) are measured in the η-ϕ plane as

$$\Delta R(i,j) = \sqrt{(\eta_i - \eta_j)^2 + (\phi_i - \phi_j)^2} \tag{3.3}$$

where η_i, η_j, ϕ_i, ϕ_j are the pseudorapidity and the azimuthal angle of the particles i and j.

3.2.2 Detector sub-systems

Inner Detector

The sub-detector closest to the interaction point is called Inner Detector (ID) which reconstructs particle trajectories or tracks. This tracking system is enclosed inside a solenoidal magnetic field of 2 T that helps to bend the charged particle tracks. The tracks are measured precisely to get a high resolution of charged particles. The ID consists of three layers arranged radially outwards: pixel detector, semiconductor tracker and transition radiation tracker. The pixel detector, closest to the interaction point is made of high-granularity silicon detectors that help to reconstruct the position of the interaction vertices and to discriminate the primary vertex from the secondary vertices. The ID covers the region op to $|\eta| < 2.5$.

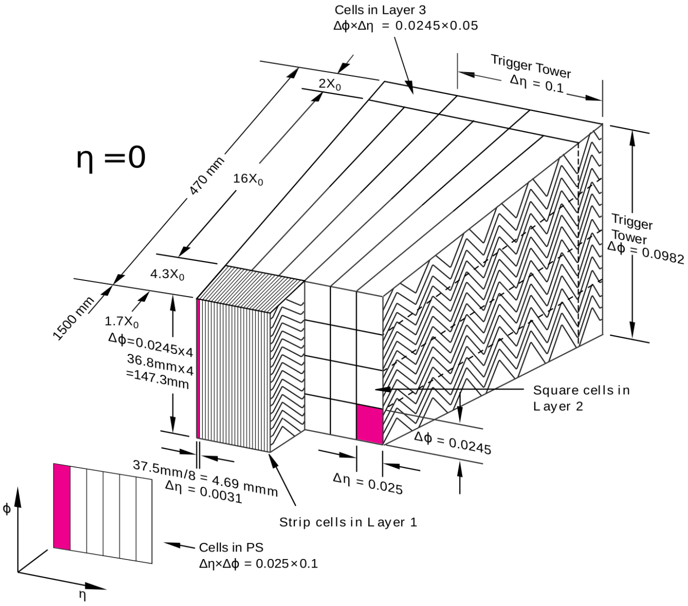


Figure 3.5: Sketch of the lateral and longitudinal segmentation of the ATLAS electromagnetic calorimeter around η=0 [38]

Calorimeter system

The ID is surrounded by the calorimeter system. This is divided into two sections: the electromagnetic (EM) calorimeter and the hadronic calorimeter. The hadronic calorimeter surrounds the EM calorimeter from its outer layer. The EM calorimeter provides a good measurement by containing the EM showers. The hadronic showers are measured by both the EM and hadronic calorimeter. The calorimeters cover the region up to $|\eta| < 3.2$ while there is a separate forward calorimeter in the region $3.1 < |\eta| < 4.9$. The EM calorimeter consists of 3 different concentric layers along with a pre-sampler (PS) layer at the front, each with their corresponding number of segmentation or *cells*. The first layer with fine granularity in η helps with accurate position measurement of electron and photon especially over the η region also covered by the ID ($|\eta| < 2.5$). The first and second layers together are best utilised for determining the direction of the particles. The energy of the shower is mostly contained in the second layer with a tail in the third layer. In Figure 3.5, the various layers of the EM calorimeter are shown along with the segmentations of the cells.

Muon detector

The outermost layer of the ATLAS detector is a muon spectrometer (MS) along with toroid magnets that surrounds the hadronic calorimeter. This provides the reconstructed tracks of muons. Muons pass the ID and the calorimeters without being absorbed and hence is measured by the MS. There are three layers of muon chambers. The system of toroid magnets is placed after the first inner layer of muon chambers. The toroid magnets consist of one barrel and two endcap toroids. These provide a magnetic field only to bend the muon tracks so that they can be precisely measured at high rapidity (outside the scope of the solenoidal magnet) by the muon chambers. The muons are reconstructed using either the ID or MS or in a combination with them. The muon chambers also trigger events containing muon candidates up to the range of $|\eta| < 2.4$.

3.2.3 The trigger system

Even though the collision rate at the interaction point is very high at 40 MHz, only a tiny part of that is interesting for physics analyses. The proton-proton cross-section is on the order of 10^{11} pb [39], while the expected cross-section of the $t\bar{t}\gamma$ production is around 2 pb. To make an efficient analysis by finding interesting events with limited storage of data, a trigger system is used that rejects a large part of the unwanted events and reduces the event rate from 40 MHz to 1 kHz.

The trigger system has two levels: a hardware-based trigger called first level trigger or *L1* and a software-based trigger called high-level trigger or *HLT*. The L1 trigger makes its decision quickly based on the information from the calorimeter system and muon detector reducing the event rate to 100 kHz. For such events, a region of interest (ROI) is built with the triggering objects identified and provided for the HLT to decide. The HLT uses algorithms for decision-making while narrowing down the event rate to 1 kHz. Only single-electron and single-muon triggers are used in the analysis.

CHAPTER 4

Description of the Monte Carlo based predictions

4.1 Monte Carlo simulated samples

All kinematic distributions needed for both the signal and background processes are simulated by Monte Carlo (MC) event generators. The event selection criteria and the optimisation of the analysis strategy are made based on these MC samples. The MC samples are also used to estimate the experimental and modelling-related systematic uncertainties. All MC samples are normalised to the most accurate theoretical prediction available of the corresponding process cross-section.

4.1.1 Event simulation

The physics processes in the pp collisions are complex. The resulting final states exhibit multi-dimensional kinematic features. These are propagated through various stages of physics evolution and eventually material interactions with the detector to be finally detected as electrical signals. It brings a challenging task to simulate the processes overall. These stages of the simulation procedure starting from the hard scattering of the process. They are explained in the following:

1. *Hard Scattering*

The physics processes at the core of the pp collisions involve large momentum transfers, referred to as the *hard scattering*. The matrix elements (ME) of the physics processes including all possible Feynman diagrams at a fixed order are taken into account to formulate a factorised multi-dimensional integration for the computation of the cross-section of the process in terms of various kinematic features of the final state particles. The initial state parton density functions (PDFs) are taken into account. The cross-section thus depends on the choice of the PDF as well as the scale of the factorisation (μ_F) and renormalisation (μ_R). The final state after this step is usually mentioned as the *parton level* in the following.

2. *Parton Showering*

The hard scattering describes the final state particles only with a fixed order approximation. The higher order effects are further approximated using a *parton shower* algorithm. The parton level final states, e.g. quarks and gluons are coloured and they radiate gluons initiating more radiations. The charged lepton final states can also similarly radiate photons. The parton shower algorithm simulates this evolution of the final state objects.

3. *Hadronisation*

The radiated quarks and gluons from the parton shower are coloured. So eventually they combine collectively in a colour-connected fashion to be confined and form colourless hadrons. The unstable hadrons decay into more stable hadrons. The hadronisation models simulate this step and they are largely independent for a given coloured system from the previous steps. The objects after this step are regarded as *particles* that are relevant for detectors and so this stage is referred to as the *particle level*.

4. *Underlying Events and Pile-up*

This is an additional corrective step in the simulation procedure. There are remnant coloured objects in the previous steps and they can further interact and evolve to create more final state particles. These are called *underlying events* and they are very important to simulate for a correct estimation of the physics process. The pileup as explained in Chapter 3.1, is also simulated for the MC.

5. *Detector Simulation*

The particle level events are finally simulated as an evolution interacting with the ATLAS detector components. This is done with Geant4 [40] software framework. The detector components are modelled in great detail with their shape, material and position. Since this step takes a long time, some of the simulated background samples are simulated using a fast simulation (called ATLFASTII [41]) approach using machine learning that mimics the Geant4 like output but at a much faster speed.

From the resulting simulated detector signals the interesting physics objects are reconstructed using the ATLAS reconstruction software [42] as described in Chapter 5.

4.1.2 Simulation of signal process

Production of the $t\bar{t}\gamma$ process is simulated with MADGRAPH5_AMC@NLO 2.7.3 [43] as a 2 → 3 process (Feynman diagrams shown in Chapter 2) from pp collisions at NLO QCD. The ME calculation employed the NNPDF3.0NLO set of PDFs [44]. The renormalisation (μ_R) and factorisation (μ_F) scales are dynamic and correspond to half of H_T, the sum over transverse masses of all final-state particles:

$$\mu_R = \mu_F = \frac{H_T}{2}, \quad H_T = \sum_f \sqrt{m_f^2 + p_{T,f}^2}, \tag{4.1}$$

where f runs over all final-state particles, and m_f and $p_{T,f}$ are the rest mass and the transverse momentum of particle f, respectively. Top quarks were decayed at LO using MADSPIN [45, 46] to preserve spin correlations. To avoid infrared and collinear singularities due to the photon radiation, kinematic selections are applied on the ME level. Final state leptons and quarks are required to have a minimum transverse momentum of 20 GeV and 1 GeV respectively. Final state photons are required to have a minimum transverse momentum of 15 GeV and be isolated according to a smooth-cone isolation (Frixione Isolation [47]) criterion. The cross-section with the above mentioned selection criteria applied at the ME level is 1.57 pb.

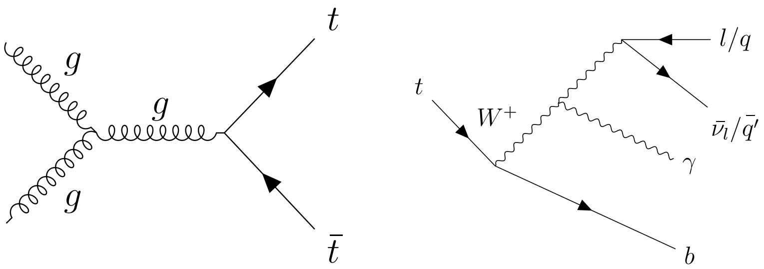
The event generation is interfaced to PYTHIA 8.240 [48] using the NNPDF2.3LO PDF set to model parton shower, hadronisation, fragmentation and underlying event. The decays of bottom and charm hadrons were simulated using the EVTGEN 1.6.0 program [49].

4.1.3 Simulation of background processes

$t\bar{t}$ production with a photon from other sources

The production of $t\bar{t}$ is one of the largest background processes (around 44% of all background events) showing quite similar signatures as the signal $t\bar{t}\gamma$ production. The production of $t\bar{t}$ with a photon radiated from one of the on-shell top quark or its decay products is called $t\bar{t}\gamma$ decay as shown in Figure 4.1. This is simulated with MADGRAPH5_AMC@NLO 2.7.3 as 2 → 2 LO $t\bar{t}$ production followed by the decay of top quarks at LO where either of the top quarks decays with a photon.

The NNPDF3.0NLO set of PDFs is used in the ME calculation. The renormalisation and factorisation scales are set to $H_T/2$ (Equation 4.1). The kinematic and isolation criteria at ME are the same as of $t\bar{t}\gamma$ production. Since the sample is only available at LO, an inclusive k-factor of 1.5 is used with the LO inclusive cross-section of 2.26 pb to normalise to NLO QCD. This k-factor was derived by comparing the normalisation of the sum of the NLO $t\bar{t}\gamma$ production sample and the NLO normalised (with k) LO $t\bar{t}\gamma$ decay sample with the theory prediction of [23]. The parton shower and hadronisation models are the same as that of $t\bar{t}\gamma$ production.



(a) (b)

Figure 4.1: Feynman diagrams of $t\bar{t}$ production with a photon radiating from one of the top decays where (a) the production stage is actually $t\bar{t}$ production, and (b) one of the top quark radiates a photon while decaying

$t\bar{t}$ production in hard scattering can also have reconstructed photons mainly from electron or hadronic sources misidentified as photons. For this, only the $t\bar{t}$ production process is simulated on ME level at NLO QCD using POWHEG-BOX-V2 [50–52]. The ME calculation uses the NNPDF3.0NLO PDF set, with the top-quark mass fixed to 172.5 GeV. The internal parameter h_{damp} to control the probability for gluon emission is set to 1.5 times the top-quark mass. By applying a k-factor, the events are normalised to a cross-section value calculated with the TOP++2.0 programme at NNLO QCD, including soft-gluon resummation to NNLL (see [53] and references therein), again assuming a top-quark mass of 172.5 GeV. The resulting cross-section for $t\bar{t}$ production at $\sqrt{s} = 13\,\text{TeV}$ is $\sigma_{t\bar{t}} = 831\,\text{pb}$. The parton shower and hadronisation models are the same as that of $t\bar{t}\gamma$ production.

Further, since it is well known that $t\bar{t}$ production also has asymmetric effects at NLO QCD, this is taken into account to further correct the $t\bar{t}\gamma$ decay LO MC as described in Chapter 9.5. The $t\bar{t}$ production MC also does not have exactly the same asymmetric value as predicted in the theory at the higher order. To account for this effect systematically the final distributions used for these processes are reweighted according to the corresponding truth top quarks with the $t\bar{t}$ charge asymmetry of 0.0064 that was calculated at NNLO QCD + NLO EW accuracy [20].

Single top quark + γ production

Single top quark production in association with a photon is one of the rare processes that appear as a background to the $t\bar{t}\gamma$ production process. There are 3 different channels of production mode (s, t and tW) and for each of them, the photon can be produced both at the production or the decay of the top quark. They are generated separately for top-quark and anti-top-quark production.

The s and t channel production modes are simulated at ME level without a photon at NLO QCD with POWHEG-BOX and the NNPDF2.3LO PDF set. The sample cross-sections are normalised to NNLO precision using k-factors [54–56]. The parton shower and hadronisation models are the same as that of $t\bar{t}\gamma$ production. The photon is expected to be modelled by the parton shower.

The tW channel production is done with separate generation for a photon from production and decay (tWγ) using MADGRAPH5_AMC@NLO 2.7.3 at LO QCD. The renormalisation and factorisation scales are set at $H_T/2$ (Equation 4.1).The parton shower and hadronisation models are the same as that of $t\bar{t}\gamma$ production. Since this process is not expected to show any asymmetry between the reconstructed top quarks (Chapter 8), the NLO QCD of tWγ process is not vital for the analysis. A large normalisation uncertainty is still taken into account in the fit (Chapter 11) due to a lack of knowledge of its cross-section.

Another tW channel sample without a photon is also produced to mainly account for the tW production process with possible fake photons. This is simulated at ME level at NLO QCD using POWHEG-BOX-V2 with the NNPDF3.0NLO PDF set, with the top-quark mass fixed to 172.5 GeV and the h_{damp} parameter is set to 1.5 times the top-quark mass. The parton shower and hadronisation models are the same as that of $t\bar{t}\gamma$ production. By considering the NLO QCD, the tW process has interference with the $t\bar{t}$ process, the interfered diagrams are removed using the *Diagram Removal* [57] method for this sample.

W,Z + γ + jets

Events with $W\gamma$ and $Z\gamma$ final states with additional jets are simulated with SHERPA 2.2.8 [58, 59] at NLO QCD using the NNPDF3.0NNLO PDF set. The simulation with SHERPA includes all steps of the event generation, from the hard process to the observable particles. All samples are matched and merged to the SHERPA-internal parton showering based on Catani-Seymour dipoles [60, 61] using the MEPS@NLO prescription [62–64]. Virtual corrections at NLO QCD in the ME are provided by the OpenLoops library [65, 66].

Additionally, events with W and Z bosons without photons (in hard-scattering) in association with additional jets are also generated to account for the possible events with fake reconstructed photons. These samples are simulated with SHERPA 2.2.1 at NLO QCD. The simulation uses NNPDF3.0NLO PDF set. The samples are normalised to the cross-section at NNLO QCD [67]. The events with additional prompt photons are simulated with the parton shower and only counted after the overlap removal with the corresponding W, Z + γ + jets samples.

WW,WZ,ZZ + γ

Events with two vector bosons, that is WW, WZ and ZZ (without photon at hard scattering) are very rare processes acting as backgrounds. These samples are generated with SHERPA versions 2.2.2 (purely leptonic decays) and 2.2.1 (for all other decays) at LO QCD with the NNPDF3.0NNLO PDF set. The samples are normalised to NLO QCD cross-sections [68]. The photons are simulated using the parton shower.

$\mathrm{t\bar{t}}$ + W,Z + γ

Events with a $\mathrm{t\bar{t}}$ pair with an associated W or Z boson (without photon at hard scattering) ($t\bar{t}V$) are very rare processes acting as backgrounds. These samples are simulated at NLO QCD with MADGRAPH5_aMC@NLO using the NNPDF3.0NLO PDF set. The photon is expected to be modelled by the parton shower. The samples are normalised to NLO QCD + NLO EW theoretical estimation [69]. The parton shower and hadronisation models are the same as that of $\mathrm{t\bar{t}}\gamma$ production.

4.2 Check of orthogonality between $\mathrm{t\bar{t}}\gamma$ production and $\mathrm{t\bar{t}}\gamma$ decay

The signature of the two processes: $\mathrm{t\bar{t}}\gamma$ production (the signal) and $\mathrm{t\bar{t}}\gamma$ decay (one of the backgrounds) are quite similar in their final states. This can raise doubt about the modelling of the two processes at the ME level.

The orthogonality of the two samples from each other is checked, based on the origin of the photon from their truth particle history records in the Monte Carlo events. This is not a very accurate estimate as certain modes may have interference between themselves. However, from the theory, the two processes are not known to be interfering with each other.

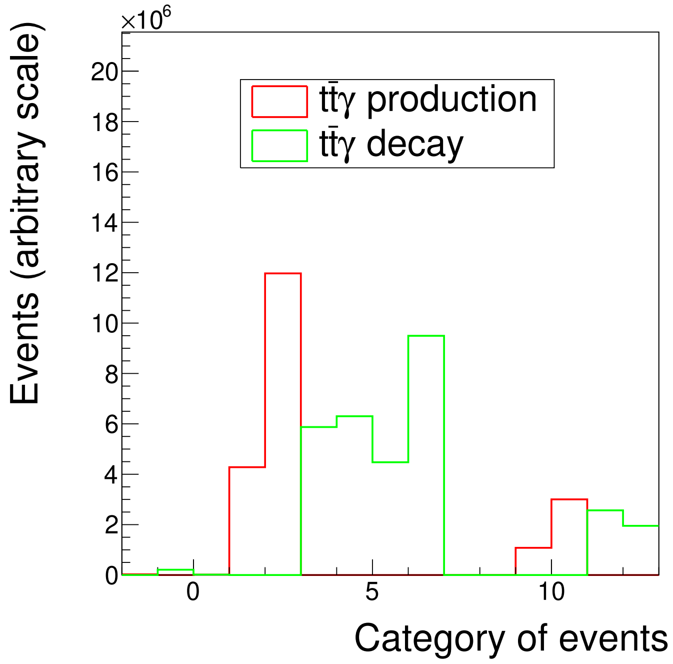
The events are categorised in 12 modes, other than the undefined ones:

- <=0: undefined
- 1,2,9,10: photon from production vertex or initial quark radiation or off-shell top quark

- 3-6,11,12: photon from on-shell top quark or its decay products

Further, category 1-6 refers to the single lepton decay channel and 9-12 to the dilepton channel. The categorisation of the two MC samples is shown in Figure 4.2. It can be clearly understood that the two samples have photon origins orthogonal to each other and are not overlapping.

The difference between the two samples relies on the fact that the photon in the $t\bar{t}\gamma$ production sample is generated at the production stage or from an off-shell top quark, while in the $t\bar{t}\gamma$ decay sample, it is generated as radiation at the decay stage from of an on-shell top quark or its decay products. The modelling of both samples uses a small width of the top quark while decaying. In theoretical calculations [23], a narrow-width approximation is used for the same. However, for a sample generated with a large width of the top quark, the definition of on-shell-ness would change and hence also the cross-section of the $t\bar{t}\gamma$ decay sample that could modify the fractions of the different modes as shown in the Figure 4.2. Still, the orthogonality would hold.



(a)

Figure 4.2: Check of orthogonality between $t\bar{t}\gamma$ production and $t\bar{t}\gamma$ decay processes

4.3 Removal of overlapping events

For all of the simulated samples described above where the photon is generated at the ME level ($X\gamma$), events outside the kinematic requirements are missed. To account for the missing events the corresponding sample where the photon is generated by the parton shower (X) is considered. This causes a large overlap region between the $X\gamma$ and X sample. So the overlap removal procedure is applied to all such samples as follows.

- All events from the $X\gamma$ samples are selected since the photon radiation simulated on the ME level comes with higher accuracy than the radiation accounted for in the showering algorithm. Also, the $X\gamma$ samples have a larger number of events for the overlapped phase space.
- Events from the X samples if they overlap with the $X\gamma$ simulation are removed.

The overlap removal selection criteria are applied based on the truth origin particle of the photon: $\gamma^{\text{orig}}_{\text{truth}}$ and truth leptons. The following requirements are applied: $p_{\text{T}}(\gamma^{\text{orig}}_{\text{truth}}) > 15\,\text{GeV}$ and $\Delta R(\ell, \gamma^{\text{orig}}_{\text{truth}}) > 0.2$, where $\gamma^{\text{orig}}_{\text{truth}}$ should be a photon and the leptons (ℓ) refer to any leptons at the truth level.

The events are considered to be in the overlap region if the above criteria are satisfied. Events from X samples, for example, $t\bar{t}$, W+jets and Z+jets are *vetoed* if they fall into the overlap region. This overlap region is large and the fraction of events that are kept (the non-overlap region) corresponds only to about 5.6%, 3%, and 3% for the $t\bar{t}$, W+jets and Z+jets processes, respectively, of the total number of selected events after following the event selection described in Chapter 6.2.

4.4 Categorisation of photons and grouping of processes

The photon or electron generated only at the ME level are of physics interest. These are called *prompt* objects. However, there can be reconstructed photon or electron objects that are generated from other sources or there are other objects that are misidentified as electrons or photons. These are called *fake* objects. The origin of the objects is found in their truth particle history records in the Monte Carlo events of the corresponding process. The fake photons are of two types:

- ***Electron Fake*** (e-fake): an electron which fakes a photon signature in the detector. It is denoted as *e-fake* γ in the following. The truth origin particle of the fake photon should either have a PDG ID of an electron or a photon. In the case of a photon origin, there should be a truth electron within $\Delta R(\gamma^{\mathrm{orig}}_{\mathrm{truth}}, e_{\mathrm{truth}})$ = 0.1

- ***Hadron Fake*** (h-fake): Non-prompt photons originating from hadrons (e.g. $\pi^0 \to \gamma\gamma$ decays) or hadronic energy depositions in the calorimeter mimicking/faking as photon signatures. It is denoted as *h-fake* γ in the following.

The photon candidates with none of the above two truth origin descriptions are considered as ***prompt*** photon. The estimations of the fake photons in data and comparison with the MC estimations are described in Chapter 9 in detail.

4.4.1 Grouping of processes

For the convenience of the analysis, the signal and background processes are grouped according to a category of the photon origin, i.e., prompt or fake photon, and also by merging the small backgrounds. The grouping of all the processes is described in Table 4.1.

Table 4.1: The grouping of processes

Groups	Processes
	Prompt γ from
$t\bar{t}\gamma$ production	$t\bar{t}\gamma$ production
$t\bar{t}\gamma$ decay	$t\bar{t}$ +γ from other sources
Prompt γ	W,Z + γ + jets
	single top quark +γ [s,t,tW channels]
	WW,WZ,ZZ + γ
	$t\bar{t}$ + W,Z + γ
h-fake photon γ	Hadron fake γ from all processes
e-fake photon γ	Electron fake γ from all processes

4.5 Definition of fiducial region at the particle level

The particle level (Section 4.1.1) definitions of objects are used to define a *fiducial* region where the measurement of the charge asymmetry is extracted. This is necessary to obtain a measurement where the effects of the ATLAS detector are removed so that it can be easily compared with other experiments or theoretical predictions. This fiducial region at the particle level is used in the unfolding of the measurement described in Chapter 11.

The event selection is performed with the particle level objects and the selection criteria are similar to the ones at the reconstruction level described in Chapter 6

Table 6.2. The events are required to have either one electron or one muon with p_{T} >25 GeV and $|\eta|$<2.5, one photon with p_{T} >20 GeV and $|\eta|$<2.37, at least four jets with p_{T} >25 GeV and $|\eta|$<2.5 and at least one of the jets with B-hadron. The charge asymmetry A_C calculated for the signal MC is **-0.014** ± **0.001** (scale uncertainty) in this region.

CHAPTER 5

Reconstruction of physics objects

In a particle collision event at the ATLAS experiment, different particles are produced at the point of collision and are detected by the various sub-detector components. These physics objects are identified from the detector's electronic signals and their kinematics are determined using reconstruction algorithms. Their kinematic properties are further used to select the physics processes of interest. Thus, for a precise measurement the expected physics objects are necessary to be well reconstructed. The reconstruction of the photon (γ), charged lepton (ℓ) i.e, electron (e) and muon (μ)[1], jet (j) including flavour tagging, missing transverse energy ($E_{\mathrm{T}}^{\mathrm{miss}}$) are briefly discussed in this chapter.

5.1 Primary vertex

Primary vertices are the proton-proton interaction points. They are determined from the charged particle tracks reconstructed via the Inner Detector hits. The tracks are required to be spatially compatible with the interaction region, i.e., small transverse impact parameter. Each event can contain multiple primary vertex candidates. The number of vertices increases with higher pileup. The hard-scatter primary vertex is chosen to be the one with at least two associated tracks and the highest $\sum p_T^2$, where the sum extends over all tracks with $p_{\mathrm{T}} > 400$ MeV matched to the vertex. The other primary vertex candidates are considered pileup.

---

[1] Tau (τ) lepton reconstruction is not done separately and the leptonic decay of Tau leptons to e or μ are considered as part of the corresponding process.

5.2 Photon and electron

Disclaimer: The combination of the photon identification efficiencies is the main contribution of the author in this Chapter and hence the content is written intending to provide a pretext for explaining the Section 5.2.2

Electrons and photons are reconstructed from the energy deposited in calorimeter cells (Figure 3.5) using a dynamical topological cell clustering algorithm [38]. The variable-sized topological clusters can include the low-energy *bremsstrahlung* photons clusters (satellite) with the electron cluster (seed) to form a *supercluster*. The clustering algorithm uses the individual cell energy significance, i.e, absolute cell energy over the cell noise (electronic and pileup noise) to create the clusters. The satellite clusters are accepted in the supercluster only if it falls within a window of $\Delta\eta \times \Delta\phi = 0.075 \times 0.125$ (3×5 calorimeter cells) around the seed cluster barycenter. Since both hadronic and electromagnetic calorimeter cells are used in the algorithm, only the clusters with electromagnetic (EM) energy fraction above 0.5 and the EM energy above 400 MeV are considered for the algorithm to reject pileup.

The reconstruction of electrons and photons also relies on tracking information. The matching of reconstructed tracks to the calorimeter clusters is needed in identifying the object:

- **electron**: supercluster (with $E_T > 1$ GeV) matched with track with at least 4 hits in the Inner Detector
- **converted photon**: topoclusters associated with a secondary conversion vertex ($\gamma \rightarrow e^+e^-$ in the Inner Detector) with loosely matched tracks
- **uncoverted photon**: supercluster (with $E_T > 1.5$ GeV) that cannot be associated to any track or vertex

The energy scale of the calorimeter cells is calibrated depending on their detector geometry using the reconstructed Z mass in Z $\rightarrow e^+e^-$ decays. Additionally, they are validated with $J/\psi \rightarrow e^+e^-$ and radiative Z boson decays for electron and photon respectively.

For the physics analysis, the electron and photon candidates are further required to pass more stringent requirements, to be identified as the prompt objects that can come only from the hard scatter and not from hadron decays.

5.2.1 Photon identification

A set of quantities in the EM calorimeter characterising the shape of the shower development are used in the optimisation of the photon identification. These are called *shower shape variables*. Two reference sets of selection criteria labelled as *loose* and *tight* are defined with increasing background rejection power. The loose selection uses a set of optimised cuts with few of the shower shape variables that take into account the leakage to the hadronic calorimeter and the backward layers (layer-2 and layer-3 as mentioned in Chapter 3) of the EM calorimeter. The tight selection additionally uses the shower shape variables from the front layer of the EM calorimeter that has fine granularity. The shower shape variables for the converted photons can be distinguished from those of the unconverted photons using information from this layer. The tight identification is therefore optimised separately for the converted and unconverted photons. In total 10 shower shape variables are exploited to define the tight identification criteria. This ensures maximum background rejection while keeping the prompt photon efficiency high. The *working point* is chosen from the multivariate discriminant based on the signal significance calculated as $\frac{\epsilon(\mathrm{sig})}{\sqrt{\epsilon(\mathrm{background})}}$. The efficiencies ($\epsilon$) are calculated as the ratio of the events passing the corresponding tight identification and the events selected from the loose identification criteria. Since there is a dependence of the shower shape variables on the photon E_T that leads to lower signal efficiency in low-E_T and non-optimal tight identification at high-E_T, a E_T-dependent tight identification is performed by using different methods for photons in the different ranges of photon E_T: [10,25], [25,100] and [100,1500] GeV.

5.2.2 Measurements of photon identification efficiency

The performance of the photon identification method is assessed by measuring its efficiency in data. Three different methods are used to measure the photon identification efficiency:

- **Radiative-Z method** that includes the low energy photons (10-100 GeV)
- **Matrix method** mainly for higher energy range (25-1500 GeV)
- **Electron extrapolation method** for the mid energy range (25-250 GeV)

They are discussed briefly in the following. A detailed description of the methods can be found in [38]. It follows with the comparison and combination of the methods as the primary focus of this section.

Radiative-Z method

A Z boson decaying to a lepton pair along with the final state radiation (FSR) photon (Z→ $\ell\ell\gamma$) gives a low background environment to study the photon identification efficiency at low energies. A loose event selection criteria is used for this method.

Additionally, cuts on $m_{\ell\ell\gamma}$ and $m_{\ell\ell}$ are used to only select FSR events and reject initial state radiation events (Z($\rightarrow \ell\ell$) + γ): $m_{\ell\ell\gamma}$ ϵ[80,100] GeV and $m_{\ell\ell}$ ϵ[40,83] GeV. The identification efficiency is calculated as the fraction of selected photon candidates that pass the tight identification criteria. Since there are background events from Z($\rightarrow \ell\ell$)+jets processes where a jet is misidentified as a photon in the selected region, a background correction is done before calculating the efficiency. The background contamination is estimated using MC prediction to be about 15% for $10 < E_T < 25$ GeV and $\leq$ 1% for higher energies. A template fit on the $m_{\ell\ell\gamma}$ distribution in the low energy region is used to evaluate the signal purity to calculate the identification efficiency. The fits are performed separately for ee and $\mu\mu$ events. A weighted average method is used to get the final identification efficiency from the two measurements.

Several sources of systematic uncertainties are taken into account for efficiency measurement. They are evaluated by repeating the measurement with variations and then taking the resulting difference as the uncertainty. Most of them are only applicable to low E_{T} regions with higher background contamination. The uncertainties are listed below with explanations of their variation:

- **Background:** Different fit ranges. Only applicable for [10,25] GeV of E_{T} bins.
- **Fudge factor:** Variation of shower shape variables. Only applicable for [10,25] GeV of E_{T} bins.
- **Generator:** Alternate MC event generator sample.
- **Closure:** Use of true photon matching. Only applicable for [10,25] GeV of E_{T} bins.

Matrix method

This method is based on the inclusive single prompt photon production process (mainly via qg $\rightarrow$ γq). The highest p_{T} photon matched to a trigger photon is considered for this study. The trigger matching requires the reconstructed photon to be within $\Delta R \leq 0.07$ of the trigger photon. The identification efficiency is the number of selected loose photon events to pass the tight identification criteria. The selected dataset also contains background events from the di-jet process. A track isolation criteria are used to discriminate between prompt and fake photons, i.e, there should be no track within $\Delta R = 0.4$ of the photons, however excluding tracks within $\Delta R = 0.1$ to account for conversion and bremsstrahlung photons. Four regions are defined for the selected events based on the combination of the two criteria: pass/fail of the tight identification and pass/fail of the track isolation requirement.

The identification efficiency is extracted by calculating the track isolation efficiencies in the four regions.

Several sources of systematic uncertainties are taken into account for this efficiency measurement as well and are listed below with explanations of their variation:

- **Track isolation:** They are evaluated by repeating the measurement with the size of the track isolation cone changed to $\Delta R = 0.2$ and then taking the resulting difference as the uncertainty.
- **Fudge factor:** Uncertainty from the variations of shower shape variables
- **Closure:** Uncertainty to account for the mismatch between the efficiencies in their derived region and the applied region.
- **MC statistics:** Uncertainty to account for the limited events in the MC samples used for the measurement.

Electron extrapolation method

This method uses Z→ee events in a region with two electron candidates within the invariant mass range: $70 < m_{\mathrm{ee}} < 110$ GeV. There are also background events from W+jets and multijet processes in this region. To get the efficiency for the photon identification efficiency the electron shower shape variables are modified to resemble photon shower shapes using a Smirnov Transform [70] derived from the MC samples of Z→ee and inclusive-photon production. The identification efficiency is calculated by fitting the data in the region with the simulated signal and the background templates estimated from control regions.

Several sources of systematic uncertainties are also considered for this efficiency measurement and explained below:

- **Closure:** Uncertainty to assess whether the transformed electron can reproduce the expected photon efficiency from MC.
- **Fragmentation photon:** The measurement is repeated by changing the fraction of fragmentation photons by ±50% for the Smirnov Transform and the difference is taken as the uncertainty. This assesses the impact of the simulation.
- **Fudge Factor:** Uncertainty from the variations of shower shape variables.
- **MC Statistics:** Uncertainty to account for the limited events in the MC used for the measurement.
- **Background:** The measurement is repeated with the selected region changed to $80 < m_{\mathrm{ee}} < 100$ GeV and the difference is taken as uncertainty.

Combination of photon identification measurements

All three methods measure the efficiency as a function of the bins of the photon's kinematic variables: E_{T}, $|\eta|$ and conversion type. The final values are found by combining the three measurements in the same bins to improve the precision. This is only applicable in the bins where more than one measurement is available i.e. overlapped regions.

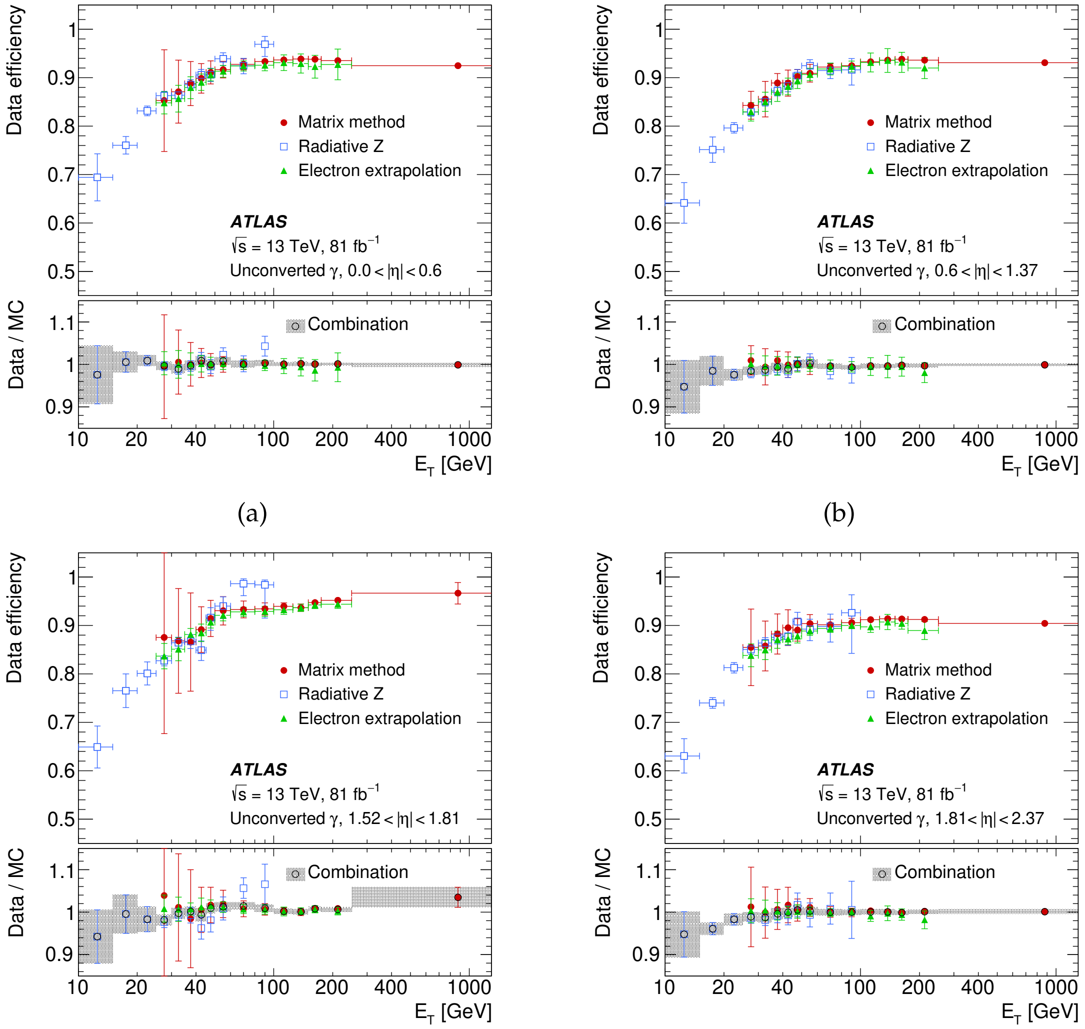
The individual efficiencies are measured for both data and MC and their ratio i.e. scale factor is calculated so that the efficiencies in MC can be corrected with the ones derived in the data. The scale factors are combined for the overlapped regions. The combination of the scale factors is done using the BLUE [71] method. This method gives an average of the measurements while taking care of all possible correlations among the same systematic uncertainties of the individual measurements. The systematic uncertainties of the individual methods are present as discussed above and are listed in Table 5.1, where the possible correlations considered are also indicated. The possible correlations between the measurements are considered only for the two systematic uncertainties: Closure and Fudge Factor from the knowledge of the individual measurement methods [38]. Any possible correlations between different systematic uncertainties of the same measurements are not considered for the combination process as they are assumed to be uncorrelated.

Table 5.1: List of the systematic uncertainties present for all of the three measurements and also the uncertainties for which the correlation is considered are indicated

	Methods			
Systematic Uncertainty	Radiative-Z Method	Matrix Method	Electron Extrapolation Method	Correlation considered
Background	✓		✓	
Closure	✓	✓	✓	✓
Fudge Factor	✓	✓	✓	✓
MC Statistics		✓	✓	
Generator	✓			
Track Isolation		✓		
Fragmentation γ			✓	

The combination is done for the three scale factor measurements $y_\alpha \pm \sigma_\alpha$ ($\alpha = 1, 2, 3$), where

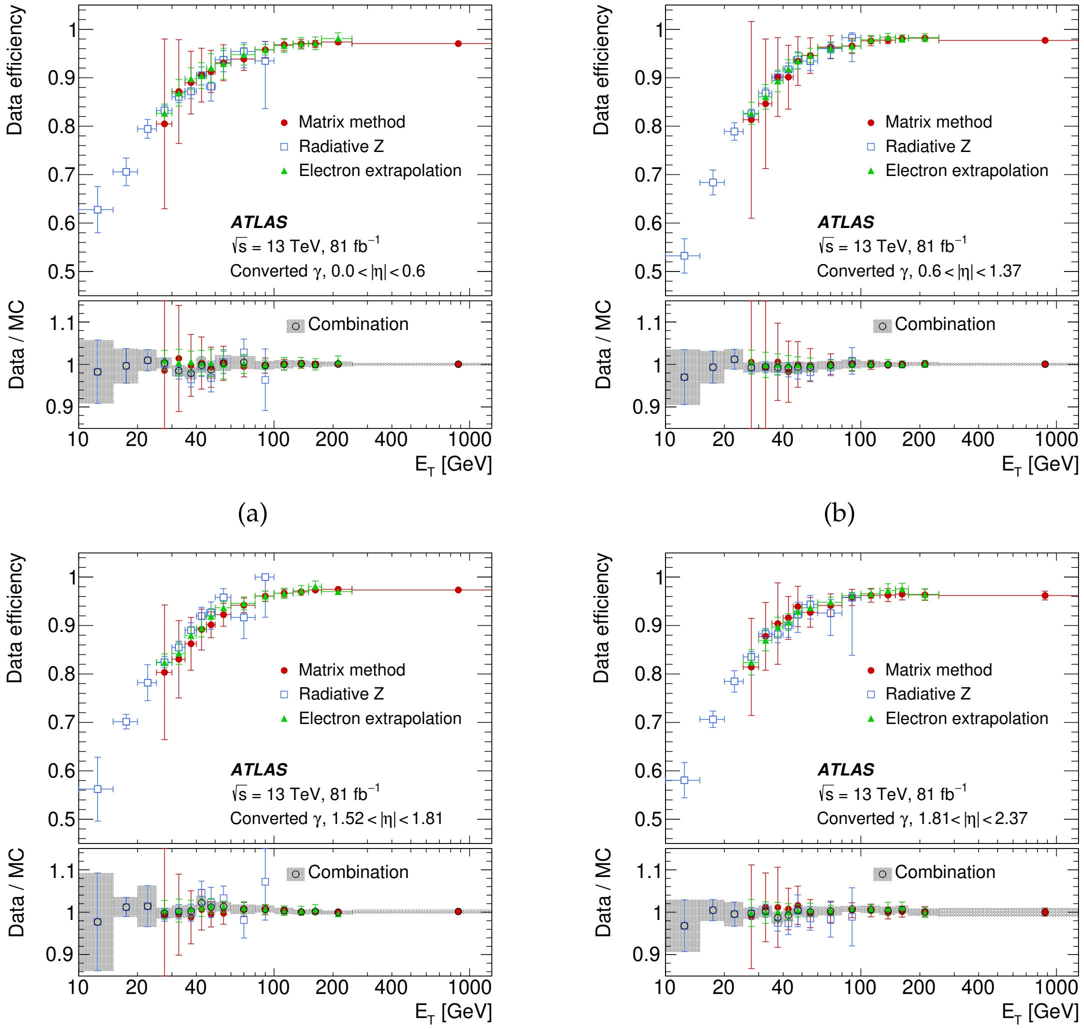
$$\sigma_\alpha = \sqrt{\left(\sum_{s=1}^{N_{syst}} \sigma_{s,\alpha}^2\right) + \sigma_{\mathrm{stat},\alpha}^2}$$



(a) (b)

(c) (d)

Figure 5.1: The photon identification efficiency, and the ratio of data to MC efficiencies, for unconverted photons with a Loose isolation requirement applied as preselection, as a function of E_T in four different $|\eta|$ regions. The combined scale factors, obtained using the BLUE method are also presented; the band represents the total uncertainty. Published in [38].



(a)

(b)

(c)

(d)

Figure 5.2: The photon identification efficiency, and the ratio of data to MC efficiencies, for converted photons with a Loose isolation requirement applied as preselection, as a function of E_T in four different $|\eta|$ regions. The combined scale factors, obtained using the BLUE method are also presented; the band represents the total uncertainty. Published in [38].

is the total error of the individual measurements that take into account N_{syst} number of systematic uncertainties as well as the statistical uncertainty. The individual systematic uncertainty of individual measurement (α) is expressed as $\sigma_{s,\alpha}$ and for all three measurements can be expressed together as a 1-dimensional vector: $\sigma_s = (\sigma_{s,\alpha})\,|_{1\times 3}$. The correlation matrix between the measurements for a particular systematic uncertainty is $\rho_s|_{3\times 3}$ with values between 0 and 1. In case of no correlation, ρ_s is a diagonal unit matrix.

The combined measurement, $\hat{y} \pm \sigma(\hat{y})$ is calculated as:

$$\hat{y} = \sigma^2(\hat{y}) \times \left(\mathbf{1}^T V_y^{-1} \mathbf{y}\right) \tag{5.1}$$

and

$$\sigma(\hat{y}) = \frac{1}{\sqrt{\mathbf{1}^T V_y^{-1} \mathbf{1}}} \tag{5.2}$$

where

$$V(\hat{y}) = \left(\sum_{s=1}^{N_{syst}} \sigma_s^T \rho_s \sigma_s\right) + \sigma^2_{\text{stat}} \tag{5.3}$$

and **1** are the 1-dimensional vector of order 3 with all elements being 1.

For the uncertainties Closure and Fudge Factor, a suitable value of the correlation is indirectly determined, by varying the non-diagonal elements of one such systematic uncertainty, ρ_s all together as a single variable and its impact on the combined measurement is tested. No significant impact is found for both the uncertainties and hence no correlation is considered for the final measurement.

A number of different approaches regarding bin-to-bin correlations of the photon's E_{T} are also considered, however, are not used in the analysis of the thesis.

In Figures 5.1 and 5.2, the three efficiency measurements are compared for all bins of kinematic variables (E_{T}, $|\eta|$ and conversion type) and their scale factor are shown in the bottom panels, where the combined measurements are also shown. The combined scale factors are close to 1 and their uncertainties are smaller than the smallest among the three measurements. These final scale factors are used in the analysis denoted as SF (ID) for a photon in Chapter 6.3.

5.2.3 Electron identification

The identification of electrons assures separation from similar objects: converted photons, hadronic jets and electrons produced from heavy-flavour hadron decays. Track information as well as the shower shape variables (same as in the photon identification) are used for this discrimination. The discriminant is a ratio of likelihoods created for the signal and background processes. The likelihoods are created from the normalised distributions of discriminating variables in different bins of $|\eta|$ and E_{T} of the electron object. The distributions for signal are derived from $J/\psi \rightarrow$ ee for E_{T} <15 GeV and the Z→ee process for E_{T} >15 GeV. A *tag-and-probe* [72] method is used to select a reconstructed electron. A tight selection of the likelihood discriminant is used for the electron identification for the analysis of this thesis. The recommended requirements on the impact parameter variables are also applied. d_0 significance, $|d_0|/\sigma(d_0) < 5$ and $|\Delta z_0 \sin(\theta)| < 0.5\,\mathrm{mm}$ is required.

5.2.4 Photon and electron isolation

The selected photon and electron objects are needed to be isolated from nearby particles. Both track isolation and calorimeter isolation requirements are used.

The calorimeter isolation is calculated by adding the transverse energy of all the topological clusters centred within a specified cone of the same barycentre of the electron or photon supercluster, and then subtracting the particle core energy (calorimeter cells contained in a $\Delta\eta \times \Delta\phi = 5 \times 7$ window around the barycenter of the particle). Also, corrections from pile-up, underlying events and estimated leakages are applied. The calorimeter isolation depends on the chosen cone size. For electron isolation a cone size of $\Delta R = 0.2$ is used and for photon isolation, both $\Delta R = 0.2$ and 0.4 are used depending on the isolation working point.

The track isolation is calculated by adding the p_{T} of the tracks within a specified cone around the electron track or photon cluster. For electron and converted photons, the tracks matched to them are excluded from the isolation calculation. For electron, a p_{T} dependent variable cone size is used that shrinks for larger p_{T}. The isolation working points for both electron and photon are summarised in Table 5.2.

Table 5.2: The requirements of the calorimeter and track isolation for both electron and photon [38] to define the loose and tight working points

Physics Object	Working Point	Calorimeter Isolation	Track Isolation
Photon	Loose	$E_T^{\mathrm{iso}}\|_{\Delta R<0.2}$ $< 0.065 \times E_{\mathrm{T}}$	$p_T^{\mathrm{iso}}\|_{\Delta R<0.2}$ $< 0.05 \times E_{\mathrm{T}}$
Photon	Tight	$E_T^{\mathrm{iso}}\|_{\Delta R<0.4}$ $< 0.022 \times E_{\mathrm{T}} + 2.45\,\mathrm{GeV}$	$p_T^{\mathrm{iso}}\|_{\Delta R<0.2}$ $< 0.05 \times E_{\mathrm{T}}$
Electron	Loose	$E_T^{\mathrm{iso}}\|_{\Delta R<0.2}$ $< 0.2 \times p_{\mathrm{T}}$	$p_T^{\mathrm{variable-iso}}\|_{\Delta R<0.2}$ $< 0.15 \times p_{\mathrm{T}}$
Electron	Tight	$E_T^{\mathrm{iso}}\|_{\Delta R<0.2}$ $< 0.06 \times p_{\mathrm{T}}$	$p_T^{\mathrm{variable-iso}}\|_{\Delta R<0.2}$ $< 0.06 \times p_{\mathrm{T}}$

5.3 Muon

To reconstruct muons, either or both, the inner detector tracks and the track segments in different layers of the Muon Spectrometer are used by combining them using track fitting algorithms [73]. Near the vicinity of $\eta = 0$, calorimeter information is used to identify muons, although with lower efficiency. The recommended requirements on the impact parameter variables are also applied: d_0 significance, $|d_0|/\sigma(d_0) < 3$ and $|\Delta z_0 \sin(\theta)| < 0.5\,\text{mm}$ is required. A tight isolation criterion mainly based on track isolation and p_T-dependence is used to select the muons. With the chosen identification criteria the rate of misidentification of hadrons as muons is very low.

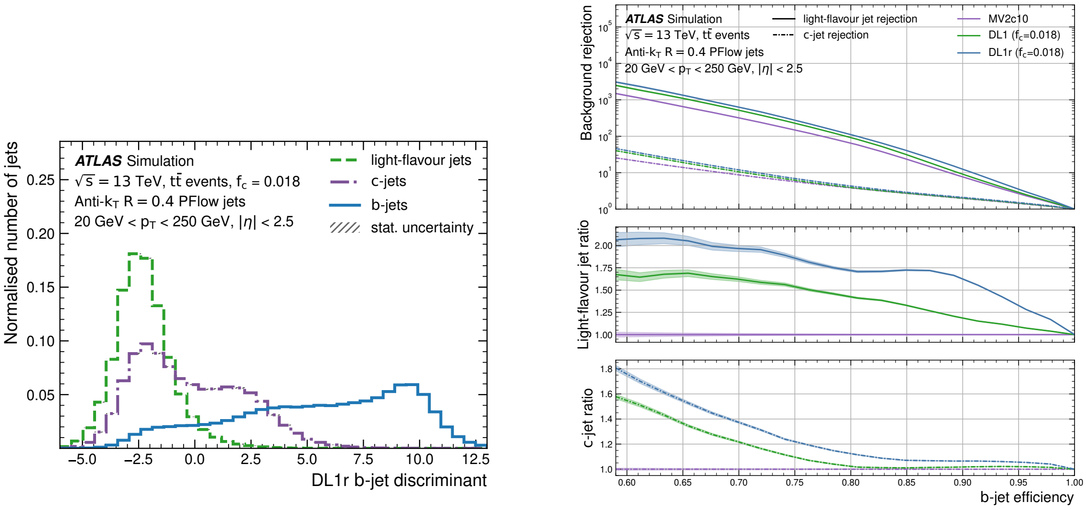
5.4 Jets

Jets are reconstructed using the anti-k_t algorithm [74] in the FASTJET implementation [75] with a radius parameter $R = 0.4$ (in the η–ϕ plane). Their reconstruction is performed on particle-flow objects (PFlow) [76]. The PFlow algorithm directly combines measurements from both the tracker and the calorimeter to form the input signals for jet reconstruction, which are intended to approximate individual particles. They are calibrated to be consistent with electromagnetic cluster shapes using corrections determined in simulation and inferred from test-beam data. The jet energy and direction are calibrated using an energy and η-dependent simulation-based calibration scheme with in situ corrections based on data [77]. The jets are required to have $p_\text{T} > 25\,\text{GeV}$ and $|\eta| < 2.5$. To reject jets from pile-up or other primary vertices, jets are required to have a *Jet Vertex Tagger* (JVT) discriminant [78] value larger than 0.59 for jets with $p_\text{T} < 60\,\text{GeV}$ and $|\eta| < 2.4$. The JVT efficiency is corrected by a scale factor extracted from a comparison of data and simulation.

5.4.1 Flavour tagging

Since the top quark always decays with a bottom (b-) quark that creates a B-hadron containing jet or *b-jet*, identifying such jets helps in the event selection of the analysis. Jets with other flavour origins (up, down, charm or strange quark) are not necessary to be distinguished among themselves for this analysis.

To identify b-jets, a flavour tagging algorithm called DL1r [79] using a deep neural network was used. The DL1r output (the output distribution of the DL1r algorithm) is shown in Figure 5.3 (a). The efficiency of the algorithm to reject light- and c-flavour jets is shown as a function of the b-jet tagging efficiency in Figure 5.3 (b). DL1r performance is also compared with the two algorithms developed previously by the ATLAS collaboration: DL1 and MV2c10. The flavour-tagging efficiencies of *b*-jets as well as of *c*-jets and light jets are calibrated with data. Jets are considered as b-tagged if the value of the DL1r-algorithm output is larger than a certain threshold that provides a 77% b-jet tagging efficiency when measured with simulated SM $t\bar{t}$ events, referred to as the 77% working point (WP). The use of 77% b-tagging WP (among the four available WPs: 60%, 70%, 77% and 85%) is optimised by keeping a balance between the signal statistics and the background discrimination after comparing the total signal-to-background ratio and the ratio of $t\bar{t}\gamma$ production and $t\bar{t}\gamma$ decay events for various combinations of WP and the number of b-tagged jets.



(a)

(b)

Figure 5.3: (a) Distribution of the output of the DL1r b-tagging algorithm for b-jets, c-jets and light-flavour jets in $t\bar{t}$ simulated events. (b) The light-flavour jet and c-jet rejection factors as a function of ϵ_b for the high-level b-taggers MV2c10, DL1, and DL1r. The lower two panels show the ratio of the light-flavour jet rejection and the (c)-jet rejection of the algorithms to MV2c10 [79]

5.5 Missing transverse momentum

Neutrinos do not interact with the materials in the detector and hence they leave no signal. From the conservation of momentum in the transverse plane of the collision, we can find the missing momentum from the negative sum of all other detected particles. This is called Missing Transverse Momentum (also known as Missing Transverse Energy. The magnitude is denoted as $E_{\mathrm{T}}^{\mathrm{miss}}$), $\vec{E}_{T,\mathrm{miss}} = E_{\mathrm{x}}^{\mathrm{miss}}\hat{x} + E_{\mathrm{y}}^{\mathrm{miss}}\hat{y}$. The sum includes the momenta of all the reconstructed objects as well as the tracks associated with the primary vertex that is not part of any of the reconstructed objects.

The latter part is called *Soft Term*. The missing momentum's x- and y-component are calculated separately as following:

$$E^{\text{miss}}_{\text{x(y)}} = -\{E^{\text{lepton}}_{\text{x(y)}} + E^{\text{photon}}_{\text{x(y)}} + \sum_{jets} E^{\text{jet}}_{\text{x(y)}} + E^{\text{SoftTerm}}_{\text{x(y)}}\} \tag{5.4}$$

CHAPTER 6

Event selection

From the collection of detector signals to the physics analysis, there are a few steps that are followed:

- Usage of Triggers to select only events with leptonic final states
- Reconstruction of objects (Chapter 5)
- Removal of the reconstructed objects in overlap
- Applying the selection criteria on the reconstructed objects
- Calculation of weights for individual events applied to the Monte Carlo events to correctly match the prediction with data

In this chapter, these steps are discussed in detail except for the reconstruction of objects that are discussed in Chapter 5.

A set of single-lepton triggers are used for the event pre-selection. Events are required to fire either one of the single-electron or single-muon triggers with different p_{T} thresholds and identification qualities. A lepton must be selected above the listed p_{T} thresholds for the corresponding data collection period and matched to that trigger. The lepton p_{T} thresholds are 25, 27 and 28 GeV for the periods 2015, 2016 and 2017-18 respectively. The different p_{T} threshold values are chosen to account for the different turn-on curves caused by differences in pileup events. A detailed description can be found in [80] and [81] for electron and muon triggers respectively.

6.1 Overlap removal of reconstructed objects

An overlap removal procedure is applied to avoid the same calorimeter energy deposit or the same track being reconstructed as part of two different objects. The procedure is performed in the following order:

- Electrons sharing their track with a muon candidate are removed.
- Jets within a $\Delta R = 0.2$ cone of an electron are removed.
- Electrons within a $\Delta R = 0.4$ cone of a remaining jet are removed.
- When a muon and a jet are close, the jet is removed if it has no more than two associated tracks and is within $\Delta R < 0.2$ of the muon, otherwise the muon is removed if it is within $\Delta R < 0.4$ of the jet and the jet has more than two associated tracks.
- Photons within a $\Delta R = 0.4$ cone of a remaining electron or muon are removed.
- The jets within a $\Delta R = 0.4$ cone of a remaining photon are removed.

6.2 Event selection criteria

The event selection is performed with a set of requirements on the kinematic properties of the reconstructed objects targeting the topology of the signal process. This selection criterion is referred to as the primary event selection and the selected region of phase space is referred to as the signal region in the following. The selection is based on the targeted region of a single-lepton channel from $t\bar{t}$ decays with one additional prompt photon as discussed in Chapter 2. Events from the τ+jets channels are also automatically selected when the τ lepton decays to an electron or a muon. Table 6.1 summarises the event selection. A further classification of regions is performed later for the analysis as explained in Chapter 10.

Table 6.1: Summary of the event selection.

Requirement	e+jets	μ+jets
	Primary Vertex	
1 Lepton	$\lvert\eta\rvert < 1.37$ or $1.52 < \lvert\eta\rvert < 2.47$ $p_{\mathrm{T}} > 25$ GeV Track and calorimeter isolation Criteria on Impact parameters	$\lvert\eta\rvert < 2.5$
1 Photon	$p_{\mathrm{T}} > 20$ GeV $\lvert\eta\rvert < 1.37$ or $1.52 < \lvert\eta\rvert < 2.37$ Tight Identification and Isolation criteria	
Anti-k_t R=0.4 jets	$N_{\mathrm{jets}} \geq 4$ $p_{\mathrm{T}} > 25$ GeV, $\lvert\eta\rvert < 2.5$ JVT cut (if $p_{\mathrm{T}} < 60$ GeV and $\lvert\eta\rvert < 2.4$)	
b-tagging	≥ 1 jets with DL1r at 77%	
m(e,γ) veto	$\lvert$m(e,γ)-91.19 GeV$\rvert$>5 GeV	–

6.3 Normalisation of Monte Carlo simulated events

To compare the MC simulated expected events with the observed data, the MC events are normalised to the data using scale factors, to correct the differences. They are all multiplied for each event and the product is called the *event weight*. They are listed and explained below:

- Monte Carlo weight (weight$_{\mathrm{MC}}$): The weight assigned by the Monte Carlo event generator at the hard-scattering stage. For the MC samples generated at NLO QCD accuracy, a part of the events can have negative weights because of destructive interference.
- Cross-section of the process normalised per event (σ_{norm}), i.e.

$$\frac{\text{Cross-section}}{\text{Total Weighted Events in MC}}$$

- Integrated luminosity of data (L)
- Pileup scale factor (SF$_{\mathrm{pileup}}$): This is used to adjust the pileup (Chapter 3.1) used in the MC.
- Lepton selection efficiency scale factor (SF$_{\mathrm{lepton}}$): The lepton (both e and μ) selection efficiencies are corrected to their data efficiencies with these scale factors. There are several efficiencies for leptons: reconstruction (reco), identification (ID), isolation (Iso), and trigger(trig). The combined scale factor is calculated as SF = SF(reco)*SF(ID)*SF(Iso)*SF(trig).

- Photon selection efficiency scale factor ($\text{SF}_{\text{photon}}$): similar to leptons, the combined scale factor is calculated as SF = SF(ID)*SF(Iso)*SF(trig)
- JVT scale factor (SF_{JVT}): The JVT efficiency is scaled with data using this scale factor as mentioned in Chapter 5.4.
- B-tagging scale factor ($\text{SF}_{\text{b-tag}}$): This corrects the b-jet tagging efficiency with data.

This finally leads to the event weight being

$$\begin{aligned}\text{event weight} = \text{weight}_{\text{MC}} \times \sigma_{\text{norm}} \times \text{L} \times \text{SF}_{\text{pileup}} \times \text{SF}_{\text{lepton}} \\ \times \text{SF}_{\text{photon}} \times \text{SF}_{\text{JVT}} \times \text{SF}_{\text{b-tag}}\end{aligned}$$

CHAPTER 7

Theoretical and experimental systematic uncertainties

Various theoretical and experimental sources of systematic uncertainties affecting both the signal and background predictions are considered for the analysis and discussed in this chapter. A summarised list of all the systematic uncertainties is shown later in Table 11.1 and Table 11.2. Here detailed explanations are given for them.

7.1 Theoretical uncertainties

For the main contributing processes to the analysis, namely $\mathrm{t\bar{t}}\gamma$ production, $\mathrm{t\bar{t}}\gamma$ decay along with $\mathrm{t\bar{t}}$ production, and single top ($\mathrm{tW}\gamma$), alternative generators are tested for the simulation. These alternatives are then used as systematic variations. An overview of the assigned systematic uncertainties is given separately for the signal and background processes in the following.

7.1.1 Modelling of signal process

Renormalisation and factorisation scale

To estimate the uncertainty due to the renormalisation scale (μ_R), the scale value is varied to $\frac{H_T}{4}$ and H_T (Equation 4.1) while keeping the factorisation scale (μ_F) value fixed to the original value $\frac{H_T}{2}$. Similarly, the uncertainty due to the factorisation scale is estimated by the same variation while keeping the renormalisation scale fixed. Simultaneous variations of both μ_R and μ_F scales are not considered.

Parton distribution function

The uncertainty due to parton distribution functions is propagated using the PDF4LHC15 prescription [82], using a set of 30 components. The PDF variations are propagated by using alternative MC generator weights corresponding to the PDF4LHC15 variations. The systematic uncertainty is obtained by comparing the individual PDF variations to the PDF4LHC15 baseline prediction.

Parton shower radiation

An uncertainty on the simulation of the radiation of the parton shower algorithm is evaluated by a variation of the α_S value in the parton shower (PYTHIA 8.240 [48], with the option SPACESHOWER:ALPHASVALUE) of the process.

Parton shower and hadronisation uncertainties

Uncertainty is considered by comparing the parton shower and hadronisation algorithm HERWIG 7.2.1 [83] with PYTHIA. There are several differences between PYTHIA and HERWIG. PYTHIA has k_T-ordered initial and final state DGLAP (From [14] pages 201-202) based shower while HERWIG has angular ordered and dipole shower. PYTHIA uses Lund String ([84]) based hadronisation model and HERWIG uses cluster based hadronisation model. These differences result in low multiplicity jet activity for the HERWIG generated events.

7.1.2 Modelling of background processes

Modelling uncertainties are considered for relevant background processes.

$t\bar{t}$ production with a photon from other sources

The same set of uncertainty sources used for the signal process is also used for the $t\bar{t}\gamma$ decay process. The uncertainties due to the μ_R and μ_F scales, parton shower radiation and hadronisation modelling are also used for the $t\bar{t}$ production process. Additionally, for $t\bar{t}$ production a generator uncertainty is considered by using a different MC generated with MADGRAPH5_AMC@NLO instead of POWHEG-BOX. Also, a variation on the h_{damp} parameter in POWHEG-BOX setup is used as an additional uncertainty for $t\bar{t}$ production.

tWγ process

The scale and parton shower and hadronisation uncertainties are considered for this process similar to the signal.

Background normalisation uncertainty

A normalisation uncertainty of 20% is currently being used for the t$\bar{\text{t}}\gamma$ decay. It is based on its NLO k-factor's uncertainty [85]. 50% uncertainty is used for the processes with small cross-sections, i.e, tWγ, Zγ, Wγ and other smaller processes (t$\bar{\text{t}}V$, diboson and other modes of single top production). The uncertainty value is decided from the knowledge of the Wγ validation region as discussed in Chapter 9.4. For the backgrounds estimated with data-driven methods, the corresponding uncertainties on the SFs are considered separately as mentioned in Section 7.2.5.

7.2 Experimental systematics

The individual sources of uncertainties for all physics objects are explained in brief in the following.

7.2.1 Uncertainties for electron and photon

The electron efficiency scale factors are used as a factor in the expression of the event weight of the MC (as explained in Chapter 6.3). The following uncertainties for each of these scale factors are considered: reconstruction, identification, isolation and trigger. Similarly, for photons, the uncertainties on the identification and isolation efficiency are considered.

Since the electron and photon energy are calibrated using MC, correction factors derived from the study of dileptonic decays of the Z boson, are applied to correct possible detector mis-modelling in the calibration. Separate uncertainties are considered for the calibration of the energy scale and its resolution. An additional energy scale uncertainty is considered for the samples with fast simulation.

7.2.2 Uncertainties for muon

Similar to electrons, uncertainties related to the different muon selection efficiencies are considered [86]: reconstruction efficiency, separate statistical and systematic uncertainties for each of identification, isolation and trigger efficiencies. For low p_{T}

muons, separate statistical and systematic uncertainties are considered for identification efficiency. Additionally, statistical and systematic uncertainty corresponding to the muon track to vertex association is considered.

Two variations on the track resolution of the combined muon track are considered separately for the Inner Detector (*Muon ID*) and the Muon Spectrometer (*Muon MS*). The muon energy scale variation and two energy resolution uncertainties (*Sagitta**) are also considered.

7.2.3 Uncertainties for jets

Several uncertainties considered for jets are described below based on their origin. The uncertainties on the jet energy scale and jet energy resolution are described in detail in [87].

Jet energy scale

The calibration of the jet energy scale (JES) is used to get the energy of the jet corrected to the particle level. The calibration is done in a few steps:

- First, the excess energy from the pileup is removed.
- Then an absolute JES calibration correction is done for the jet to match in energy and direction (η) with the corresponding particle level jet.
- Next, the global sequential calibration [87] improves the jet p_{T} resolution and associated uncertainties by removing the dependence on the tracking, calorimeter, and muon chamber.
- A final calibration is applied to correct for remaining differences between data and MC simulation.

The following uncertainties are considered for the different steps of JES calibration and are arranged sequentially, with the names summarised later in Table 11.2 highlighted.

For the **pileup** removal the μ offset and N_{PV} (number of primary vertex) offset uncertainties are considered due to the μ and N_{PV} modelling respectively in the MC simulation. The uncertainty due to per-event p_{T} density modelling in MC simulation is referred to as the *pileup ρ-topology uncertainty*. Another pileup related uncertainty called *pileup p_{T} term* is considered to account for the residual p_{T} dependence.

At the absolute JES calibration step, a bias is created for the reconstructed jets in terms of their η being shifted from their truth η. So a parameterised second correction step is applied to address this, called **η inter-calibration**. This step introduces 6 uncertainties: a statistical uncertainty, a systematic uncertainty to address all mismodelling and four systematic uncertainties to address different non-closure.

A **flavour** composition uncertainty is considered to address the uncertainty on the proportional sample composition of quark- and gluon-initiated jets. This is estimated from the $t\bar{t}\gamma$ production (signal) process. Jet energy response is calculated by balancing the p_T of a jet against that of a well-calibrated reference object or system. The response is defined as the average ratio of the jet p_T to the reference object p_T bins. The differing responses of quark- and gluon-initiated jets are considered as the flavour response uncertainty. An additional response uncertainty addresses the difference for b-quark-initiated jets called BJES response uncertainty.

A Single Particle **High p_T** uncertainty is derived from studies of the response to individual hadrons to cover the region beyond 2.4 TeV.

A **'Punch Through'** uncertainty accounts for the mis-modelling of jets which pass through the calorimeter into the muon system.

A **non-closure** uncertainty is considered for the fast simulation based samples by comparing their jet response to the corresponding full detector simulation.

The final **calibration** step introduces many independent uncertainties. However, for two jets at different η and p_T, these uncertainties can be partially correlated. An **eigenvector** decomposition is performed on a covariance matrix of these uncertainties to reduce them to 15 meaningful components with a minimum loss of correlations as necessary for the analysis. They are grouped according to their origin: 6 statistical, 4 modelling, 2 detector and 3 mixed components.

Jet energy resolution

A precise measurement of the resolution of the jet energy (JER) is highly important for the reconstruction of top quarks. The JER in MC is calibrated to match the data. Several uncertainties are considered for this step and they are transformed via an eigenvector decomposition method to 7 component uncertainty along with a data-MC offset uncertainty.

B-tagging

The sources of uncertainties considered for the b-tagging scale factors used for weighting events include statistical uncertainties as well as modelling and experimental systematic uncertainties [88]. An eigenvector decomposition method is performed on these uncertainties and they are transformed into 19 component uncertainties: 9 related to b-jets, 4 related to each of charm- and light-jets and 2 related to extrapolations.

Jet vertex tagger

The uncertainty is considered by varying the JVT (Chapter 5.4) cut value up and down.

7.2.4 Uncertainties for missing energy

Since $E_{\mathrm{T}}^{\mathrm{miss}}$ is calculated from other physics objects, the uncertainties due to the other objects are already propagated for $E_{\mathrm{T}}^{\mathrm{miss}}$. However, additional uncertainties [89] are considered due to the calibration of the *soft term* (as defined in the equation 5.4). Two resolution uncertainties are considered corresponding to the parallel or perpendicular direction of the vector sum of the transverse momenta of all the high-p_{T} physics objects. Also, the uncertainty on the energy scale of the soft term is considered.

7.2.5 Other experimental uncertainties

Pile-up uncertainty

The pile-up profile in MC is corrected to match the data by a reweighing factor (Chapter 6.3), a variation of this factor is considered as the pile-up uncertainty.

Uncertainty on the integrated luminosity

A normalisation uncertainty of $\pm 1.7\,\%$ [35] in the integrated luminosity of the combined data (2015–2018) is considered as mentioned in Chapter 3.1.

Uncertainties on the fake object backgrounds

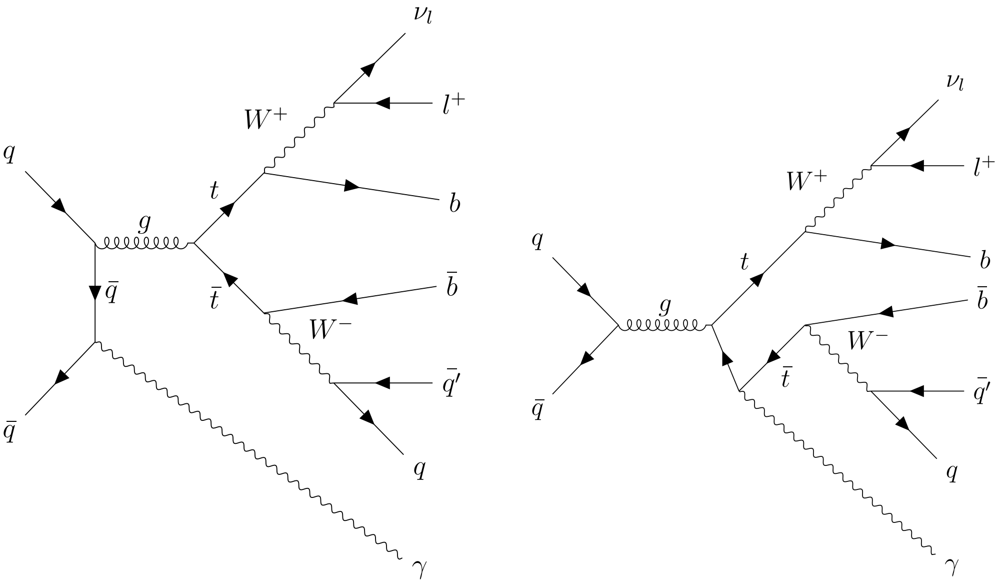
For the electron and hadron fake photon backgrounds (Chapters 9.1 and 9.2), the total uncertainty of their estimated scale factors are considered as systematic uncertainties for the analysis.

For the fake lepton background, two uncertainties on the efficiencies are considered: i) the variation on the parametrisation, ii) a quadratic sum of the other two variations as explained in Chapter 9.3.

CHAPTER 8

Reconstruction of top quark pairs

For the measurement of any observable involving top quark pairs like the charge asymmetry (A_C), it is necessary to reconstruct the four-momenta of the top quarks from their decay products.



(a) (b)

Figure 8.1: A leading-order example Feynman diagrams of $t\bar{t}\gamma$ production as (a) initial quark radiation and (b) final state radiation where the top quark pair decays one charged lepton, one photon, and four quarks among which two should be b-quarks

In a $t\bar{t}\gamma$ production event, as shown in Figure 8.1, the top quark pairs decay and the photon is not emitted as part of any of the top quark decays. In other words, the photon is not part of the $t\bar{t}$ system and so it is not needed to be associated with the top quark reconstruction unlike in the case of $t\bar{t}\gamma$ decay process. To evaluate the kinematic variables of the top quarks, the $t\bar{t}$ system is reconstructed using a kinematic likelihood fit implemented with *KLFitter* [90], which assesses the compatibility of the observed event with the decay of a $t\bar{t}$ pair.

8.1 Kinematic likelihood fit

For reconstructing the 4-vectors of the two top quarks from their decay products of a semi-leptonic decay, a likelihood is defined based on the decay topology of the signal process in a leading order model signature (like Figure 8.1). The four momenta of the reconstructed objects are used as input to the likelihood.

The following assumptions are considered for the likelihood construction:

- The constraints on the top quark or W boson mass from the invariant mass (m_{inv}) of their decay final state particles[1] are used in terms of a *Breit-Wigner* function parametrised by the corresponding peak mass m and decay width Γ as:
$$B(m_{inv}|m,\Gamma) = \frac{\text{constant}}{(m_{inv}^2 - m^2)^2 + (m\Gamma)^2} \tag{8.1}$$
- The directions of the particle momenta are assumed to be well measured.
- Out of the 24 four-momentum parameters of the 6 objects [leading 4 jets (in p_T), 1 charged lepton and the neutrino], only the 4 jet energies, the charged lepton energy and the three-momentum of the neutrinos are the free parameters in the fit.
- The energies of the reconstructed particles (E^{meas}) are corrected to their corresponding final state particle energies (E) using the *transfer functions*, F^t. Each transfer function is determined by using dedicated fits for the corresponding object using data. Transfer functions for each object are multiplied in the likelihood. For the neutrino x- and y-components of the momentum (denoted as p_x^ν and p_y^ν), transfer functions for the x- and y-components of the missing energy are used.
- The selected jets are considered in different permutations (association of objects to final state particles) to build the likelihood. The highest likelihood value is assumed to be from the permutation best fitting the assumed description of the decay topology of the signal process and any other permutation with a wrong assignment will lead to a lower value of the likelihood.

---

[1] parton level final state particles: four quarks q_i, lepton ℓ and neutrino ν

- The charm jets are considered as light jets for simplicity

Hence, the likelihood function is constructed as:

$$\begin{aligned} L = {} & B(m_{q_1q_2q_3}|m_t,\Gamma_t) \times B(m_{q_1q_2}|m_W,\Gamma_W) \times B(m_{q_4\ell\nu}|m_t,\Gamma_t) \times B(m_{\ell\nu}|m_W,\Gamma_W) \\ & \times \prod_{i=1}^{4} F^t_{jet}(E^{meas}_{jet,i}|E_{jet,i}) \times F^t_{\ell}(E^{meas}_{\ell}|E_{\ell}) \times F^t_{miss}(E^{miss}_{x}|p^{\nu}_{x}) \times F^t_{miss}(E^{miss}_{y}|p^{\nu}_{y}), \end{aligned} \tag{8.2}$$

where

- $m_{q_1q_2q_3}$, $m_{q_1q_2}$, $m_{q_4\ell\nu}$, $m_{\ell\nu}$ are the invariant masses of the different combination of final state particles as top quark or W boson decay products
- m_t and m_W are the pole masses of the top quark and W boson respectively
- Γ_t and Γ_W are the decay width of the top quark and W boson respectively, where the values are the same as used in the signal event generator.

A maximum of 5 jets is considered for all possible permutations. For the events with more than 5 jets, only the 5 jets with the highest transverse momentum are chosen for the permutations. After calculating the likelihood L for all possible permutations, each of the permutation **i** is given an event probability, P_{i}. They are normalised by the sum over the likelihood values of all permutations and weighted with the b-tagging factors, f_{b-tag} as:

$$P_{\mathrm{i}} = \frac{L_i \times f^i_{b-tag}}{\sum_j L_j \times f^j_{b-tag}} \tag{8.3}$$

The b-tagging factor, f_{b-tag} is determined from the b-tagging efficiency, ϵ_b and the light jet mis-identification rate, ϵ_l as follows:

$$\begin{aligned} f_{\text{b-tag}} = {} & \begin{Bmatrix} \epsilon_b, & b_{had} \text{ is b-tagged} \\ (1-\epsilon_b), & b_{had} \text{ is not b-tagged} \end{Bmatrix} \times \begin{Bmatrix} \epsilon_b, & b_{lep} \text{ is b-tagged} \\ (1-\epsilon_b), & b_{lep} \text{ is not b-tagged} \end{Bmatrix} \\ & \times \begin{Bmatrix} \epsilon_l, & q_1 \text{ is b-tagged} \\ (1-\epsilon_l), & q_1 \text{ is not b-tagged} \end{Bmatrix} \times \begin{Bmatrix} \epsilon_l, & q_2 \text{ is b-tagged} \\ (1-\epsilon_l), & q_2 \text{ is not b-tagged,} \end{Bmatrix} \end{aligned} \tag{8.4}$$

where

- ϵ_b and ϵ_l have the values 77% and 1/163 respectively from the chosen working point of the b-tagging algorithm,
- b_{lep} - the jet assigned as the b-jet decayed from the *leptonic* top quark,
- b_{had} - the jet assigned as the b-jet decayed from the *hadronic* top quark,
- q_1 and q_2 - the jets assigned as the light quark jets decayed from the *hadronic* top quark

The permutation with the highest event probability P_i is chosen for event reconstruction. The assigned jets of the chosen permutation are used to get the corresponding momenta directly. The hadronic top quark is reconstructed from the two jets assigned as the light quark jets (q_1 and q_2) and the jet assigned as the hadronic b-jet (b_{had}) in the chosen permutation of *KLFitter*. For the leptonic top quark, the jet assigned as the leptonic b-jet (b_{lep}) and the reconstructed lepton momentum are used but lack the complete four-momentum of the neutrino. Only the x- and y-components of the missing energy are known which are assumed to correspond to the sole expected neutrino in the event, given that the neutrino emission from other sources like hadrons is neglected. The z-component of the neutrino momentum, P_z^ν remains unknown. To determine the suitable P_z^ν value additional constraints are used based on the decay kinematics. Three different methods are used for this. Their performance is described in the following.

8.2 Solutions of P_z^ν

Considering the W boson decaying to the charged lepton and the neutrino to be on-shell, the invariant mass of the lepton-neutrino system can be constrained to be equal to the pole mass of the W boson, $m_W = 80.4$ GeV. Using the P_z^ν with this constraint leads to a quadratic equation in terms of P_z^ν:

$$A \times \left(P_z^\nu\right)^2 + B \times P_z^\nu + C = 0\,, \tag{8.5}$$

$$\begin{aligned} \text{where } A &= \left(P_z^\ell\right)^2 - E_\ell^2\,, \\ B &= 2 \times P_z^\ell \times \mu\,, \\ \text{and } C &= \mu^2 - E_\ell^2\left(P_T^\nu\right)^2 \\ \text{with } \mu &= \frac{m_W^2}{2} + p_T^\ell \times p_T^\nu \times cos(\Delta\Phi_{\ell\nu}) \end{aligned}$$

E_ℓ, p_T^ℓ and p_z^ℓ are the energy, transverse component and the z-component of the charged lepton momentum respectively, p_T^ν is the transverse momentum of the neutrino, $\Delta\Phi_{\ell\nu}$ is the difference in the azimuthal angle of the charged lepton and the neutrino.

The equation gives solutions for the P_z^ν as:

$$P_z^\nu = \frac{-B \pm \sqrt{D}}{2A}\,, \tag{8.6}$$

where

$$D = B^2 - 4AC = 4E_\ell^2 \times \left[\mu^2 - \left(P_T^\ell\right)^2 \left(P_T^\nu\right)^2\right] \tag{8.7}$$

This implies that the equation provides real solutions for P_z^ν only when D$\geq$ 0. That further implies:

$$\mu \geq p_T^\ell p_T^\nu\,, \tag{8.8}$$

$$\implies m_W^2 \geq 2p_T^\ell p_T^\nu \left(1 - cos(\Delta\Phi_{\ell\nu})\right)\,, \tag{8.9}$$

$$\implies m_T^{\ell\nu} \leq m_W \tag{8.10}$$

The transverse mass of the charged lepton-neutrino system ($m_T^{\ell\nu}$) can be defined as:

$$m_T^{\ell\nu} = \sqrt{2p_T^\ell p_T^\nu \left(1 - cos(\Delta\Phi_{\ell\nu})\right)} \tag{8.11}$$

On the other hand, the solution becomes imaginary when $m_T^{\ell\nu} \geq m_W$.

In the case of real solutions, since there are two solutions, the one among them that gives the reconstructed leptonic top quark mass closest to the top quark mass of 172.5 GeV is chosen. In the case of an imaginary solution, three different methods can be considered to get a suitable solution for P_z^ν, described in the following:

8.2.1 Method 1

A simple solution is achieved by considering the discriminant D to be zero. This does not satisfy Equation 8.5 and can give a slightly incorrect value for the reconstructed W-boson and top quark masses when compared to their corresponding pole masses.

8.2.2 Method 2

Another method would be by changing the $m_T^{l\nu}$ value such that it does not exceed m_W. In Equation 8.11, all the arguments on the right-hand side are measured quantities and hence, constants for the given event. However, the lepton is measured with much better precision than the missing energy components. If it can be assumed that the true azimuthal angle of the neutrino and hence $\Delta\Phi_{l\nu}$ does not exactly correspond to the angle calculated from missing energy transverse components, then the $\Delta\Phi_{l\nu}$ can be modified the least such that $m_T^{l\nu}$ decreases to m_W to obtain a real solution of P_z^ν. The discriminant D would become zero in this case. From Equation 8.6, it can be easily derived that the solution of P_z^ν would be $\frac{-B}{2A}$, and can be simplified from Equation 8.11, to be:

$$p_z^\nu = \frac{p_z^l}{p_T^l} \times p_T^\nu \tag{8.12}$$

The $\Delta\Phi_{l\nu}$ value satisfying this solution would be $cos^{-1}\left(1 - \frac{m_W^2}{2p_T^l p_T^\nu}\right)$. The new solution would satisfy Equation 8.5, but it also changes the x- and y-components of the neutrino momentum while not modifying p_T^ν. However, this method also does not ensure a good resolution of the reconstructed top quark and W-boson four-momenta while modifying $\Delta\Phi_{l\nu}$ as can be seen in the performance studies in Section 8.3.

8.2.3 Method 3

Similar to method 2, if it is assumed that both the measured observables $\Phi(E_{\mathrm{T}}^{\mathrm{miss}})$ and $E_{\mathrm{T}}^{\mathrm{miss}}$ do not exactly correspond to the azimuthal angle and the transverse mass of the true neutrino respectively, then $m_T^{l\nu}$ can be modified with more freedom following Equation 8.11. It is to note that the modification of the $\Phi(E_{\mathrm{T}}^{\mathrm{miss}})$ and $E_{\mathrm{T}}^{\mathrm{miss}}$ to obtain true Φ^ν and p_T^ν, is not just an unjustified fix to avoid imaginary solutions, but in fact accounts to some extent for the sources of missing energy other than the neutrino from W-boson decay. For example, the hadrons as part of the jets can also emit neutrinos that are not accounted for as being part of the jets in the jet reconstruction algorithms. However, it is not possible to assign the missing energy to all possible sources systematically. A safe assumption similar to Method 2 can be followed, i.e. minimal changes to the missing transverse components required such that $m_T^{l\nu}$ equals m_W to obtain a real solution p_z^ν.

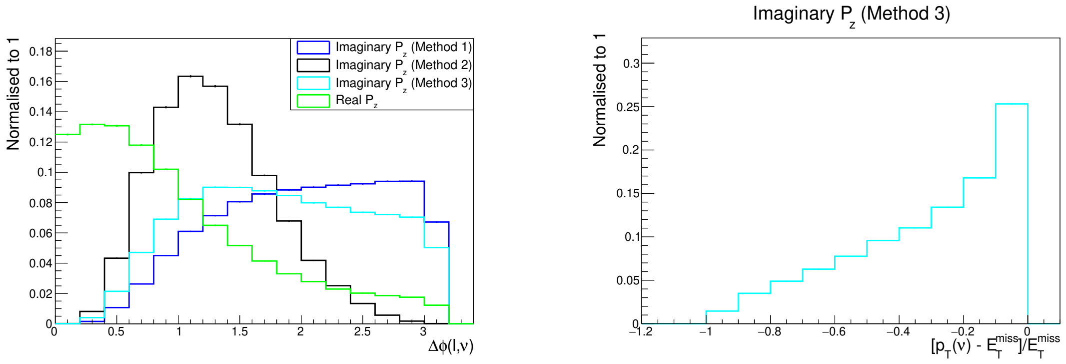
In this method [91], both p_T^ν and $\Delta\Phi_{l\nu}$ are varied in Equation 8.11 to get $m_T^{l\nu}$ to be equal to m_W. Since the two variables are related, a fit is done on $\Delta\Phi_{l\nu}$ by varying it while iteratively modifying p_T^ν by the ratio of $\frac{m_W^2}{m_T^{l\nu}}$ (function of $\Delta\Phi_{l\nu}$), so that the difference between the modified x- and y-components of the neutrino momentum and their given missing transverse momentum Δ is minimum:

$$\Delta = \left(p_x^{\nu'} - p_x^{\nu}\right)^2 + \left(p_y^{\nu'} - p_y^{\nu}\right)^2 \tag{8.13}$$

As a result, this method gives the solution of p_z^ν to be $\frac{-B}{2A}$ by construction from Equation 8.6. However, since this solution depends on both p_T^ν and $\Delta\Phi_{l\nu}$, it differs from the other two methods. Also, it does not only change the z-component of the neutrino momentum but also the x- and y-components and, thus, both p_T^ν and Φ^ν.

A comparison of $\Delta\Phi_{l\nu}$ for the cases of imaginary solutions of the three methods along with the cases of real solutions is shown in Figure 8.2 (a). For real solutions the angular difference peaks at small values. This makes sense given that large $\Delta\Phi_{l\nu}$ value would result in large $m_T^{l\nu}$ that can exceed m_W and yield imaginary solutions for p_T^ν. $\Delta\Phi_{l\nu}$ for Method 1 corresponds to the measured missing energy values as no modification is done for transverse components of the neutrino. So, the deviations shown for Method 2 and 3 compared to Method 1 can be explained as these two methods reduce the $\Delta\Phi_{l\nu}$ value to obtain a lower value of $m_T^{l\nu}$ and to be equal to m_W. Since the aim of Method 2 is to reduce this variable to obtain a real solution, the variable gets peaked at a much lower value. Whereas, since Method 3 modifies this along with the p_T^ν, the overall change in $\Delta\Phi_{l\nu}$ with respect to Method 1 is not large.

For Method 3 (for imaginary solutions only), since the p_T^ν value is also modified with respect to its original value, i.e. $E_{\mathrm{T}}^{\mathrm{miss}}$, the fractional change is shown in Figure 8.2 (b). This shows a reduction of p_T^ν with respect to $E_{\mathrm{T}}^{\mathrm{miss}}$. This works as per expectation since p_T^ν is supposed to be less than $E_{\mathrm{T}}^{\mathrm{miss}}$ due to other possible missing energy sources as explained before.



(a)

(b)

Figure 8.2: (a) $\Delta\Phi_{l\nu}$ obtained with the three methods for the case of imaginary solutions of p_z^ν, compared with the case of real solutions. (b) The relative difference of the fitted p_T^ν from its original value, i.e. $E_{\mathrm{T}}^{\mathrm{miss}}$ for Method 3. For Method 1 and 2 p_T^ν does not change from $E_{\mathrm{T}}^{\mathrm{miss}}$.

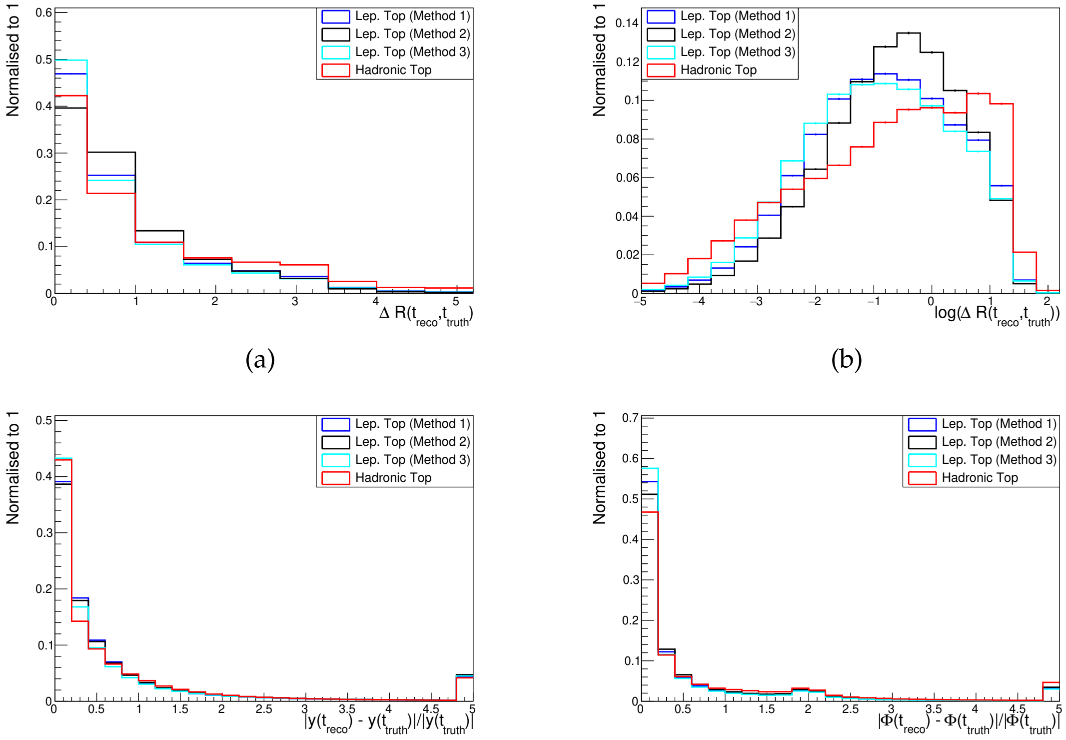
8.3 Performance

The performance of the top quark reconstruction algorithms is evaluated by comparing the kinematics of the reconstructed top quarks with the corresponding truth top quark, i.e. the *parton* level top quark after radiation and before decay.

For this study, several variables are chosen. In Figure 8.3, the comparison is done for the angular variables: (a) ΔR and (b) its logarithm, (c) relative difference of rapidity and (d) the azimuthal angle, Φ between the truth and reconstructed top quarks. It is quite evident that the leptonic top quarks are better reconstructed than the hadronic top quarks. In the case of rapidity, the performance is close to each other, considering that the P_z^ν is not directly measured. It can also be pointed out that Method 3 performs the best for leptonic top quarks, not just for the rapidity but also for the azimuthal angle. Method 2 performs worse than Method 1.

The reconstructed top quark masses are compared between the different methods for leptonic and hadronic top quarks in Figure 8.4. Although for all the methods the leptonic top quark and the hadronic top quark peaks near the top quark mass (172.5 GeV), their resolutions differ from each other. Method 3 gives the best resolution for the leptonic top quarks, while the hadronic top quarks have the worst resolution. Method 2 also does not perform well.

In Figure 8.5, the relative difference of transverse momentum (p_T) of the reconstructed and truth top quarks is shown. The relative difference of transverse momentum is also checked for different ranges of the distribution of transverse momentum as shown in the 2D plots in Figure 8.6. Method 3 performs the best for leptonic top quarks, as evident also from the other variables in the transverse direction. Hadronic top quarks have the worst performance similar to other variables.



(a) (b)

(c) (d)

Figure 8.3: (a) Comparison of ΔR between the leptonic top quark reconstructed with the three methods (blue, black and cyan) and the truth top quark. The hadronic top quark is also compared (red) (b) Natural logarithm of the same variable. The absolute relative difference of the reconstructed and truth top quarks as a function of their (c) Rapidity and (d) Azimuthal angle

The resolution of the hadronic top quarks distribution is worse than the leptonic top quark as can be seen for other variables. From the 2D plots, it can be understood that the majority of the small relative difference comes from the p_T^t range of 100-200 GeV for all the cases. The main difference between them is visible in the broadness of the distribution in the 2D plots along the p_T^t axis. For example, hadronic top distribution is more widely spread than its leptonic top counterparts.

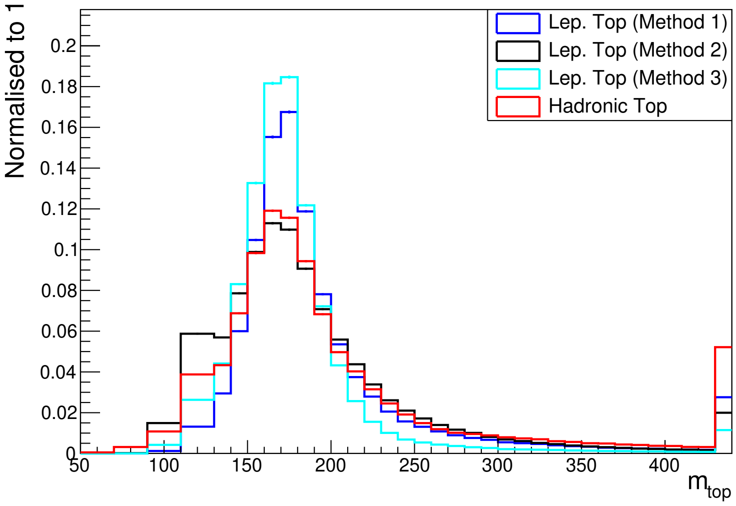
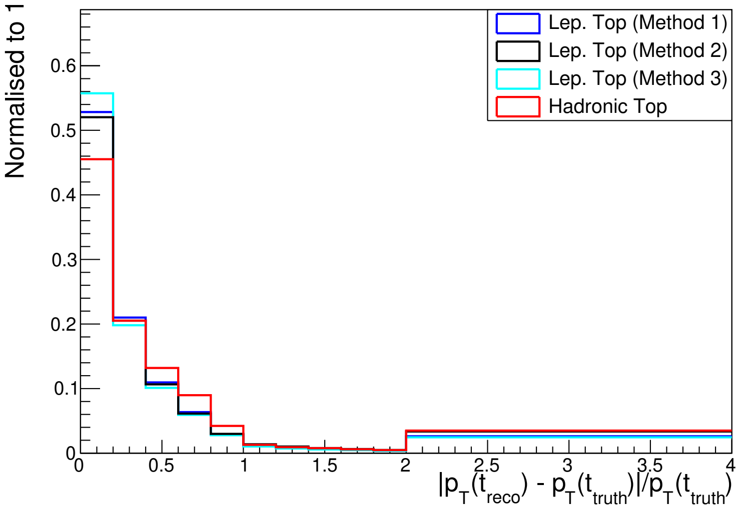
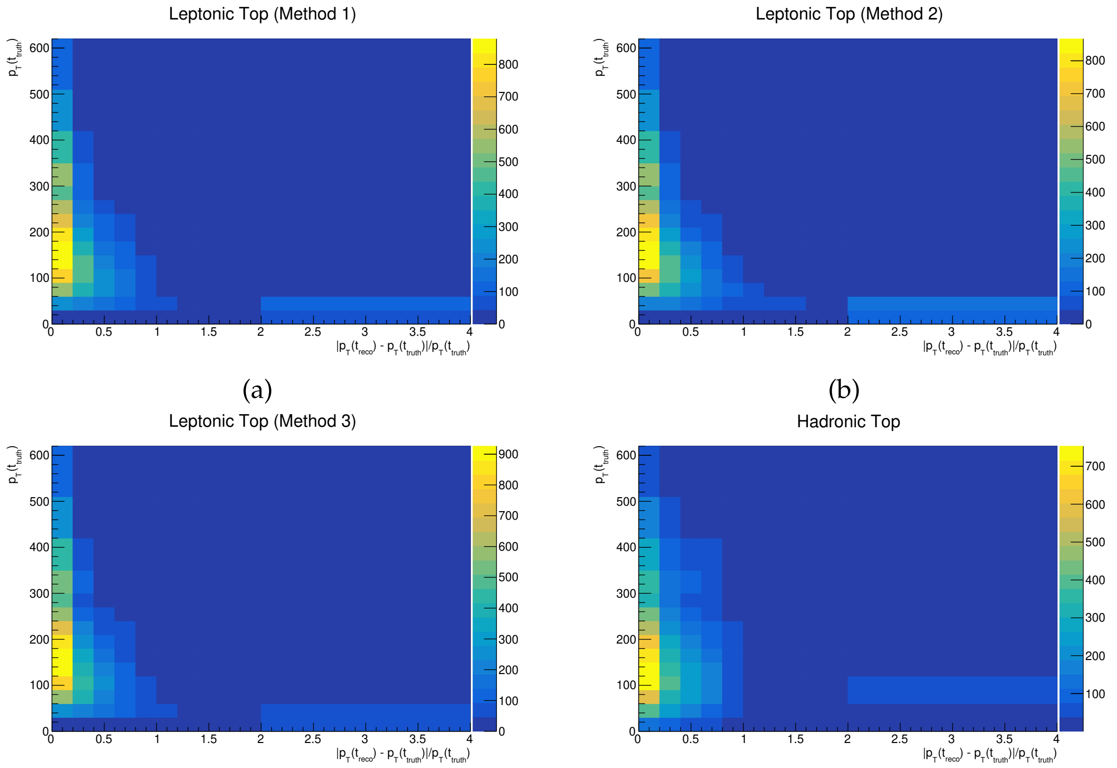


Figure 8.4: The mass of the reconstructed leptonic top quark obtained with the three different methods and compared with the mass of the reconstructed hadronic top quark



(a)

Figure 8.5: The absolute relative difference between the reconstructed and truth top quarks as a function of their transverse momentum p_T. The last bin has different width and contains the overflow.



(a) (b) (c) (d)

Figure 8.6: Absolute relative difference between the p_T of reconstructed and truth top quark Vs the p_T of the truth top quark (a) Method 1 for the leptonic top quark (b) Method 2 for the leptonic top quark (c) Method 3 for the leptonic top quark (d) the hadronic top quark

8.4 Limitations

The above discussed methods for the top quark reconstruction show good performance overall. Some limitations, difficult to avoid, may be considered for future improvements:

- Since a leading order model signature is considered to build the likelihood and also to combine jets to reconstruct the top quarks, the events with more than four jets are always assumed to have the excess jet originating from the initial state gluon radiations, so that they are not supposed to be part of the kinematic reconstruction. However, that is not necessarily true as more than one jet can be formed by a final state quark and hence can modify the kinematic relations
- Even though the reconstructed jets are found to be very close (small ΔR) to their corresponding *particle* level jets (except in cases of wrong flavour tagging), they are often found shifted in their $\eta - \phi$ plane from their corresponding parton

level quarks

- The transfer functions are only considered for the energies of the objects, but not for any of the angular observables

8.5 Conclusion

The reconstruction of the top quarks is a necessary and crucial step for the measurement presented in this analysis. It is evident from the performance studies that Method 3 is the most suitable approach. However, performing the measurement (Chapter 11) separately based on the Method 1 and Method 3 does not show any significant difference. Hence, because of simplicity and convenience Method 1 is rather chosen for the final results.

CHAPTER 9

Estimation of background processes

There are several backgrounds considered for the analysis and they are used in groups as described in Table 4.1. To improve their physics modelling they are treated with dedicated methods using data either for correction or validation.

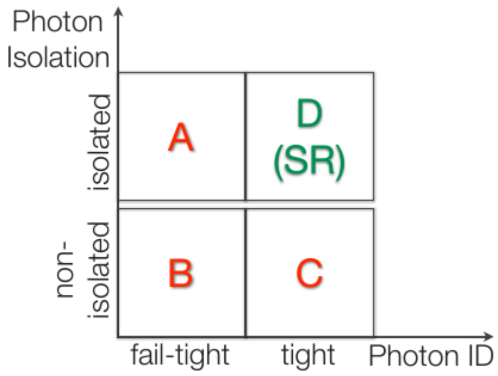
As explained in Section 4.4, selected events in MC can include a photon object that is a fake photon as determined by its MC truth information. However, this estimation may not be very accurate. So, correction factors are derived using data that can be applied to the MC-based estimations. They are described in Sections 9.1 and 9.2. The fake leptons are also estimated from data as described in Section 9.3.

Additionally, all of the backgrounds containing prompt photons also may need some corrections either because the detector simulation is not perfect or their MC prediction is missing higher-order corrections. For example, the prompt photon background modelling is validated with data as described in Section 9.4. Since the $t\bar{t}\gamma$ decay modelling is at the LO (QCD) accuracy (except the small contribution from $t\bar{t}$ sample), a re-weighting procedure is performed on the reconstructed top quark templates to include NLO QCD effects and it is described in Section 9.5.

Disclaimer: The Sections 9.1, 9.2 and partially 9.3 are not contributions of the author.

9.1 Hadron fake photon background

Fake photon background events as defined in Section 4.4 are denoted as *h-fake* in the following. Such contributions are mainly expected from the (fake) photon-associated production of $t\bar{t}$ (∼90%). A 2-dimensional sideband method, also known as *ABCD* method is used for the estimation of the hadron fake photon background in the signal region. Four orthogonal regions A, B, C and D are defined based on photon identification and isolation criteria to get different enrichment of hadron fake photon events. They are illustrated in Figure 9.1.



(a)

Figure 9.1: An illustration of the ABCD method to estimate the h-fake photon background, using four categories of photons based on their identification and isolation criteria.

The selection criteria of the photon in the four regions (A, B, C and D) are described. The rest of the event selection corresponds to those described in Chapter 6.2:

- **Region A:** The Tight identification working point should fail while also failing the requirements of at least two shower shape variables of the front layer of the EM calorimeter. These variables are effective for discriminating hadron fake photons with broader shower profiles. The Tight isolation working point should be satisfied.

- **Region C:** The Tight photon isolation working point (Table 5.2) should be failed and the Loose working point should be satisfied. The Tight identification working point should be satisfied.

- **Region B:** Both of the failing conditions on the photon identification in region A and on photon isolation in region C needs to be satisfied.

- **Region D:** Same as the signal region.

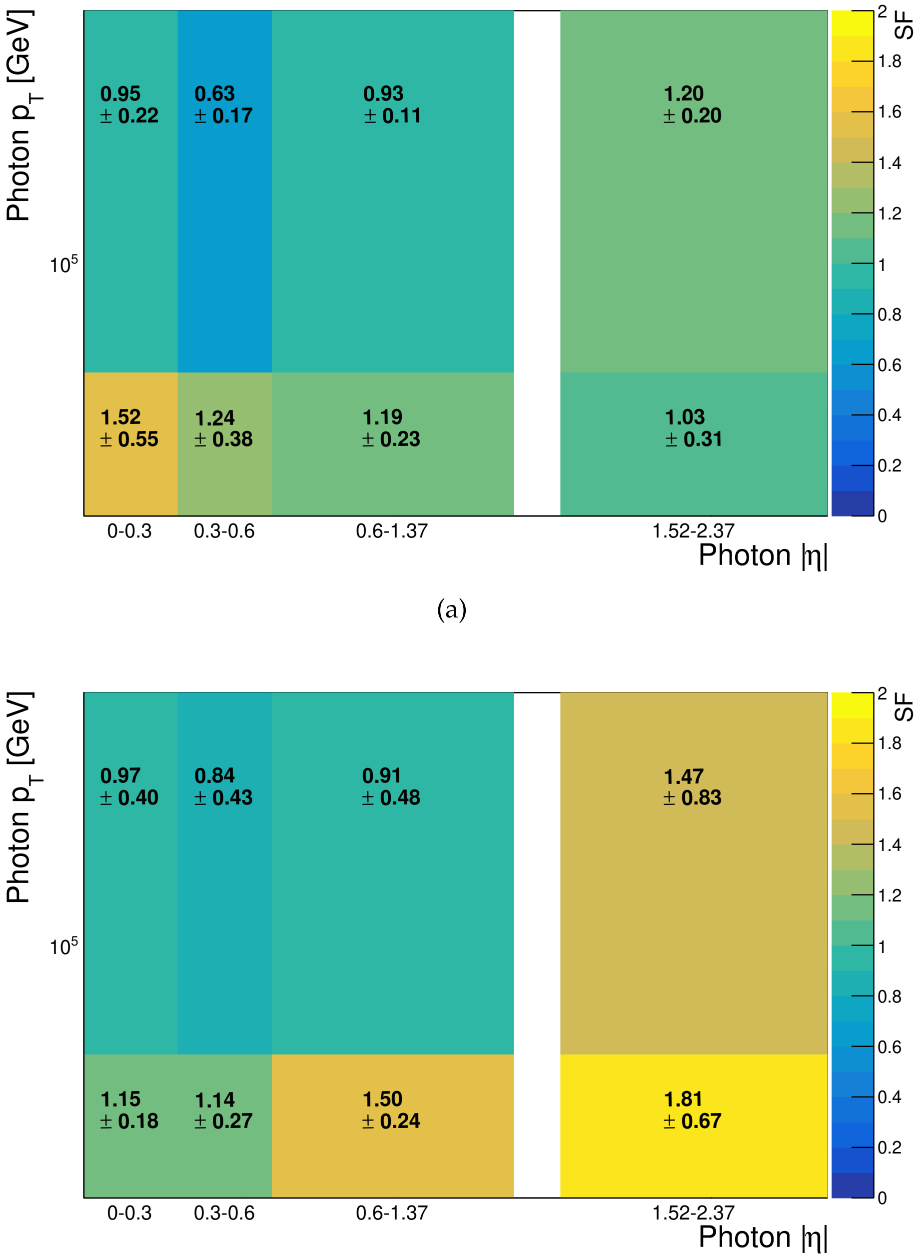
Assuming that the correlation between the photon identification and isolation criteria is negligible for hadron fake contributions in the four regions, and the number of h-fake events in region D being $N_{MC}^{\text{h-fake}}$ in MC and $N_{data}^{\text{h-fake}}$ in data, their scale factors, $SF_{\text{h-fake}}$ can be written as:

$$SF_{\text{h-fake}} = \frac{N_{data,D}^{\text{h-fake}}}{N_{MC,D}^{\text{h-fake}}} = \frac{SF_{\text{h-fake}}^{A} \times SF_{\text{h-fake}}^{C}}{SF_{\text{h-fake}}^{B}} \tag{9.1}$$

where the hadron fake estimations from data, $N_{data}^{\text{h-fake}}$ in A, B and C regions are calculated by subtracting the signal and other backgrounds. The scale factors are derived for different bins of p_{T}, $|\eta|$ and the conversion type of the photon.

Several uncertainties are considered for the scale factors. Since the hadron fake contributions are mainly due to fake photons from $\mathrm{t\bar{t}}$ events, the modelling systematic uncertainties of $\mathrm{t\bar{t}}$ process (as mentioned in Chapter 7) are the major sources of uncertainties. The normalisation uncertainty of other background processes is also considered. Effects of the experimental systematic uncertainties are negligible as they cancel out for this ratio measurement.

The resulting scale factors in the bins of photon kinematics are shown in Figure 9.2 where the total uncertainties are calculated as the quadratic sum of all the systematic and statistical uncertainties.



(a)

(b)

Figure 9.2: Scale factors for (a) converted and (b) unconverted photons as a function of their (p_{T}, $|\eta|$). The uncertainty corresponds to the total uncertainty [92].

9.2 Electron fake photon background

These fake photon background events as defined in Section 4.4 are denoted as *e-fake* γ in the following. The main contributions of such fake photon events are expected in the associated production of the processes: $t\bar{t}$ (~84%) and Z+jets (~9%).

A *tag-and-probe* [72] method is used to estimate such fake photon events both in data and MC, considering the events in $Z \to ee$ process where an electron is misidentified as a photon. Two control regions, referred to as ee and $e\gamma$ in the following, are used that are orthogonal to the signal region.

- **$e\gamma$ region**: The selection criteria for this region is one photon with $p_T > 20$ GeV (that can be either a fake photon or prompt photon from $Z \to ee\gamma$ decay) considered as the probe particle and one trigger matched electron[1] with $p_T > 25$ GeV considered as the tag particle. Further, both the electron and photon should pass tight identification and isolation requirements, they should be back-to-back with an opening angle between them $\Delta\phi(e,\gamma) > 150°$, and their invariant mass should be within 50 GeV around the Z-boson mass of 91.19 GeV.

- **ee region**: The selection criteria for the tagged electron is very same, and a probe electron is considered with similar conditions for the probe photon in $e\gamma$ region. In case both the electrons are triggered in an event, the tag and probe electrons are chosen randomly.

The probability of an electron being misidentified as a photon is calculated by comparing the events in the two regions. The rate of an electron being misidentified as a fake photon, i.e., the *fake rate* is calculated from the MC in the two regions in terms of the probe objects. The probe photon and probe electron in the $e\gamma$ and ee regions, respectively, are compared for the same values of their p_T and $|\eta|$. They are also separately calculated for converted and unconverted probe photons. However, since there are also events from the prompt photon background processes in these regions, this contribution is subtracted before calculating the fake rate. The fake rate for a single bin i can be expressed as:

$$FR(\mathrm{MC})_{\mathrm{bin\ i}} = \frac{\text{Number of fake photon originated events in } e\gamma \text{ region, bin i}}{\text{Number of events in } ee \text{ region, bin i}} \tag{9.2}$$

The fake rates are also evaluated using data. For this, the $Z \to ee$ process is modelled using a double-sided crystal ball function. The contribution of events from background processes that do not originate from a Z boson is modelled using a Bernstein 4th-order polynomial function. The prompt photon backgrounds are

---

[1] reconstructed electron matched with a trigger electron

modelled using MC. In both regions, the signal and background predictions are fitted with the data using the invariant mass of the tag and probe objects in the range of 60-120 GeV. Afterwards, the fake rate is calculated by subtracting the backgrounds from the data and comparing them for any bin i as:

$$FR(\text{data})_i = \frac{[N(\text{data}) - N_{\text{fitted}}(\text{non-Z}) - N_{\text{MC}}(\text{prompt }\gamma)]_{e\gamma,i}}{[N(\text{data}) - N_{\text{fitted}}(\text{non-Z}) - N_{\text{MC}}(\text{prompt }\gamma)]_{ee,i}}, \tag{9.3}$$

where N corresponds to the number of events in the corresponding region and bin. The scale factor of the fake rate in bin i is calculated as the ratio between the fake rates in data and MC:

$$SF_i = \frac{FR(\text{data})_i}{FR(\text{MC})_i} \tag{9.4}$$

The scale factors are parameterised in 3 bins of p_{T} and 4 bins of $|\eta|$ of the photon and also separately for converted and unconverted photons. Several uncertainties are considered for the scale factor calculation, related to the fake rate estimation from data. The difference in results with different fit setups is considered as separate uncertainties by:

- using a MC generated distribution for the $Z \to ee$ process instead of using the crystal ball function
- using a Gaussian function instead of a Bernstein function for the non-Z backgrounds
- Changing the fit mass range to 65-115 GeV

The quadratic sum of all the uncertainties including the statistical uncertainty on the scale factor is considered the total uncertainty.

9.3 Fake lepton background

There are non-prompt or misidentified leptons (electron or muon) that can come from the decay of a heavy flavour hadron (b- or c-hadron), light flavour hadron decays faking the electron shower, decays of pions/kaons to muons etc. They are known as *fake leptons*. Such leptons are expected mainly from multi-jet processes. The prompt and isolated leptons are considered as *real leptons*. Since it is difficult to model multi-jet processes, a control region enriched with fake lepton is used to extrapolate the fake lepton contribution in a region with tighter lepton selection (as in the signal region) using the *Matrix Method*. For electrons, a looser definition of identification and isolation is used, while for muons same identification criteria are used with no isolation to define the loose control region. The tight control region with the same selection except with tight leptons is also used. The control regions[2] contain events with:

- lepton triggers and object pre-selection same as the signal region
- 1 loose/tight lepton
- at least 4 jets, with 1 jet b-tagged with 77% working point
- $E_{\mathrm{T}}^{\mathrm{miss}} < 30$ GeV

In MC samples the real lepton can be identified from the truth information. The number of events in the two regions with real leptons gives the efficiency of a real lepton in a loose lepton region passing the tight lepton selection (*real efficiency*):

$$\epsilon_{\mathrm{real}} = \frac{N_{\mathrm{real}}^{\mathrm{tight}}}{N_{\mathrm{real}}^{\mathrm{loose}}} \tag{9.5}$$

To calculate the efficiency of a fake lepton in a loose region passing the tight lepton selection (*fake efficiency*), all MC events with real leptons in the corresponding region (same as before) are subtracted from data:

$$\epsilon_{\mathrm{fake}} = \frac{N_{\mathrm{fake}}^{\mathrm{tight}}}{N_{\mathrm{fake}}^{\mathrm{loose}}} = \frac{N_{\mathrm{data}}^{\mathrm{tight}} - N_{\mathrm{real}}^{\mathrm{tight}}}{N_{\mathrm{data}}^{\mathrm{loose}} - N_{\mathrm{real}}^{\mathrm{loose}}} \tag{9.6}$$

The efficiencies are calculated from these control regions by parameterising in terms of the bins of $p_T(\ell) : m_T(\ell\nu)$.

To estimate the fake lepton events in the data of the signal region, another control region (SRloose) is defined with the same selection as the signal region except with a loose definition of the leptons as above. The efficiencies values are assumed to be valid in these regions:

$$\begin{aligned} N_{\mathrm{real}}^{\mathrm{SR}} &= \epsilon_{\mathrm{real}} \times N_{\mathrm{real}}^{\mathrm{SRloose}} \\ N_{\mathrm{fake}}^{\mathrm{SR}} &= \epsilon_{\mathrm{fake}} \times N_{\mathrm{fake}}^{\mathrm{SRloose}} \end{aligned} \tag{9.7}$$

The events in the SRloose and SR region can be divided into real and fake contri-

---

[2] Needless to say that the tight control region is a subset of the loose control region

butions:

$$\begin{aligned} N^{\text{SRloose}} &= N^{\text{SRloose}}_{\text{real}} + N^{\text{SRloose}}_{\text{fake}} \\ N^{\text{SR}} &= N^{\text{SR}}_{\text{real}} + N^{\text{SR}}_{\text{fake}} \\ &= \epsilon_{\text{real}} N^{\text{SRloose}}_{\text{real}} + \epsilon_{\text{fake}} N^{\text{SRloose}}_{\text{fake}}, \end{aligned} \tag{9.8}$$

they can be written in a matrix format as:

$$\begin{pmatrix} N^{\text{SRloose}} \\ N^{\text{SR}} \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ \epsilon_{\text{real}} & \epsilon_{\text{fake}} \end{pmatrix} \begin{pmatrix} N^{\text{SRloose}}_{\text{real}} \\ N^{\text{SRloose}}_{\text{fake}} \end{pmatrix} \tag{9.9}$$

By inverting the matrix the fake and real lepton events in the SRloose region can be solved:

$$\begin{pmatrix} N^{\text{SRloose}}_{\text{real}} \\ N^{\text{SRloose}}_{\text{fake}} \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ \epsilon_{\text{real}} & \epsilon_{\text{fake}} \end{pmatrix}^{-1} \begin{pmatrix} N^{\text{SRloose}} \\ N^{\text{SR}} \end{pmatrix} \tag{9.10}$$

By simplifying Eq. 9.10, the estimation of fake lepton events in the signal region can be estimated in terms of the efficiencies (ϵ_{real} and ϵ_{fake}) and the total data in the two regions (N^{SRloose} and N^{SR}) as:

$$N^{\text{SR}}_{\text{fake}} = \epsilon_{\text{fake}} N^{\text{SRloose}}_{\text{fake}} = \frac{\epsilon_{\text{fake}}}{\epsilon_{\text{real}} - \epsilon_{\text{fake}}} \times \left(\epsilon_{\text{real}} N^{\text{SRloose}} - N^{\text{SR}} \right) \tag{9.11}$$

Since the efficiencies are parameterised as a function of the kinematic variables, instead of using the total estimated value of $N^{\text{SR}}_{\text{fake}}$, each event in the SRloose is assigned an event weight $w^{\text{SR}}_{\text{fake}}$ such that $N^{\text{SRloose}} = 1$ and $N^{\text{SR}} = 1$ if the tight selection is passed for the lepton or 0 otherwise, denoted by δ being 1 or 0 respectively. That is,

$$w^{\text{SR}}_{\text{fake}} = \frac{\epsilon_{\text{fake}}}{\epsilon_{\text{real}} - \epsilon_{\text{fake}}} \times (\epsilon_{\text{real}} - \delta) \tag{9.12}$$

Several uncertainties are considered for the calculation of the efficiencies where due to each variation the resulting difference in yields is considered as a separate uncertainty:

- Different parameterisations are chosen $[p_T(\ell) : |\eta(\ell)|, p_T(\ell) : \Delta\phi(\ell, \nu)]$
- For the fake efficiency calculation the $t\bar{t}$ MC normalisation is changed 5%
- The control region selections are changed for $E_{\text{T}}^{\text{miss}}$ to $30 < E_{\text{T}}^{\text{miss}} < 50$ GeV

The quadratic sum of the latter two variations is considered as one uncertainty in the analysis.

9.4 Prompt photon background

All processes excluding $t\bar{t}\gamma$ production and $t\bar{t}\gamma$ decay, with a prompt photon, are grouped together as the Prompt photon background (as explained in Table 4.1). These processes are not expected to have charge asymmetry (A_C), and therefore their $|y(t)| - |y(\bar{t})|$ distribution should not impact the fit significantly. However, the prediction of their shape and normalisation are necessary to be checked with data. For this purpose, two validation regions (VR) are defined where two of the significant processes $Z\gamma$ and $W\gamma$ are well enriched: $Z\gamma$ VR and $W\gamma$ VR.

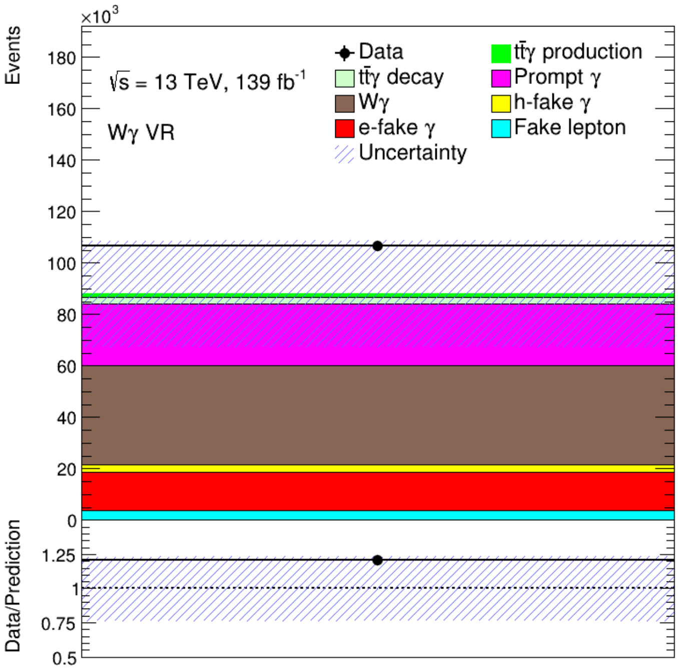
The validation regions are orthogonal to the signal region and also among themselves. They correspond to lower jet activity and different regions with less heavy flavour jets. This is done by using both 77% and 85% working points (WP) of the DL1r b-tagging algorithm. The definitions of these two regions are given in Table 9.1

Table 9.1: Definition of the prompt photon validation regions.

$Z\gamma$ VR	$W\gamma$ VR
Basic cuts of objects 1 Photon ≤ 3 jets	
2 Leptons ($e^+e^-, \mu^+\mu^-$)	1 lepton
≥ 1 b-tagged jet with 85% WP 0 b-tagged jet with 77% WP	0 b-tagged jet with 85% WP

In the $Z\gamma$ validation region, the $Z\gamma$ process constitutes around 95% of the total events. There is a good agreement between data and MC in this region and thus validates well the $Z\gamma$ process.

In the $W\gamma$ validation region, the $W\gamma$ process constitutes around 44% of the total events. The MC prediction of $W\gamma$ is expected to underestimate the number of events in data. The comparison of the yields of all the processes from MC and the data is shown in Figure 9.3. A difference in the normalisation of the $W\gamma$ process (shown separately from the Prompt γ group in the figure) can be seen, although within the uncertainties. A large normalisation uncertainty is considered in the fit (described in Chapter 11).



(a)

Figure 9.3: Event yields in the $W\gamma$ validation region; The $W\gamma$ process is shown separately from the Prompt γ group

9.5 tt̄γ decay background

The tt̄γ decay background is modelled with MC at LO QCD accuracy as described in detail in Chapter 4.1.3. This does not include the asymmetric effects that should be present at its production considering NLO QCD accuracy. Therefore, a linear re-weighting method is being used to account for this effect by modifying the reconstruction level $|y(t)| - |y(\bar{t})|$ distribution by re-weighting the corresponding truth level distribution for tt̄γ decay with the theoretical best known tt̄ asymmetry (nominal) value[3]. The theoretical value of tt̄ A_C^{NLO} is considered to be $0.0064^{+0.0005}_{-0.0006}$, calculated at NNLO accuracy in QCD with EW corrections at NLO [2, 20].

The method is done in the following steps:

- The distribution of $|y(t)| - |y(\bar{t})|$ for tt̄γ decay at truth level (from the LO MC) has two bins with yields T_1 and T_2 and having the corresponding asymmetry

---

[3] This step is done in addition to the NLO/LO k-factor applied to correct the normalisation of the tt̄γ decay process

A_C^{LO}:

$$A_C^{\text{LO}} = (T_2 - T_1)/(T_2 + T_1) \tag{9.13}$$

- The yields of the two bins are re-weighted with the asymmetry A_C^{NLO} resulting in yields T_1' and T_2', full fill the condition that the total number of events should not change with this step:

$$\begin{aligned} T_1' + T_2' &= T_1 + T_2 \\ T_1' - T_2' &= A_C^{\text{NLO}} \times (T_1' + T_2') \end{aligned} \tag{9.14}$$

- The modified bin yields are found from the yields T_1 and T_2 by multiplying them with the factors x_1 and x_2 such that:

$$T_1' = (1 + x_1) \times T_1 \tag{9.15}$$

$$T_2' = (1 + x_2) \times T_2 \tag{9.16}$$

  These factors are calculated using the above equations:

$$x_1 = \frac{1}{2} \times (T_1 + T_2) \times \frac{(A_C^{\text{NLO}} - A_C^{\text{LO}})}{T_1} \tag{9.17}$$

$$x_2 = -\frac{1}{2} \times (T_1 + T_2) \times \frac{(A_C^{\text{NLO}} - A_C^{\text{LO}})}{T_2} \tag{9.18}$$

- To propagate the modified asymmetry to the reconstructed level distribution (with yields R_1 and R_2), the factors x_1 and x_2 are used with the migration matrix, M[4] of the MC sample as:

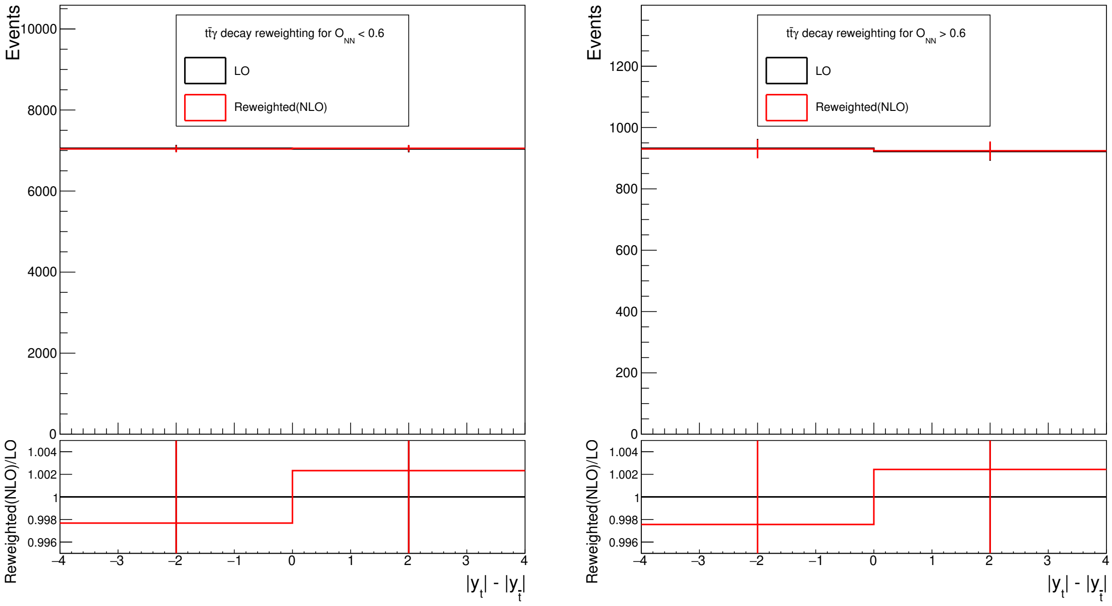
$$R_1 = (1 + x_1) \times M_{11} + (1 + x_2) \times M_{12} \tag{9.19}$$

$$R_2 = (1 + x_1) \times M_{21} + (1 + x_2) \times M_{22} \tag{9.20}$$

- The new reconstruction level distribution with yields R_1 and R_2 are used finally for the t$\bar{\text{t}}\gamma$ decay background. For consistency, the contribution obtained with the NLO QCD t$\bar{\text{t}}$ samples are also reweighted similarly to match the asymmetry value, A_C^{NLO}. As a result of this reweighting procedure, the asymmetry value of the $|y(t)| - |y(\bar{t})|$ distribution for t$\bar{\text{t}}\gamma$ decay increases slightly as can be seen in Figure 9.4.

---

[4] The migration matrices are similar to the ones described in Chapter 11.1.1, for the t$\bar{\text{t}}\gamma$ decay sample



(a) (b)

Figure 9.4: The $|y(t)| - |y(\bar{t})|$ distribution for $t\bar{t}\gamma$ decay shown for the two signal regions $O_{\text{NN}} < 0.6$(a) and $O_{\text{NN}} \geq 0.6$(b) (described in Chapter 10)

The re-weighted $t\bar{t}\gamma$ decay distributions with the modified asymmetry are used finally in the analysis. As it can be seen in Figure 9.4, the re-weighted distributions are barely modified (<0.5%) and hence do not significantly impact the analysis.

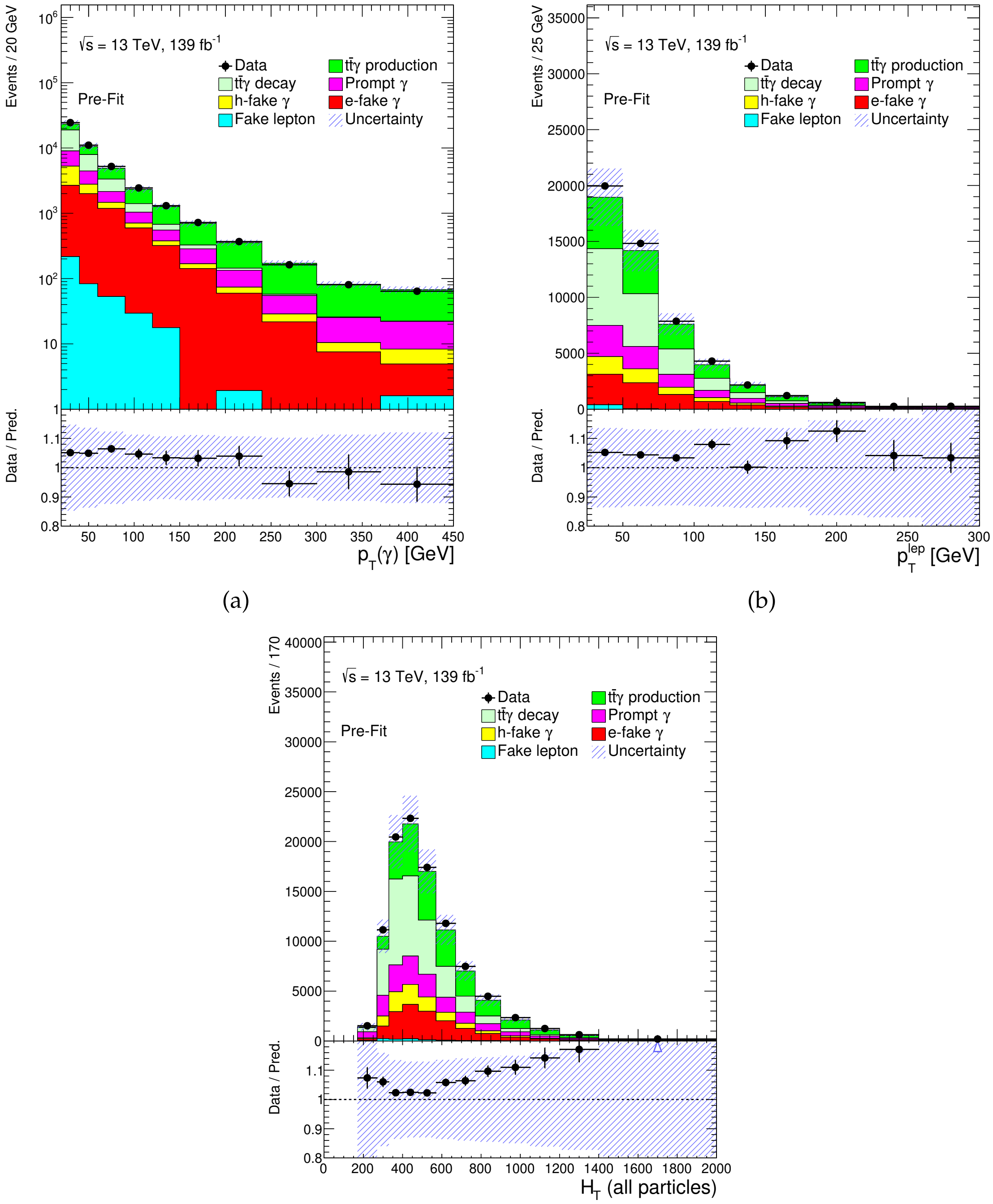
CHAPTER 10

Discrimination of signal from background processes

In this chapter, the comparison of the data and prediction are discussed in Section 10.1 for various physics observables, after following the event selection (described in Chapter 6.2). The expected number of events for the background processes are determined after applying the corrections described in Chapter 9. The proportion of signal compared to the total selected events is found to be around 27% only. To enhance the sensitivity of the measurement, the separation between the events of the signal and background processes is required. A neural network (NN) [93] based discriminator variable is constructed for this purpose. The network training is based on 21 input variables (as described in Section 10.2) with MC-generated events of both the signal and background processes. The construction of the neural network is based on a simple network architecture described in Section 10.3. The performance of the network is described in Section 10.4. Afterwards, the output discriminating variable of the network is used for the analysis (Section 10.5) to determine a region boundary to define two regions. Finally, the event yields of the two regions are discussed in Section 10.6.

10.1 Comparison of data and predictions

The comparison of data and prediction for several physical observables are discussed in this section. They are defined in Table 10.1 and their distributions are shown in the respective Figures 10.1-10.6, where their sum is compared with data and the ratio of data and prediction is shown separately. Their uncertainty band includes the total of statistical and systematic uncertainties, as described in Chapter 7.

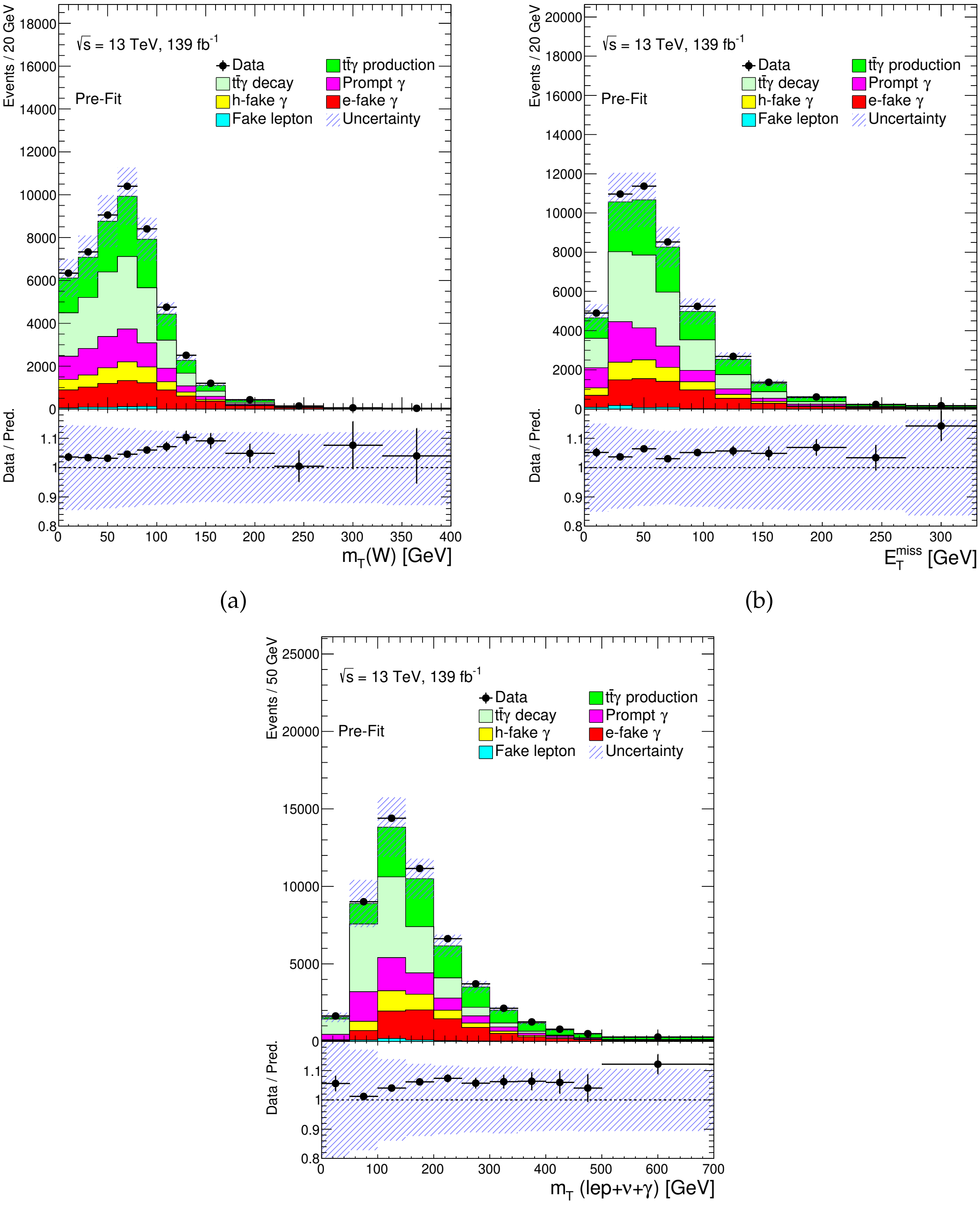


(a)

(b)

(c)

Figure 10.1: (a) $p_T(\gamma)$ (b) $p_T(\ell)$ (c) $H_T(all)$ for all the processes are shown, and their total is compared with data. The uncertainties include statistical and all experimental and theoretical systematic uncertainties added in quadrature. The lower part of the plot shows the ratio of the data to the prediction. The last bin of each distribution contains the overflow events, if applicable.

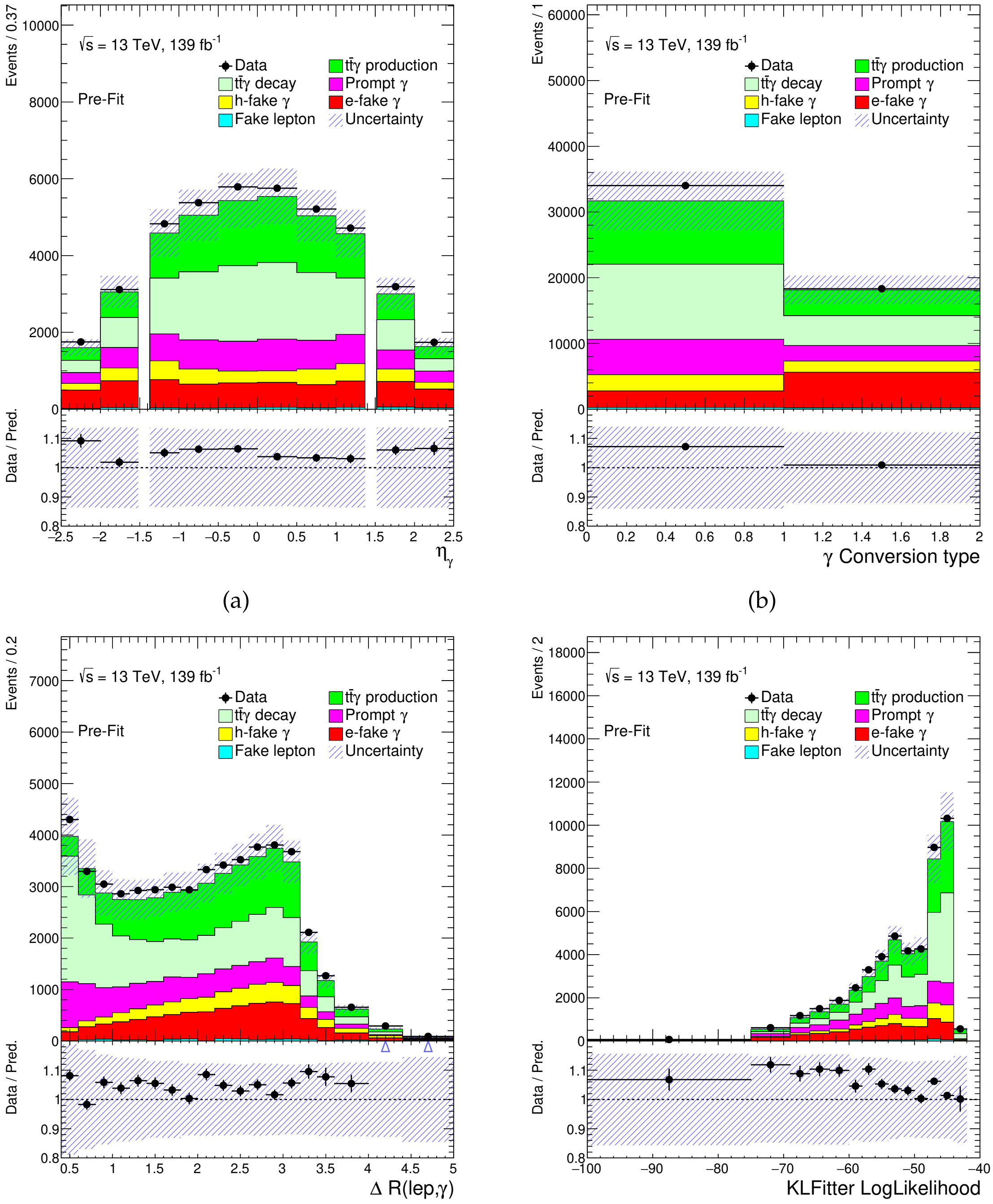


(a)

(b)

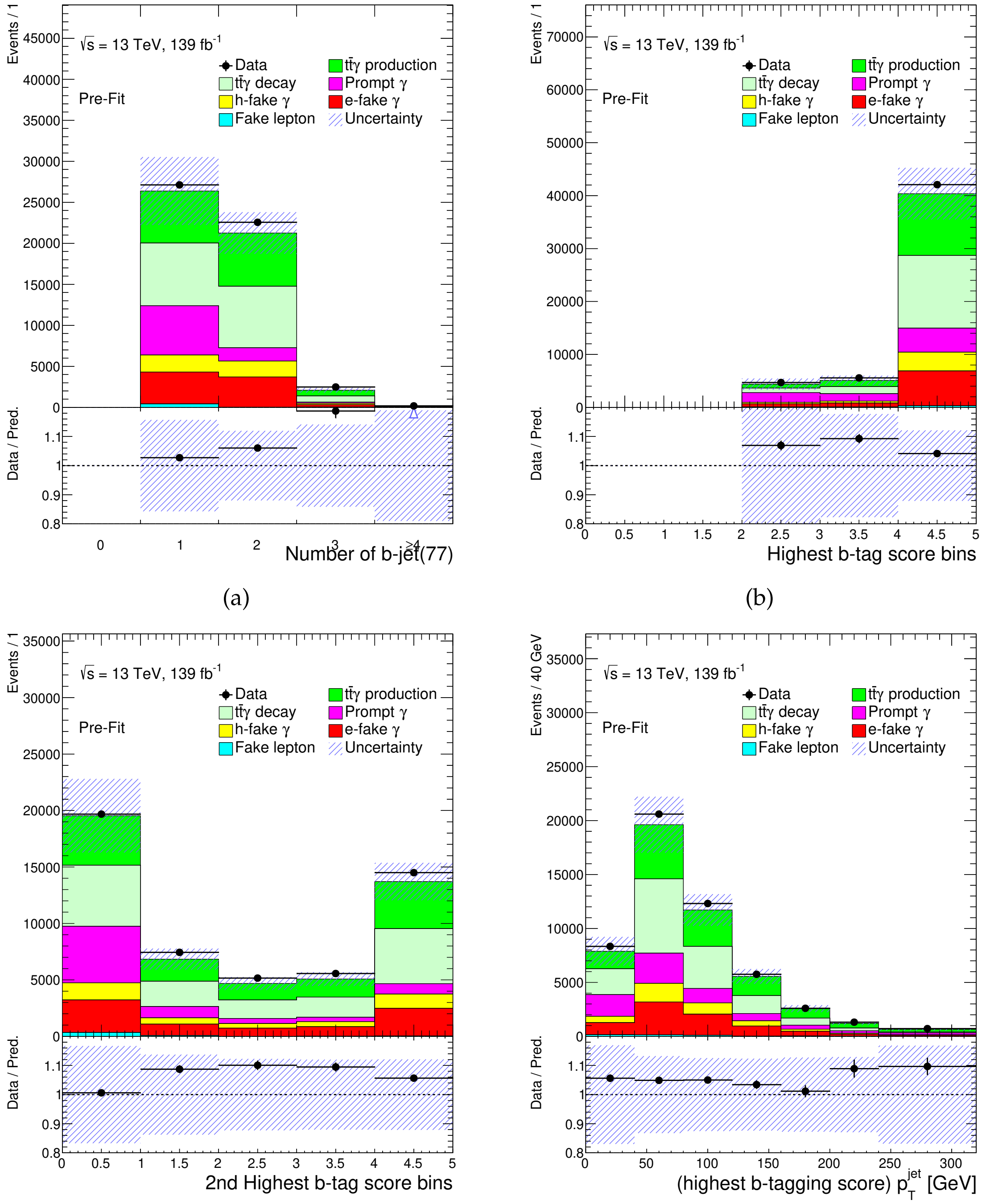
(c)

Figure 10.2: (a) $m_T(W)$ (b) $E_{\mathrm{T}}^{\mathrm{miss}}$ (c) $m_T(\ell,\nu,\gamma)$ for all the processes are shown, and their total is compared with data. The uncertainties include statistical and all experimental and theoretical systematic uncertainties added in quadrature. The lower part of the plot shows the ratio of the data to the prediction. The last bin of each distribution contains the overflow events, if applicable.



(a) (b) (c) (d)

Figure 10.3: (a) $\eta(\gamma)$ (b) γ conversion type (c) $\Delta R(\ell,\gamma)$ (d) KLFitter Log-likelihood for all the processes are shown, and their total is compared with data. The uncertainties include statistical and all experimental and theoretical systematic uncertainties added in quadrature. The lower part of the plot shows the ratio of the data to the prediction. The last bin of each distribution contains the overflow events, if applicable.



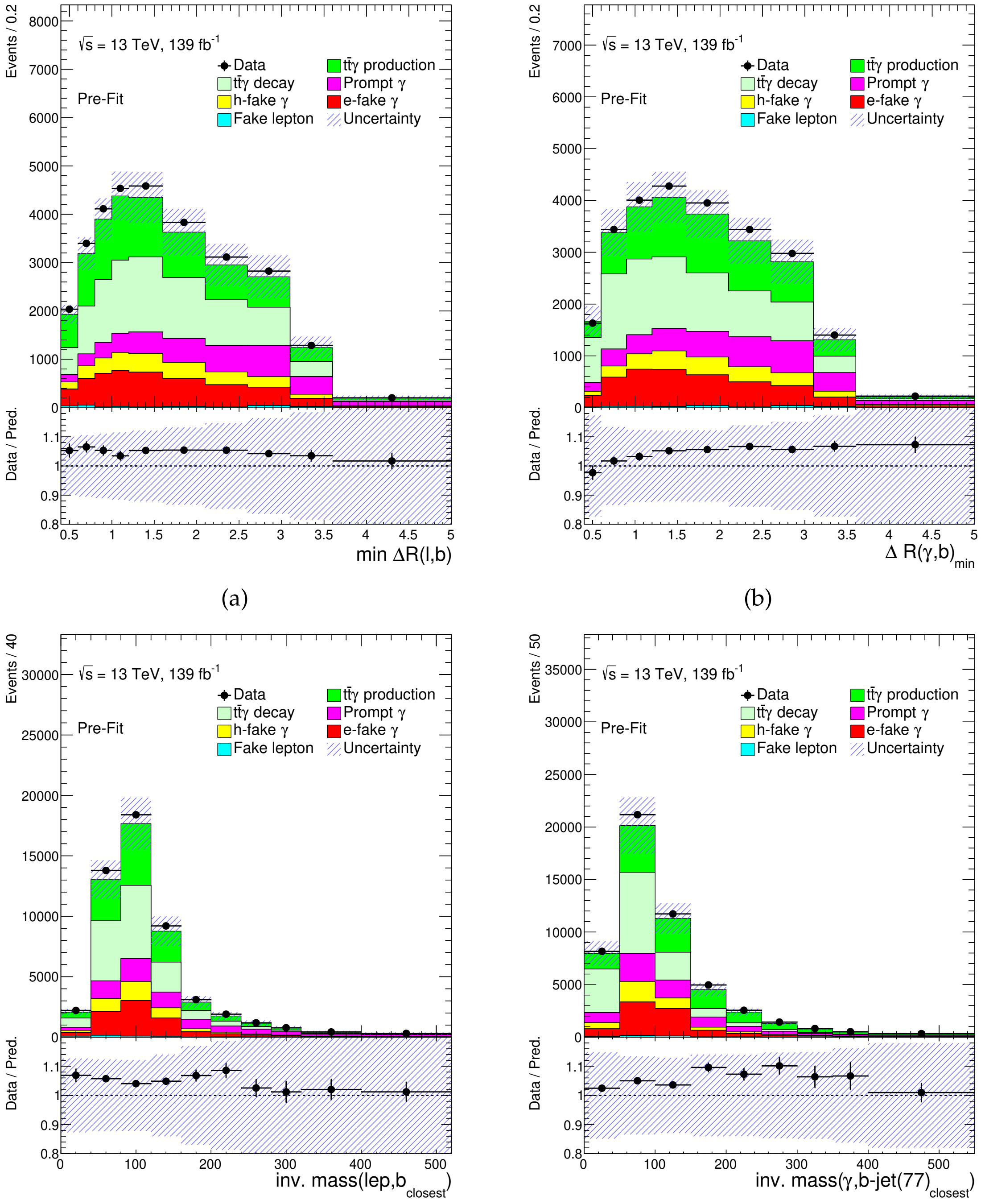
(a)

(b)

(c)

(d)

Figure 10.4: (a) $N(\text{b-jet}_{77})$ (b) Highest jet DL1r binned score (c) 2nd highest jet DL1r binned score (d) $p_T(\text{jet [highest DL1r score]})$ for all the processes are shown, and their total is compared with data. The uncertainties include statistical and all experimental and theoretical systematic uncertainties added in quadrature. The lower part of the plot shows the ratio of the data to the prediction. The last bin of each distribution contains the overflow events, if applicable.



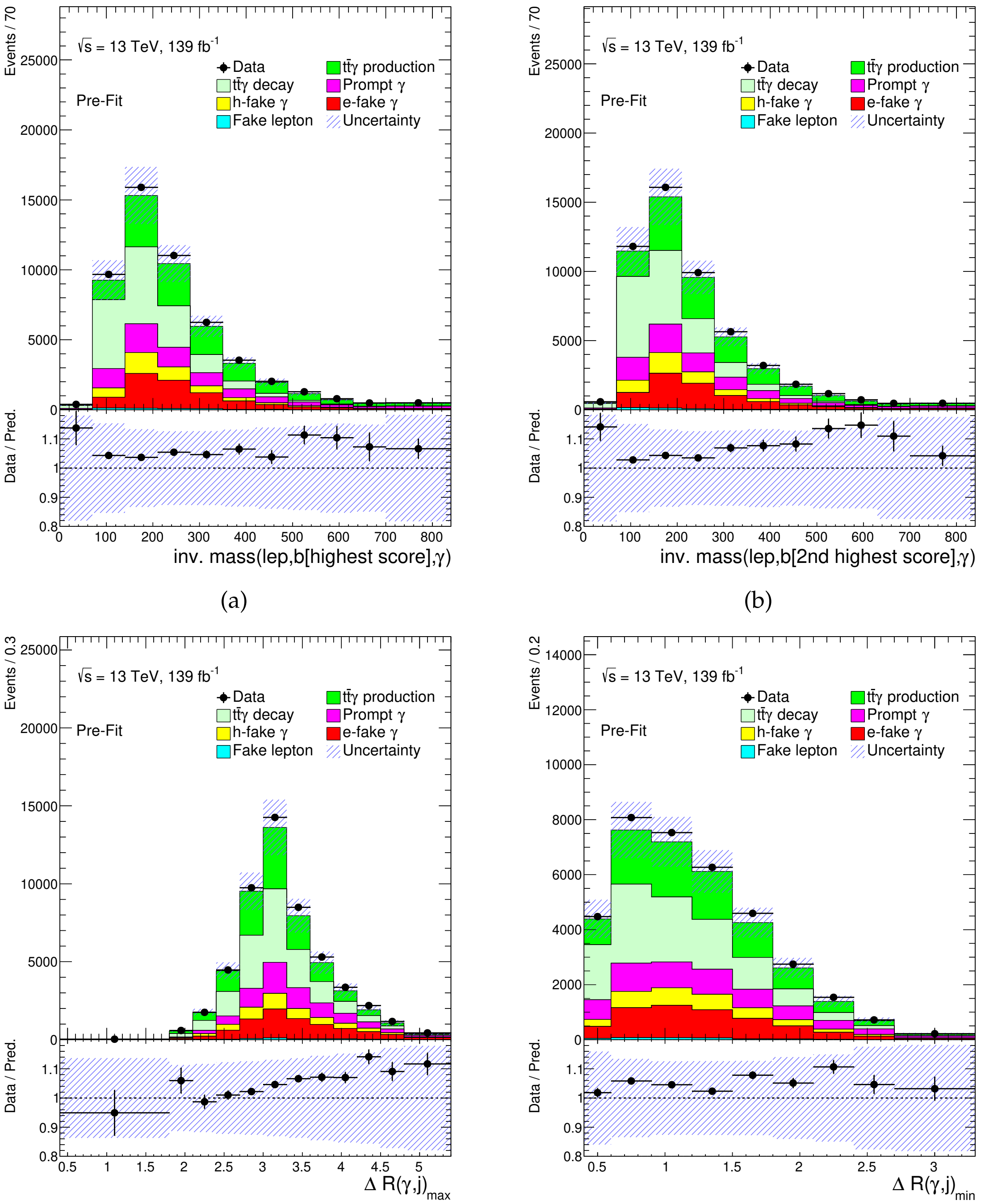
(a)

(b)

(c)

(d)

Figure 10.5: (a) minimum $\Delta R(\ell, \text{b-jet}_{77})$ (b) minimum $\Delta R(\gamma, \text{b-jet}_{77})$ (c) closest $m(\ell, \text{b-jet}_{77})$ (d) closest $m(\gamma, \text{b-jet}_{77})$ for all the processes are shown, and their total is compared with data. The uncertainties include statistical and all experimental and theoretical systematic uncertainties added in quadrature. The lower part of the plot shows the ratio of the data to the prediction. The last bin of each distribution contains the overflow events, if applicable.



(a) (b) (c) (d)

Figure 10.6: (a) $m(\ell, \gamma, \text{jet [highest DL1r score]})$ (b) $m(\ell, \gamma, \text{jet [2nd highest DL1r score]})$ (c) maximum $\Delta R(\gamma, jet)$ (d) minimum $\Delta R(\gamma, jet)$ for all the processes are shown, and their total is compared with data. The uncertainties include statistical and all experimental and theoretical systematic uncertainties added in quadrature. The lower part of the plot shows the ratio of the data to the prediction. The last bin of each distribution contains the overflow events, if applicable.

10.2 Selection of variables for neural network

A fully connected neural network is used for the discrimination of signal events from backgrounds. The choice of input variables is very crucial for this network for such network architecture. A total of 85 relevant variables are explored for this purpose. The variables are essentially based on the kinematic properties of the selected objects. Some of these are compound functions or high-level variables made from the object kinematics. Out of the 85, only 21 input variables are chosen finally, and they are described in Table 10.1. The DL1r b-jet discriminant output score for the selected jets in the analysis is used in the specific bins (-10, 0.665, 2.195, 3.245, 4.565, 20.) corresponding to few specific WPs that are calibrated with data, referred to as *jet DL1r binned score* in the table.

To judge whether a variable should be considered as input for the network, individually their normalised distributions for the signal and backgrounds are used. The difference in the shape of their distributions gives a hint of their importance. For this purpose, a merit value called *Separation* is defined in Equation 10.1, where a sum is done over the bins (N_{bins}) of the distribution of the considered variable using the squared shape differences of the signal (S_i) and each of the background (B_i) groups in each bin i.

$$\text{Separation} = \frac{1}{2}\sum_{i}^{N_{\text{bins}}} \frac{(S_i - B_i)^2}{S_i + B_i} \times 100\% \tag{10.1}$$

The Separation is calculated for all explored input variables. Except for the integer variables, all of their normalised distributions are given the same number of bins: $N_{\text{bins}} = 20$. The considered integer variables have fewer bins. The variables are selected based on the calculated Separation values separately for each of the backgrounds (as defined in Chapter 4.4.1).

As an example, in Figure 10.7 the comparisons are shown for the signal and the background groups for the variable, $\Delta R(\ell, \gamma)$ where all the distributions are given 20 bins and normalised to one. The Separation is calculated following Equation 10.1 for each of the backgrounds and are shown in the legend. For this particular variable, the $t\bar{t}$ +prompt γ (or $t\bar{t}\gamma$ decay) has the highest Separation and good separation is also shown for the other prompt γ background[1]. Similar plots for the rest of the selected input variables are shown in Appendix B.1.

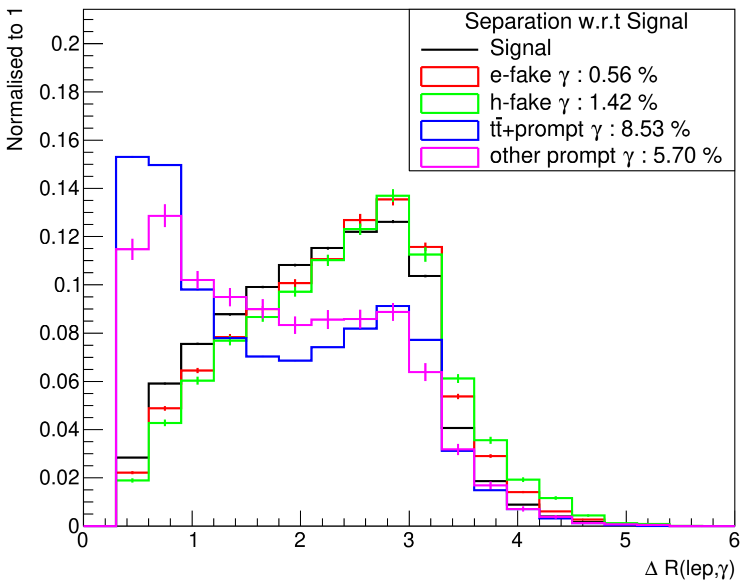
As the final selection of the variables, they are chosen only if their calculated Separation is more than 1% for at least one of the backgrounds. The Separation values for the finally selected 21 variables are shown in Figure 10.8, where higher

---

[1] In this chapter the **other prompt** γ implies the **Prompt** γ as defined in Chapter 4.4.1

Table 10.1: The final list of input variables used in the neural network training and their description

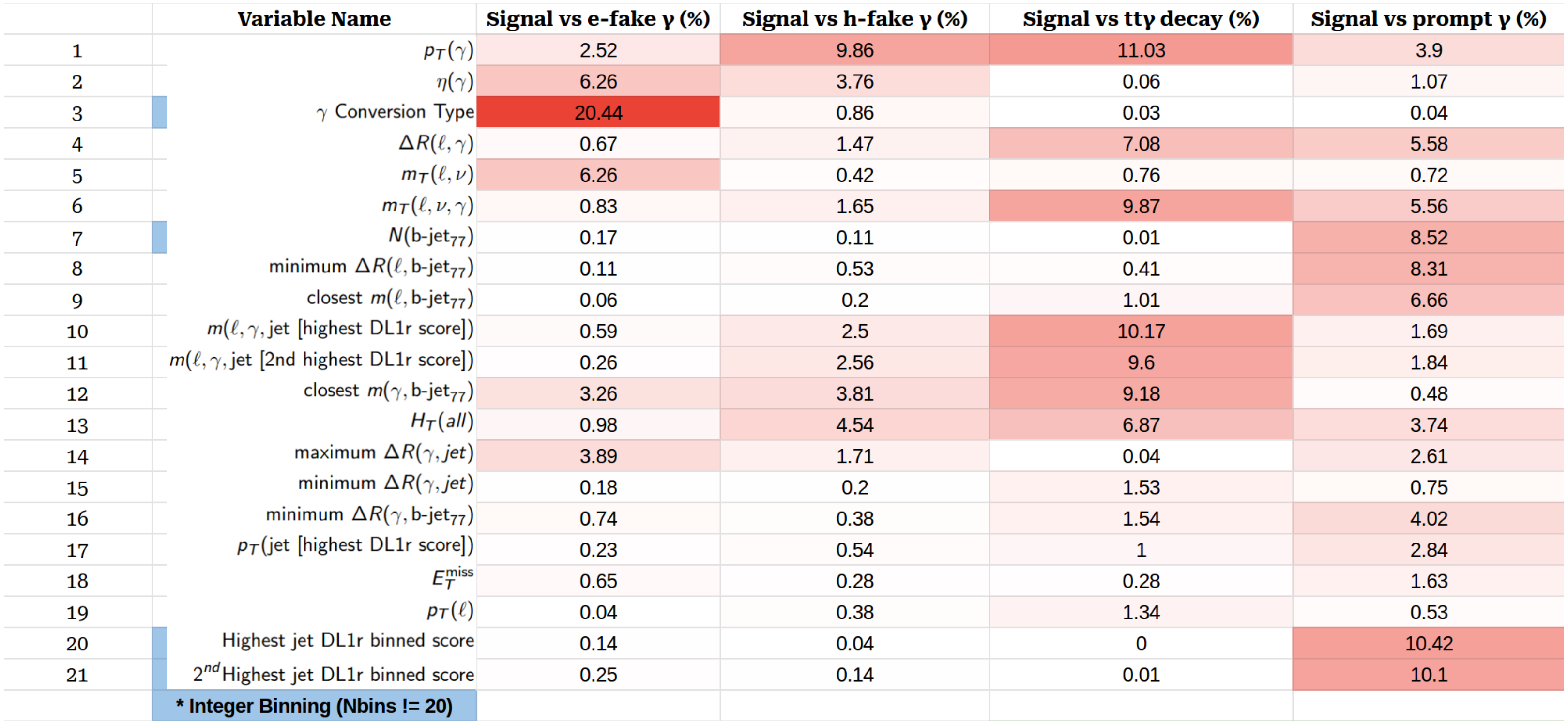
Name of Variable	**Description**
$p_T(\gamma)$	p_{T} of the photon
$p_T(\ell)$	The p_{T} of the lepton
$H_T(all)$	The sum of transverse momentum of all charged and neutral particles
$E_{\mathrm{T}}^{\mathrm{miss}}$	The missing transverse energy
$m_T(\ell, E_{\mathrm{T}}^{\mathrm{miss}})$ or $m_T(W)$	Transverse mass of lepton-transverse missing energy system
$m_T(\ell, E_{\mathrm{T}}^{\mathrm{miss}}, \gamma)$	Transverse mass of lepton-photon-transverse missing energy system
$\eta(\gamma)$	η of the photon
γ conversion type	Unconverted (0) or Converted(0) photon
$\Delta R(\ell, \gamma)$	Distance between lepton and photon
N(b-jet$_{77}$)	Number of b-tagged jets with 77% efficiency working point (WP)
Highest jet DL1r binned score	The bin value of the DL1r score of the b-tagged jet with the highest score
2nd highest jet DL1r binned score	The bin value of the DL1r score of the b-tagged jet with the 2$^{\mathrm{nd}}$ highest score
p_T(jet [highest DL1r score])	p_{T} of the b-tagged jet with the highest DL1r b-tagging score
minimum $\Delta R(\ell,$b-jet$_{77})$	The minimum of the distances between the lepton and any b-tagged jet (77% WP)
minimum $\Delta R(\gamma,$b-jet$_{77})$	The minimum of the distances between the photon and any b-tagged jet (77% WP)
closest $m(\ell,$b-jet$_{77})$	Invariant mass of the system with the minimum $\Delta R(\ell,$b-jet$_{77})$
closest $m(\gamma,$b-jet$_{77})$	Invariant mass of the system with the minimum $\Delta R(\gamma,$b-jet$_{77})$
$m(\ell, \gamma,$jet [highest DL1r score])	Invariant mass of the lepton-photon system and the jet with the highest DL1r b-tagging score
$m(\ell, \gamma,$jet [2nd highest DL1r score])	Invariant mass of the lepton-photon system and the jet with the 2$^{\mathrm{nd}}$ highest DL1r b-tagging score
maximum $\Delta R(\gamma, jet)$	The maximum of the distances between the photon and the jets
minimum $\Delta R(\gamma, jet)$	The minimum of the distances between the photon and the jets



(a)

Figure 10.7: The comparison of the normalised (to 1) distributions for the variable $\Delta R(\ell, \gamma)$ of the signal with respect to the different background categories; the calculated Separation values are shown in the legend

values of Separation are highlighted with a more red-coloured text background. The integer variables are also marked.



		Variable Name	Signal vs e-fake γ (%)	Signal vs h-fake γ (%)	Signal vs ttγ decay (%)	Signal vs prompt γ (%)
1		$p_T(\gamma)$	2.52	9.86	11.03	3.9
2		$\eta(\gamma)$	6.26	3.76	0.06	1.07
3	🟦	γ Conversion Type	20.44	0.86	0.03	0.04
4		$\Delta R(\ell,\gamma)$	0.67	1.47	7.08	5.58
5		$m_T(\ell,\nu)$	6.26	0.42	0.76	0.72
6		$m_T(\ell,\nu,\gamma)$	0.83	1.65	9.87	5.56
7	🟦	$N(\text{b-jet}_{77})$	0.17	0.11	0.01	8.52
8		minimum $\Delta R(\ell,\text{b-jet}_{77})$	0.11	0.53	0.41	8.31
9		closest $m(\ell,\text{b-jet}_{77})$	0.06	0.2	1.01	6.66
10		$m(\ell,\gamma,\text{jet [highest DL1r score]})$	0.59	2.5	10.17	1.69
11		$m(\ell,\gamma,\text{jet [2nd highest DL1r score]})$	0.26	2.56	9.6	1.84
12		closest $m(\gamma,\text{b-jet}_{77})$	3.26	3.81	9.18	0.48
13		$H_T(all)$	0.98	4.54	6.87	3.74
14		maximum $\Delta R(\gamma,jet)$	3.89	1.71	0.04	2.61
15		minimum $\Delta R(\gamma,jet)$	0.18	0.2	1.53	0.75
16		minimum $\Delta R(\gamma,\text{b-jet}_{77})$	0.74	0.38	1.54	4.02
17		$p_T(\text{jet [highest DL1r score]})$	0.23	0.54	1	2.84
18		E_T^{miss}	0.65	0.28	0.28	1.63
19		$p_T(\ell)$	0.04	0.38	1.34	0.53
20	🟦	Highest jet DL1r binned score	0.14	0.04	0	10.42
21	🟦	2^{nd}Highest jet DL1r binned score	0.25	0.14	0.01	10.1
	🟦	* Integer Binning (Nbins != 20)				

Figure 10.8: The chart of the calculated Separation (10.1) for the individual backgrounds for all the 21 input variables selected for neural network training. Except for the integer variables (γ Conversion type, $N_{\text{b-jet}}$ and binned b-tagging scores) all variables have 20 bins in their corresponding histograms.

10.3 Description and training of neural network

A simple, fully connected neural network architecture is chosen for this analysis. However, a detailed study on different hyperparameters of the network has been done to determine the optimal version of the network, as summarised in Section 10.3.3. The description and training of the finally chosen optimized network are in the following.

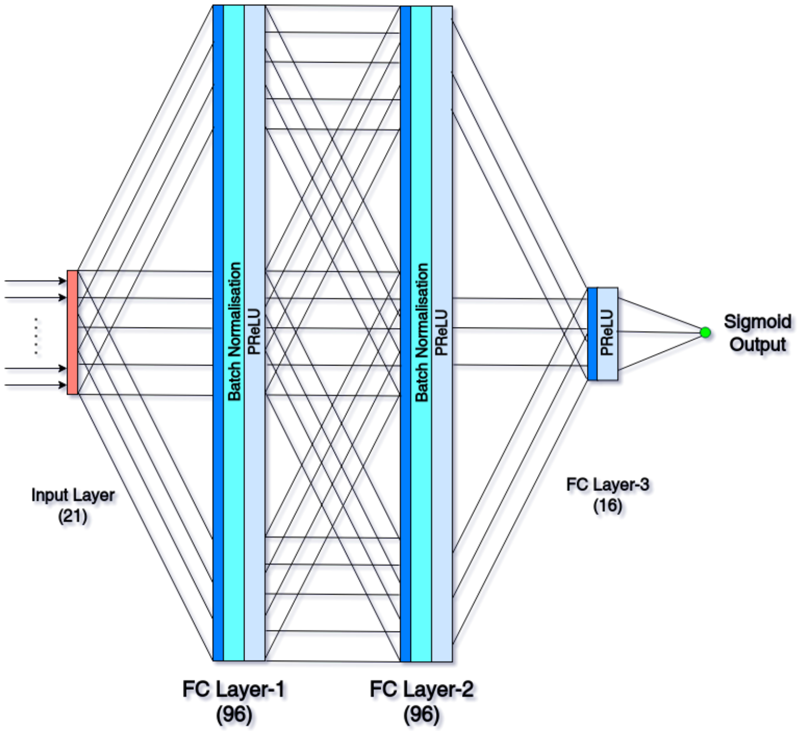


Figure 10.9: The architecture of the neural network using 21 input variables. It has three hidden layers with 96, 96 and 16 nodes in each, with PReLU activation layers and/or batch-normalisation following them. The output node of the network is used following a sigmoid activation function.

10.3.1 Network architecture

The network is mainly made of three hidden layers as it is shown in Figure 10.9 along with the input layer consisting of 21 variables (explained above) transformed by scaling to be within (0,1) using *Scikit-learn's* [94] *MinMax* scaler. Layer-1 and Layer-2 are both 96 nodes. Layer-3 has 16 nodes. All the hidden layers are followed by a *Parametric Rectified Linear Unit* or *PReLU* [95] activation function to keep parametrised non-linearity in the output of the different hidden layers, with a as an additional parameter, described as a function of a variable x as:

$$f(a,x) = x, x \geq 0$$
$$f(a,x) = ax, x < 0$$

Layer-1 and Layer-2 are also followed by a *Batch Normalisation* [96] Layer. This helps to avoid very high or very low activation values in the layer before. These two steps add additional trainable parameters. Layer-2 and Layer-3 also have a L_2 [97] regularisation term (with *kernel_regularizer* value 0.15) added for them in the Loss function to avoid over-training the network. The output node is used, following a *Sigmoid* activation function $f(x) = \frac{1}{1+e^{-x}}$, to get the probability of any event being 1 for signal or 0 for background. By this construction, a total of 13969 parameters are used in the network, out of which 13585 are used during the training.

10.3.2 Training of the network

Before training, all the events in the MC samples are weighted properly. The individual event weights (Chapter 6.3) are used since the sum of the event weights determine the proportion of events belonging to a process rather than the total number of entries in the corresponding MC samples. There are some MC samples that contain negative weighted events. These events are removed for the training as they can be treated wrongly by the network. It is shown in Appendix B.2 that it is safe to use only the positive weighted events for the training, and the difference in the shape of the network output is negligible. The normalisation of each sample is corrected after the removal of the negative weighted events so that the overall proportion of each process remains the same. However, the total events available for the signal MC and all background MC samples together are not the same. Hence, the weights of the background events are additionally reweighted to keep the balance with the signal. This weighting procedure is done for both training and test datasets.

Next, the datasets are divided between the training+validation and test sets in 85% and 15% respectively. The training and validation sets are used in a 5-fold cross-validation setup by splitting them further, with each split made of 17% of total events. The validation set is switched with each of the five splits while using the rest of the splits for training five different models. The test dataset is not used anyhow during the training and is only used to assess the performance of the models afterwards.

The training of the network is performed using Keras [98] with the TensorFlow [99] backend. For training and optimization of the network, a *Binary-Cross-entropy* function is used as the Loss function with *Adam* [100] optimizer with *AMSGrad* [101] variant. A *learning rate* of $1 \cdot 10^{-3}$ with decay rate of $15 \cdot 10^{-5}$ is used in the optimizer.

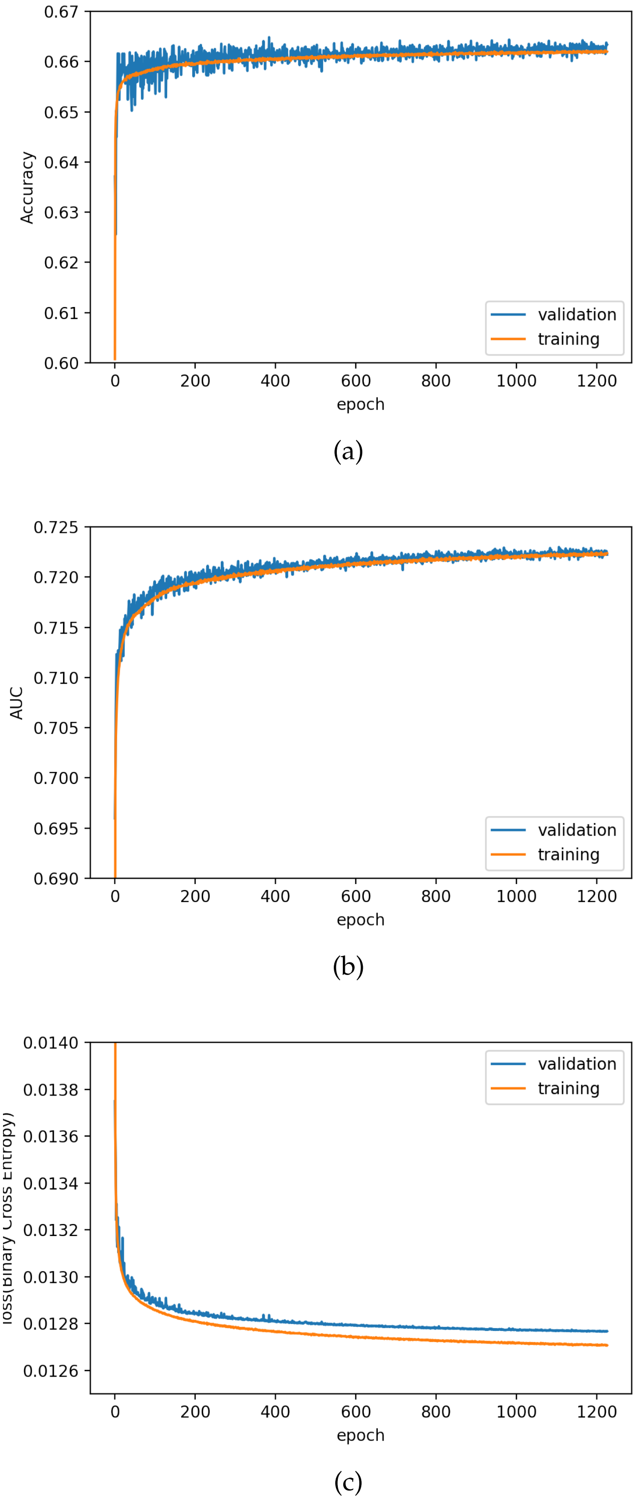


Figure 10.10: (a) Accuracy, (b) AUC and (c) loss function of the network during the training progression shown for each epoch.

A *batch size* of 512 is used for the training.

The progression of the network during the training of the Accuracy (probability of making a correct prediction) and the Area Under the Receiver-Operating-Characteristic (ROC [102]) curve or AUC on the signal and background discrimination for the training and validation datasets for each epoch are shown in Figure 10.10 (a) and Figure 10.10 (b) respectively.

The training is run with a maximum epoch of 1500 and stopped with a *EarlyStopping* [103] function (*patience*=50, *min_delta*=$1 \cdot 10^{-6}$) on the validation loss when its slope gets close to zero to avoid any over-fitting as shown in Figure 10.10 (c).

10.3.3 Hyperparameters

Several hyperparameters are considered for network training, with different combinations of their multiple values. The final chosen network has the optimal values of them, corresponding to the best performance (AUC) of the network output. The hyperparameters and their searched values are listed in the following with the chosen values underlined:

- Number of Layer-1 and Layer-2 nodes: 32, 64, $\underline{96}$, 128, 160
- Number of Layer-3 nodes: 8, $\underline{16}$, 32
- Learning rate: 0.0005, $\underline{0.001}$, 0.0015, 0.002
- Batch size: 128, 256, $\underline{512}$, 1024, 2048
- Adam's AMSGrad variant: $\underline{\text{on}}$, off
- Regularizer: Dropout layers, $\underline{\text{L2}}$

10.4 Performance of the neural network

The output of the neural network, O_{NN}, is then used as a new variable in the analysis. The O_{NN} is evaluated for all the MC samples and data as shown in Figure 10.11. Since the same MC samples are being used partially in training and also in the analysis, the five trained models (in the cross-validation approach) are used (inferred) in the analysis in a way so that they were not trained with the same events. The ROC curves of the five trained networks for the test dataset are shown in Figure 10.12. The five cross-validated training models have a negligible ensemble uncertainty of 0.003 on the AUC, implying an unbiased training.

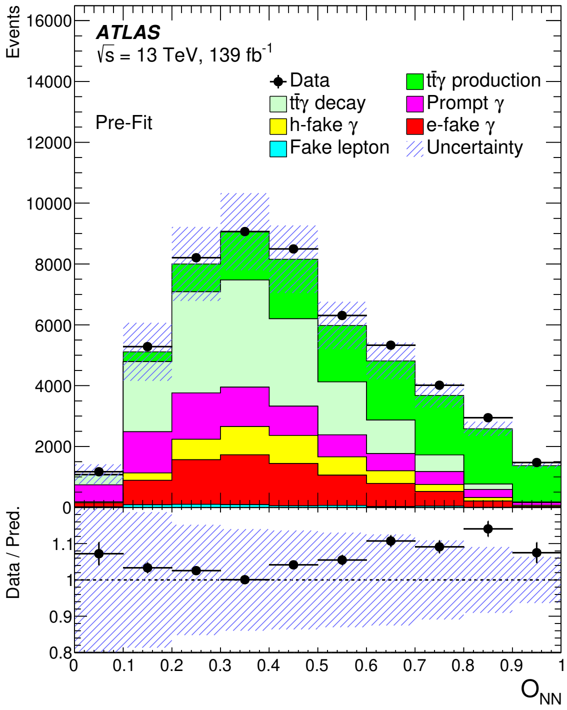
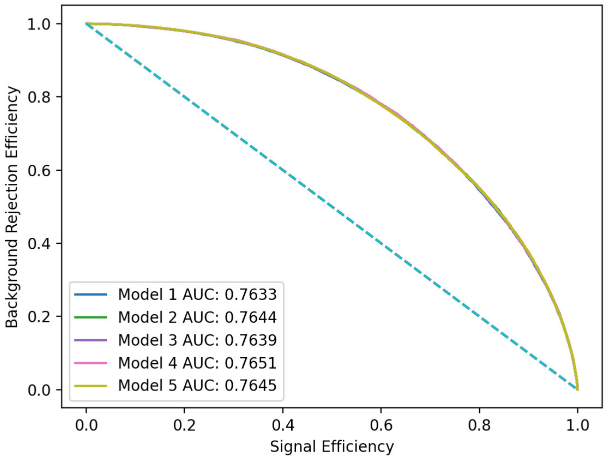


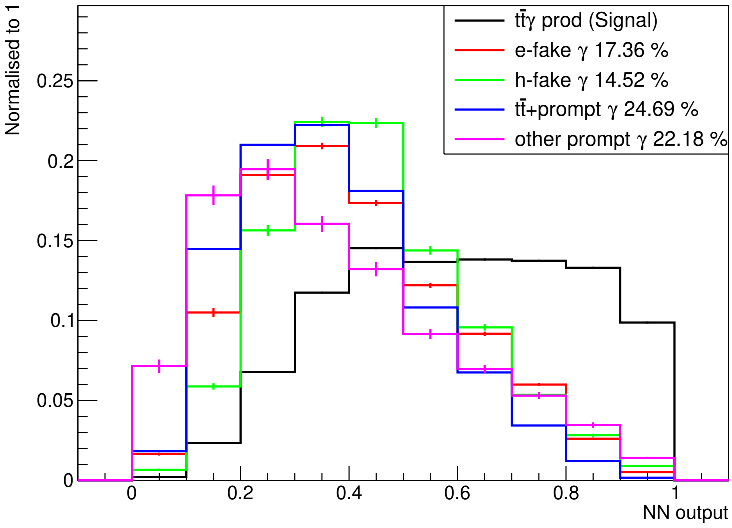
Figure 10.11: Neural network output for all the processes are shown, and their total is compared with the data. The uncertainties include statistical and all experimental and theoretical systematic uncertainties added in quadrature [3]. The lower part of the plot shows the ratio of the data to the prediction.

To understand the discrimination power of O_{NN} between the signal and the backgrounds, the corresponding normalised distributions are compared in Figure 10.13 with the Separation values calculated for each of the backgrounds similar to the input variables as in Figure 10.7. From the values in the plot, the backgrounds can be ranked accordingly from the best to worst separable as: the $t\bar{t}\gamma$ decay ($t\bar{t}$ + prompt γ), other prompt γ, e-fake γ and h-fake γ. The reason behind the fake photon backgrounds being less separable from the signal can be because of the availability of fewer input variables with good Separation values, as shown previously in Figure 10.8.



(a)

Figure 10.12: The ROC curve of the network outputs for the five trained models for signal and all backgrounds with the area under the curve (AUC) for the test (unused in training) dataset shown in the legend, with a negligible ensemble uncertainty implying an unbiased training.

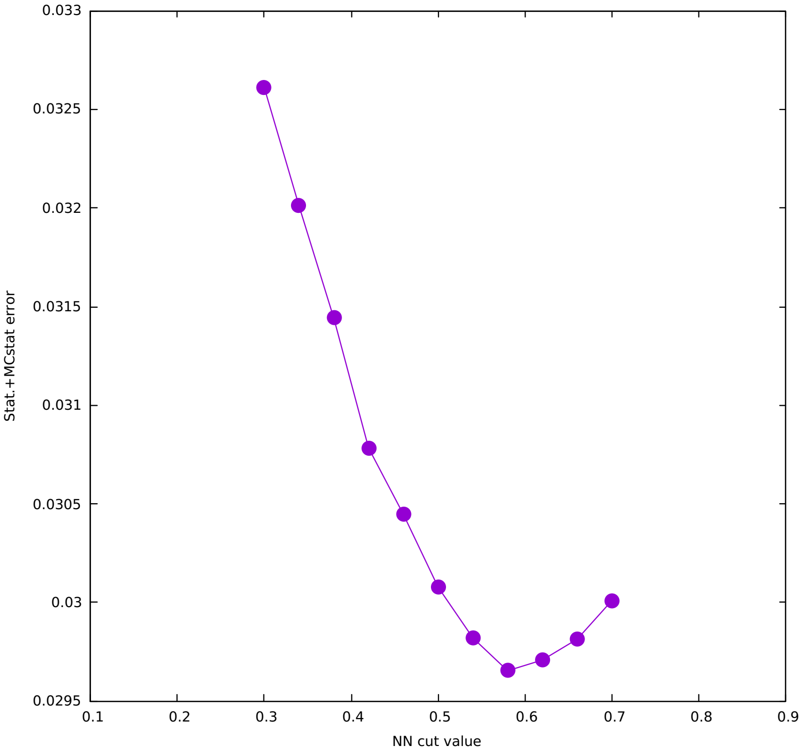


(a)

Figure 10.13: Comparison of network output for signal and the different backgrounds. In the legend, the Separation values of different backgrounds are also shown for the output variable.

10.5 Application of the neural network in the analysis

The O_{NN} variable is not directly used in the likelihood fit as described in Chapter 11, but only for further defining the signal and background enriched regions after applying the primary event selection described in Chapter 6.2. In particular, events are divided into two regions based on O_{NN} being below or above a threshold/cut value. To determine the optimal cut value, the measurement with pseudo data (explained later in Chapter 11) and without the systematic uncertainties is repeated for several cut values (11 values around 0.5, between 0.3 and 0.7). Since the uncertainties of the measurement from systematic sources are expected to be smaller than the dominant statistical and MC statistical uncertainties, they are ignored for this optimisation. The cut value with the least total uncertainty (i.e. statistical + MC statistical uncertainty) on the asymmetry A_C is chosen finally. The optimised cut value is 0.6, as can be found from Figure 10.14 that the nearest cut value indicates a minimum of the total uncertainty.



(a)

Figure 10.14: The optimisation of the cut value on O_{NN} for defining the two regions

10.6 Event yields

Event yields refer to the events that pass the event selection criteria as defined in Chapter 6. They are further divided into two regions using the neural network output value (O_{NN}) at 0.6 as a boundary. The event yields in the two regions are shown in Table 10.2. The uncertainty indicates the total of statistical and systematic uncertainties. The comparison is also shown in a logarithmic scale in Figure 10.15.

Table 10.2: Event yields in the two regions $O_{\text{NN}} < 0.6$ and $O_{\text{NN}} \geq 0.6$

Processes	$O_{\text{NN}} < 0.6$	$O_{\text{NN}} \geq 0.6$
$t\bar{t}\gamma$ production	$6\,660 \pm 350$	$6\,910 \pm 340$
$t\bar{t}\gamma$ decay	$14\,100 \pm 3\,100$	$1\,900 \pm 560$
Prompt γ	$6\,400 \pm 2\,000$	$1\,300 \pm 400$
H-fake γ	$3\,400 \pm 1\,400$	790 ± 360
E-fake γ	$6\,420 \pm 860$	$1\,480 \pm 260$
Lepton fake	410 ± 110	57 ± 35
Total	$37\,400 \pm 4\,500$	$12\,400 \pm 1\,100$
Data	38527	13763

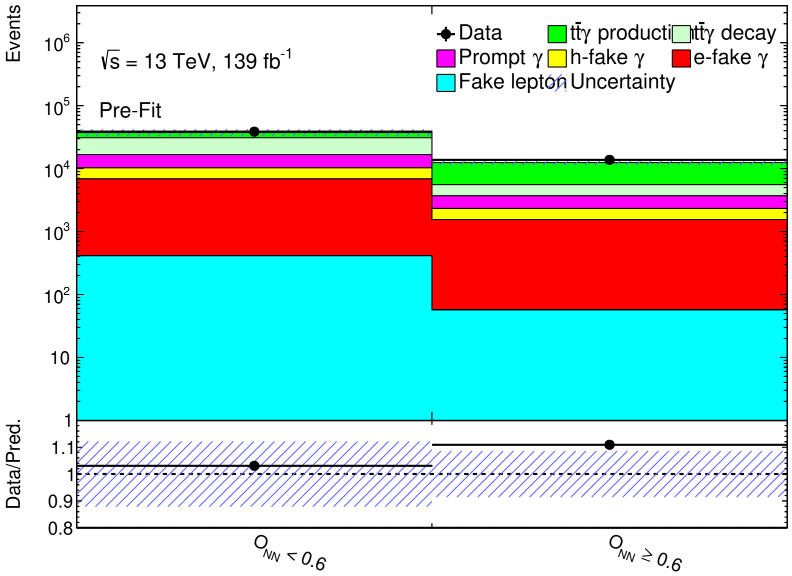
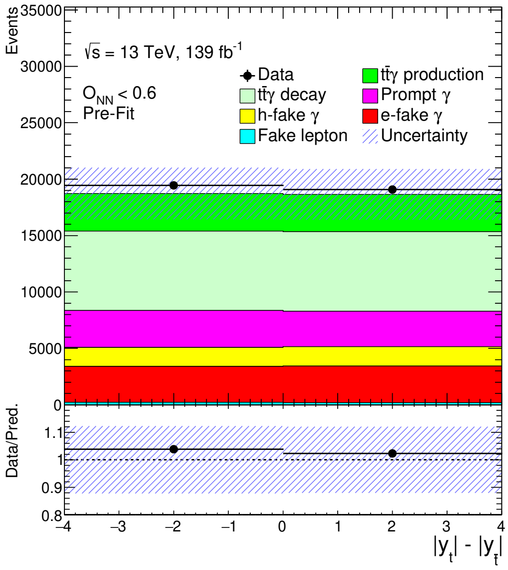


Figure 10.15: The event yields of the prediction of all the processes and their total is compared with the data divided into the two regions based on the neural network output value being less or greater than 0.6. The uncertainties include all statistical and systematic uncertainties added in quadrature. The lower part of the plot shows the ratio of the data to the prediction.

CHAPTER 11

The analysis and measurement procedure



(a) (b)

Figure 11.1: The distribution of $|y(t)| - |y(\bar{t})|$ for all the processes in the two regions $O_{\mathrm{NN}} < 0.6$ and $O_{\mathrm{NN}} \geq 0.6$ are shown and their total is compared with the data. The uncertainties include statistical and all systematic uncertainties added in quadrature. The lower part of the plot shows the ratio of the data to the prediction.

Once the MC predicted $|y(t)| - |y(\bar{t})|$ distributions (difference of the absolute rapidity of the two top quarks) at the reconstruction level are obtained for the two regions $O_{\mathrm{NN}} < 0.6$ and $O_{\mathrm{NN}} \geq 0.6$, they are fitted with the corresponding data.

The fit is performed along with unfolding the measurement by using the profile likelihood unfolding method that extracts the asymmetry A_C at the particle level fiducial region as a parameter of interest in the fit. The parameters associated with the signal and background processes and many *nuisance parameters* associated with the uncertainties are considered in the fit.

The reconstructed distribution of $|y(t)| - |y(\bar{t})|$ for the two signal regions are shown in Figure 11.1. The uncertainties include all statistical and systematic uncertainties as discussed in Chapter 7. The strategy and method of using the profile likelihood unfolding is described in Section 11.1. The treatment for all uncertainties to make them usable in the fit is explained in Section 11.2. A test fit with pseudo data is described in Section 11.3. A test to check the bias in the fit, called *linearity test* is described in Section 11.4. Finally, the measurement of A_C is described in Section 11.5.

11.1 Profile likelihood unfolding method

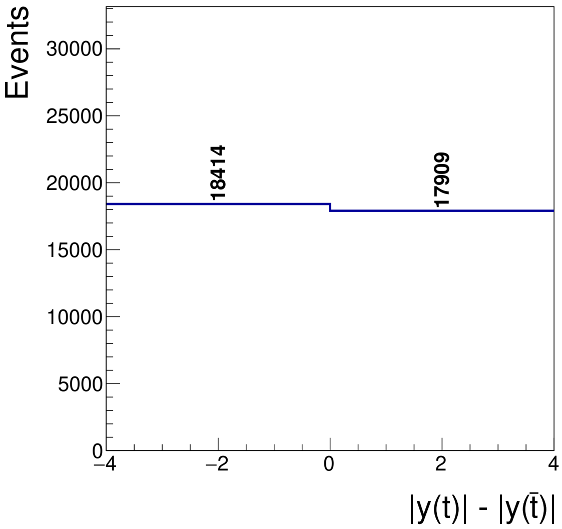
The MC predicted distributions of $|y(t)| - |y(\bar{t})|$ at the reconstruction level in the two signal regions of $O_{\mathrm{NN}} < 0.6$ and $O_{\mathrm{NN}} \geq 0.6$ for the signal and background processes are used to fit the data. A likelihood function is defined for this based on the Poisson probability built from this prediction and data.

The prediction for the signal process is slightly complicated. Although a reconstruction level distribution of the signal process is available, the same can also be expressed in terms of a particle/truth level distribution folded with its response matrices. This approach gives the advantage (over using the traditional method of using the reconstruction level distribution) of directly inferring the particle level distribution from the fit. A profile likelihood fit method [104] is used as described in the following.

11.1.1 Folding of the particle level distribution

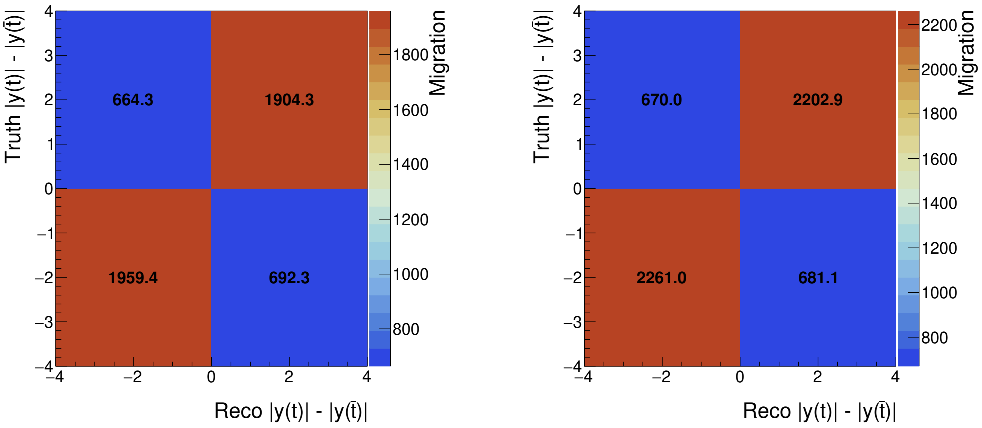
The particle level distribution of $|y(t)| - |y(\bar{t})|$ is obtained after applying the event selection on the particle level objects (as described in Chapter 4.5) for the signal process and shown in Figure 11.2. The event yields of the two bins of this distribution are denoted as T_- and T_+ corresponding to $|y(t)| < |y(\bar{t})|$ and $|y(t)| > |y(\bar{t})|$ respectively.

First, to fold the particle level distribution migration matrices ($\mathbf{M}$) are required. The migration (2×2) matrices contain the yields of the two bins of reconstructed events corresponding to each of the two bins of the particle level distribution. The elements of these matrices can be denoted as the intersection or correspondence of



(a)

Figure 11.2: The particle/truth level distribution of $|y(t)| - |y(\bar{t})|$



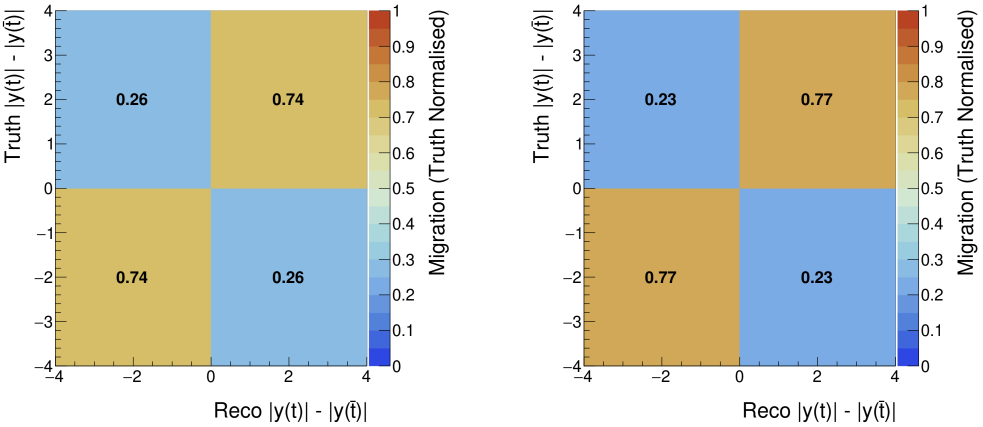
(a) (b)

Figure 11.3: The migration matrices, $N_{r \cap t}$ of the signal of $|y(t)| - |y(\bar{t})|$ in the two regions (a) $O_{\mathrm{NN}} < 0.6$ and (b) $O_{\mathrm{NN}} \geq 0.6$ created from the particle level and reconstruction level events

the reconstructed and truth (particle level) events: $M_{r,t} = N_{r\cap t}$, for reconstructed bin r and truth bin t. They are shown in Figure 11.3.

The migration matrices are not directly used for folding, but via the response matrices. The response matrices (**P**) are defined such that by folding (multiplying) them with the truth **T** gives the reconstructed distribution **R = PT**. This is not only performed for the nominal signal, but also for all systematic variations of the signal (not shown here). Therefore, the uncertainties of the signal comes with the response matrices instead of the reconstructed distributions as in the case of the backgrounds.

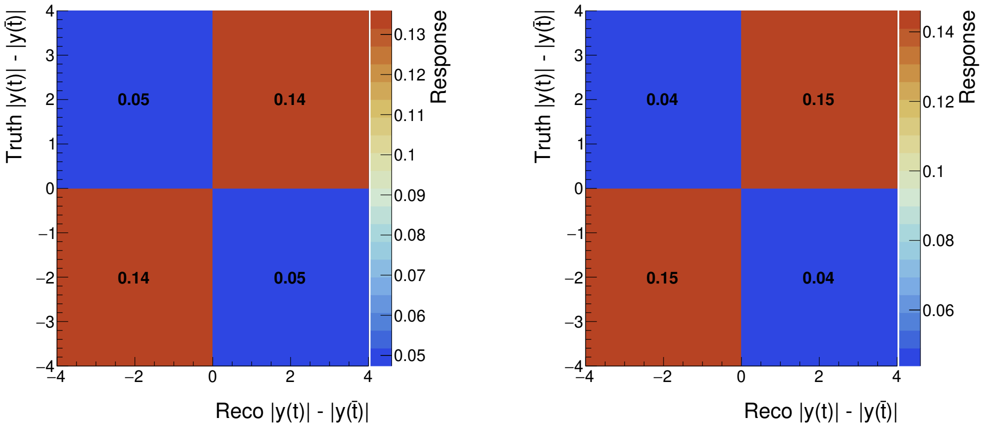
P is defined in terms of the truth normalised migration matrices: $\boldsymbol{M}^{\mathbf{norm}}$ (shown in Figure 11.4), the efficiency of selecting truth events of a particular bin(t), $\boldsymbol{\epsilon_t}$, and the acceptance of selecting reconstructed events of a particular bin(r), $\boldsymbol{f_r}$:



(a) (b)

Figure 11.4: Normalised migration matrices, M of the signal of $|y(t)| - |y(\bar{t})|$ in the two regions (a) $O_{\mathrm{NN}} < 0.6$ and (b) $O_{\mathrm{NN}} \geq 0.6$ created from the truth and reconstructed events

$$\begin{aligned} M^{\mathrm{norm}}_{r,t} &= \frac{N_{r\cap t}}{\sum_r N_{r\cap t}} \\ \epsilon_t &= \frac{\sum_r N_{r\cap t}}{N_t} \\ f_r &= \frac{\sum_t N_{r\cap t}}{N_r} \end{aligned} \tag{11.1}$$



(a) (b)

Figure 11.5: Response matrices, P of the signal of $|y(t)| - |y(\bar{t})|$ in the two regions (a) $O_{\mathrm{NN}} < 0.6$ and (b) $O_{\mathrm{NN}} \geq 0.6$ created from the truth and reconstructed events

giving the

$$\begin{aligned}\text{Response, } P_{r,t} &= \frac{M_{\mathrm{r,t}}{}^{\mathrm{norm}} \times \epsilon_t}{f_r} \\ &= \frac{\frac{N_{r\cap t}}{\sum_r N_{r\cap t}} \times \frac{\sum_r N_{r\cap t}}{N_t}}{\frac{\sum_t N_{r\cap t}}{N_r}} \\ &= N_{r\cap t} \times \frac{N_r}{N_t \times \sum_t N_{r\cap t}}\end{aligned} \tag{11.2}$$

The resulting response matrices are shown in Figure 11.5.

11.1.2 Description of the likelihood function

A likelihood function is defined taking into account the signal and background distributions of $|y(t)| - |y(\bar{t})|$ in different regions as well as all the systematic uncertainties. The likelihood considers a Poisson probability based on the prediction and data and profiled for all systematic uncertainties with Gaussian constraints. A profile likelihood fit [104] method is used with this likelihood. The likelihood can be expressed as:

$$\mathcal{L} = \prod_{\mathrm{c}} \prod_{\mathrm{b}} \mathrm{Pois}(N_{b,c}^{\mathrm{data}} | \nu_{b,c}(\vec{\theta})) \cdot \prod_{p} \mathrm{Gauss}(0|\theta_p, 1), \tag{11.3}$$

where
- c refers to the regions/channels: $O_{\mathrm{NN}} \geq 0.6$ and $O_{\mathrm{NN}} < 0.6$,
- b refers to the two reconstruction level bins: $|y(t)| < |y(\bar{t})|$ and $|y(t)| > |y(\bar{t})|$,
- $N^{\mathrm{data}}_{b,c}$ is the data in bin b and channel c,
- $\vec{\theta}$ are the *nuisance parameters (NP)* corresponding to the systematic uncertainties after following the procedures described in Section 11.2 and $\mathrm{Gauss}(0|\theta_p, 1)$ is a Gaussian constraint of the NP $\theta_p \ \epsilon\ \vec{\theta}$ with a mean of 0 and a standard deviation of 1.
- $\nu_{b,c}$ is the expected total events in bin b and channel c, can be expressed as follows:

$$\nu_{b,c} = \left[\gamma_{b,c,-} \times \mu_- \times R_{b,c,-} \times T_- + \gamma_{b,c,+} \times \mu_+ \times R_{b,c,+} \times T_+\right] + \sum_B \gamma^B_{b,c} \times N^B_{b,c}, \tag{11.4}$$

where
- γ: represents the scale factors (by default 1) for each bin to take into account the MC statistical uncertainties. The background samples each have the factors $\gamma^B_{b,c}$ and for the signal, each truth bin has a unique $\gamma_{b,c,t}$. A total of 8 γ factors for the signal and 20 (4 for each background process) for the backgrounds exist.
- $R_{b,c,t}$ are the response matrices of the signal, calculated from the truth and reconstructed events.
- μ_- and μ_+ are the signal strength corresponding to the truth bins T_- and T_+ having the value of 1 as input before performing the fit.

After the profile likelihood fit is performed, the parameters of interests (POI) (μ_- and μ_+) in the fit get a value as well as all the individual NP θ_p gets their Gaussian distributions updated with a new mean ($\hat{\theta}$) and error ($\Delta\hat{\theta}$), from of the prior mean and error values of 0 (θ_0) and 1 ($\Delta\theta$) respectively.

11.1.3 The strategy for the measurement

By performing a profile likelihood fit with data with the above described likelihood, an inference can be drawn directly for the particle level distribution of $|y(t)| - |y(\bar{t})|$ of the signal process, i.e. the cross-section for the two bins of this distribution can be extracted via the parameters of interest (POI), μ_- and μ_+. The asymmetry, A_C is defined at the particle level in terms of the distribution's yields T_- and T_+ and the signal strengths μ_- and μ_+ as

$$A_C = \frac{\mu_+ T_+ - \mu_- T_-}{\mu_+ T_+ + \mu_- T_-}. \tag{11.5}$$

Thus, μ_- can be expressed in terms of A_C as

$$\mu_- = \mu_+ \times \frac{T_+}{T_-} \times \frac{1 - A_C}{1 + A_C} \tag{11.6}$$

and can be propagated to the likelihood via $\nu_{b,c}$ as

$$\nu_{b,c} = \left[\gamma_{b,c,-} \times R_{b,c,-} \left(\frac{1 - A_C}{1 + A_C} \right) + \gamma_{b,c,+} \times R_{b,c,+} \right] \mu_+ T_+ + \sum_B \gamma^B_{b,c} \times N^B_{b,c}. \tag{11.7}$$

Hence, a profile likelihood fit is performed with this new likelihood definition to obtain the asymmetry A_C directly from the fit, i.e. A_C and μ_+ become the POI of the fit.

11.2 Treatment of the uncertainties

The individual sources of systematic uncertainties are considered to be uncorrelated. The systematic variations are considered as correlated between the regions ($O_{\mathrm{NN}} < 0.6$ and $O_{\mathrm{NN}} \geq 0.6$). A NP with a Gaussian prior is assigned to each systematic uncertainty. The up-and-down variations of the same systematic uncertainty are considered with a single NP. The impact of the systematic uncertainties on the considered distribution is symmetrised in different approaches as described in Section 11.2.1. The sources of systematic uncertainties with negligible impacts are *pruned* and are not considered for the fit as explained in Section 11.2.3. No *Smoothing* technique is used since there are only two bins present in the $|y(t)| - |y(\bar{t})|$ histograms.

11.2.1 Symmetrisation

Since the NP in likelihood is defined with Gaussian constraint, the up and down variations of the corresponding uncertainty are required to be symmetric around the nominal value of a histogram bin, and hence they are symmetrised. If both the up and down variations are provided, i.e. three point uncertainties, two different symmetrisation options are available referred to as the so-called *Two-Sided* and *Abs-Mean*. In the case of single variation uncertainties, i.e. two-point uncertainties, a

One-Sided symmetrisation option is used.

Two-Sided

In this approach, the difference between the up and down variations is shifted such that they become symmetric around the nominal value with their original signs intact. The '+' and '-' of the following expression are then taken as symmetrised up and down variations, respectively:

$$\text{symmetrised up/down} = \text{nominal} \pm \frac{(\text{up} - \text{down})}{2} \tag{11.8}$$

Abs-Mean

In this alternate approach, the average of the differences of the up/down with respect to the nominal are calculated and that average difference is mirrored around the nominal value:

$$\text{symmetrised up/down} = \text{nominal} \pm \left(\frac{|\text{up} - \text{nominal}| + |\text{down} - \text{nominal}|}{2} \right) \tag{11.9}$$

This symmetrisation technique provides a fixated direction of the up and down variations being always larger and smaller than the nominal value respectively. Even in those cases, where both variations would go in the same direction in a bin, e.g. due to large statistical fluctuations, *AbsMean* symmetrisation provides a symmetrised uncertainty. In those cases, it is superior to the *Two-sided* symmetrisation, as that would potentially underestimate the variation in those bins. Thus, in the case of two opposite side up and down variations with the up value being larger than the nominal, *AbsMean* symmetrisation would act exactly as *Two-sided* symmetrisation. The *AbsMean* symmetrisation is used for those variations with large statistical fluctuations, in particular in the following analysis it has been used only for the modelling uncertainties on the radiation of the signal process as described later.

One-Sided

In the case of a single variation uncertainty, the variation is mirrored around the nominal to obtain the symmetrised variation.

A summarised list of all the systematic uncertainties with their NPs and the corresponding symmetrisation options used are shown in Table 11.1 and Table 11.2.

A detailed explanation of the individual uncertainties is given in Chapter 7.

Table 11.1: The summary list of all theoretical systematic uncertainties with their number of individual NPs corresponding to the group origin and the symmetrisation option used for them

Systematic Uncertainty	Symmetrisation	Number of NPs
$t\bar{t}\gamma$ production μ_R	Two-sided	1
$t\bar{t}\gamma$ production μ_F	Two-sided	1
$t\bar{t}\gamma$ production PDFs	One-sided	30
$t\bar{t}\gamma$ production PS model	One-sided	1
$t\bar{t}\gamma$ production PS radiation	AbsMean	1
$t\bar{t}\gamma$ decay μ_R	Two-sided	1
$t\bar{t}\gamma$ decay μ_F	Two-sided	1
$t\bar{t}\gamma$ decay PDFs	One-sided	30
$t\bar{t}\gamma$ decay PS model	One-sided	1
$t\bar{t}\gamma$ decay PS radiation	Two-sided	1
$t\bar{t}$ μ_R	Two-sided	1
$t\bar{t}$ μ_F	Two-sided	1
$t\bar{t}$ PS model	One-sided	1
$t\bar{t}$ Generator model	One-sided	1
$t\bar{t}$ Simulation	One-sided	1
$t\bar{t}$ PS radiation	Two-sided	1
$t\bar{t}$ *hdamp*	One-sided	1
$tW\gamma$ μ_R	Two-sided	1
$tW\gamma$ μ_F	Two-sided	1
$tW\gamma$ PS model	One-sided	1
Background Normalisation Uncertainties		
$t\bar{t}\gamma$ decay	One-sided	1
$W\gamma$	One-sided	1
$Z\gamma$	One-sided	1
$tW\gamma$	One-sided	1
other smaller processes	One-sided	1

Table 11.2: The summary list of all experimental systematic uncertainties with their number of individual NPs corresponding to the group origin and the symmetrisation option used for them

Systematic Uncertainty	Symmetrisation	Number of NPs
Electron efficiency	Two-sided	4
Photon efficiency	Two-sided	2
e-γ energy scale	Two-sided	2
e-γ energy resolution	Two-sided	1
Muon efficiency	Two-sided	10
Muon track resolution	Two-sided	2
Muon energy scale	Two-sided	1
Muon energy resolution	Two-sided	2
Jet Energy Scale		
– pileup	Two-sided	4
– η-intercalibration	Two-sided	6
– flavour related	Two-sided	3
– high-p_{T}	Two-sided	1
– *Punch Through*	Two-sided	1
– simulation non-closure	Two-sided	1
– calibration eigenvectors	Two-sided	15
Jet Energy Resolution		
– calibration eigenvectors	Two-sided	7
– data/MC offset	Two-sided	1
B-tagging		
– b-jets related eigenvectors	Two-sided	9
– c-jets related eigenvectors	Two-sided	4
– light-jets related eigenvectors	Two-sided	4
Jet vertex tagger	Two-sided	1
MET (soft) scale	Two-sided	1
MET (soft) resolution	One-sided	2
Pileup	Two-sided	1
Integrated Luminosity	Two-sided	1
Electron fake γ	Two-sided	1
Hadron fake γ	Two-sided	1
Fake lepton	Two-sided	2

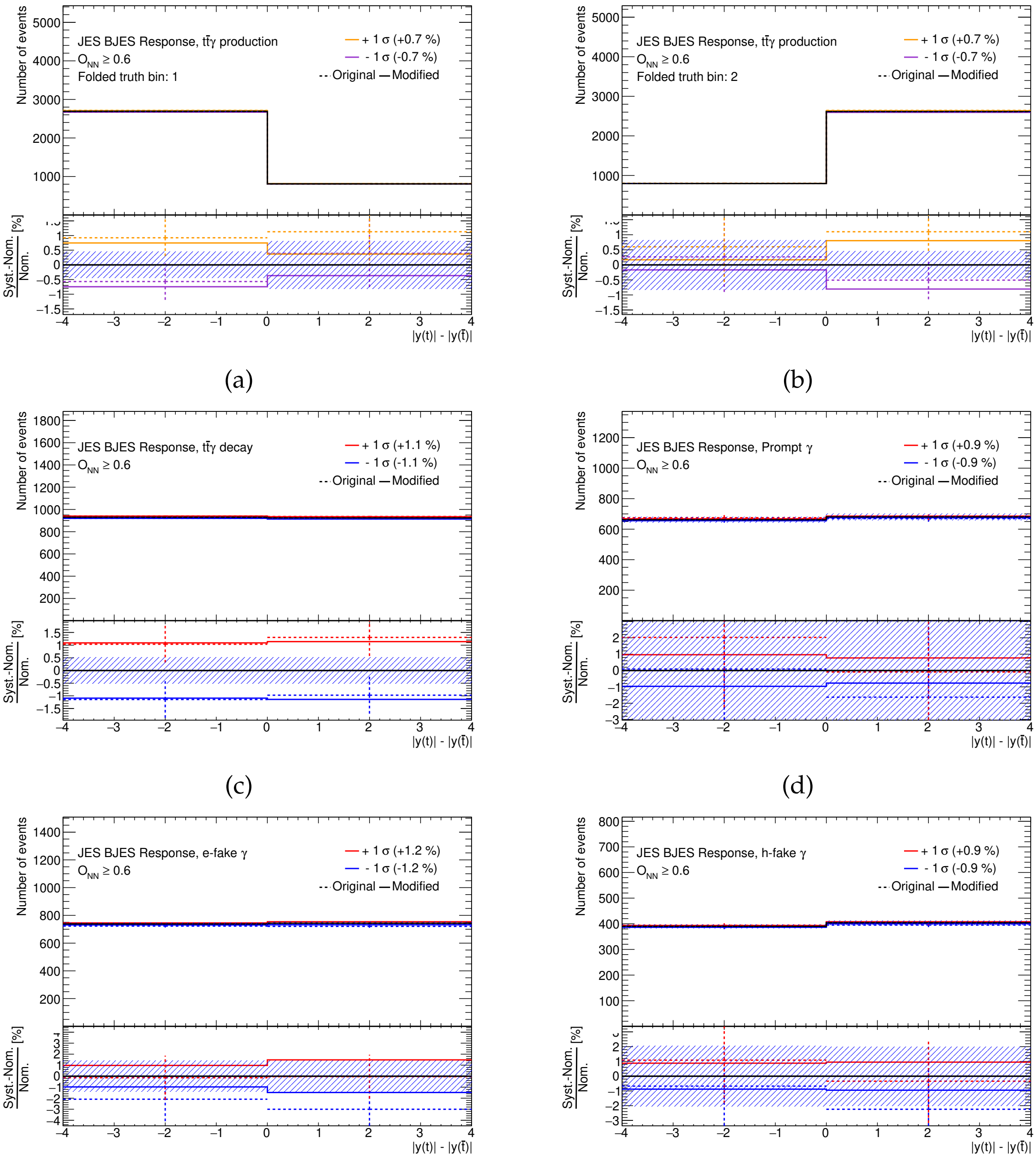
11.2.2 Effect of symmetrisation

A set of example plots are shown in Figure 11.6 for one source of systematic uncertainty (b-jet flavour uncertainty for jet energy scale) for the signal region with $O_{NN} \geq 0.6$ for the signal and all the background processes, where a comparison of the up and down variations with the nominal distribution and the effect of the Two-sided symmetrisation are shown. The signal distributions are found by folding the particle/truth level distribution and results in two separate plots because of the two truth bins T_- and T_+. This is explained in Section 11.1.

11.2.3 Pruning

Since there are many uncertainties present that have a negligible impact on this analysis, a *pruning* procedure is followed to prune them making the fit computationally faster. The pruning is applied separately for the *shape* and normalisation difference between a variation and its corresponding nominal histograms. For pruning based on normalisation, the relative difference of the integral of the distributions is checked to be above the specified threshold value to keep the variation. For pruning based on shape, the relative difference for any bin of the distribution is checked to be above the specified threshold value to keep the variation. The threshold values of 0.1% are used for both the shape and normalisation based pruning. The list of pruned and kept uncertainties are listed in Figures 11.7-11.10. For a better visibility, they are divided into four parts and explained with labels in Figure 11.7. The different colours indicate which aspects of the systematic variation are kept or pruned:

- green for both normalisation and shape kept
- yellow for normalisation only (shape pruned)
- orange for shape only (normalisation pruned)
- red when pruned entirely



(a) (b)

(c) (d)

(e) (f)

Figure 11.6: The variations (up and down) on $|y(t)| - |y(\bar{t})|$ for systematic uncertainty of b-jet flavour for jet energy scale are shown and compared to the nominal distribution for all the processes in the signal region with $O_{NN} \geq 0.6$. The original variations are shown with dotted lines and the symmetrised variations are shown with solid lines. The plot corresponds to the processes (a) and (b) folded signal distributions corresponding to T_{-} (truth bin 1) and T_{+} (truth bin 2) respectively, (c) tt̄γ decay, (d) prompt γ, (e) e-fake γ and (f) h-fake γ.

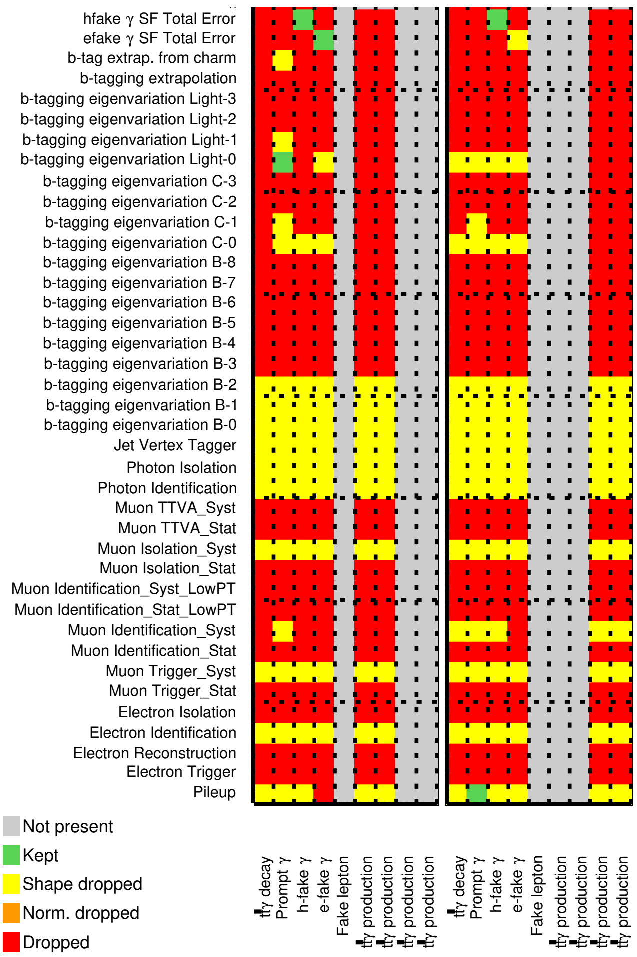


Figure 11.7: The nuisance parameters (1) considered for normalisation and shape uncertainties and whether they are kept or pruned in the fit. The left set of columns belongs to the $O_{NN} < 0.6$ region and the right set of columns belongs to the $O_{NN} \geq 0.6$ region. For the signal, the individual truth bin folded contributions are shown for both regions, where two columns correspond to the same region and redundantly the other two columns correspond to the other region - being empty

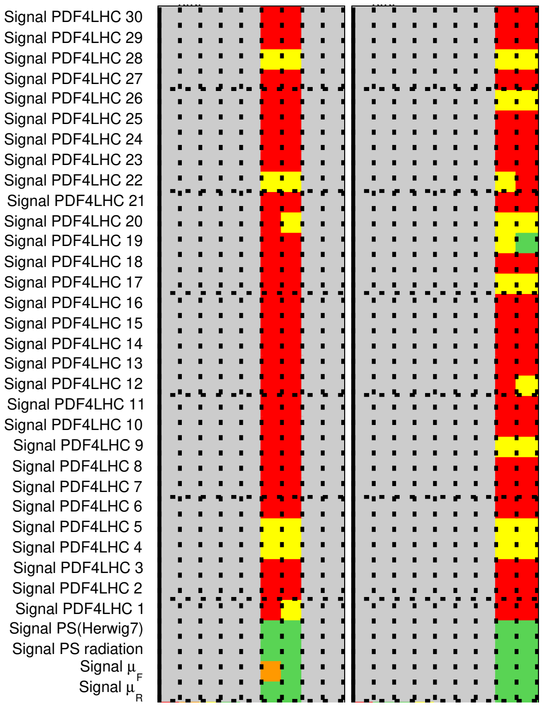


Figure 11.8: The nuisance parameters (2) considered for normalisation and shape uncertainties and whether they are kept or pruned in the fit. The labels of the processes are the same as in Figure 11.7. The left set of columns belong to the $O_{NN} < 0.6$ region and the right set of columns belong to the $O_{NN} \geq 0.6$ region.

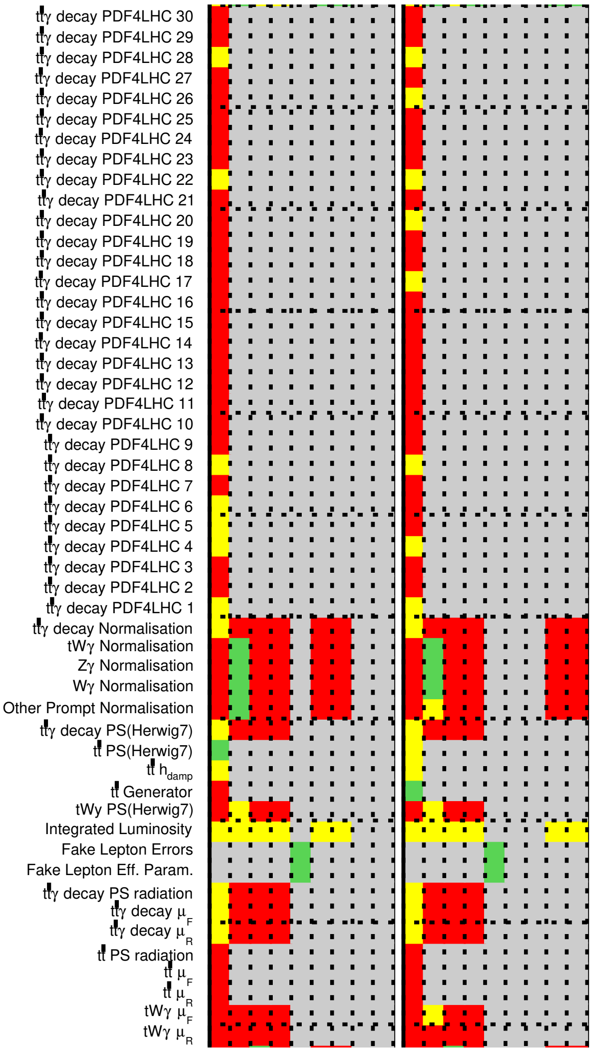


Figure 11.9: The nuisance parameters (3) considered for normalisation and shape uncertainties and whether they are kept or pruned in the fit. The labels of the processes are the same as in Figure 11.7. The left set of columns belong to the $O_{NN} < 0.6$ region and the right set of columns belong to the $O_{NN} \geq 0.6$ region.

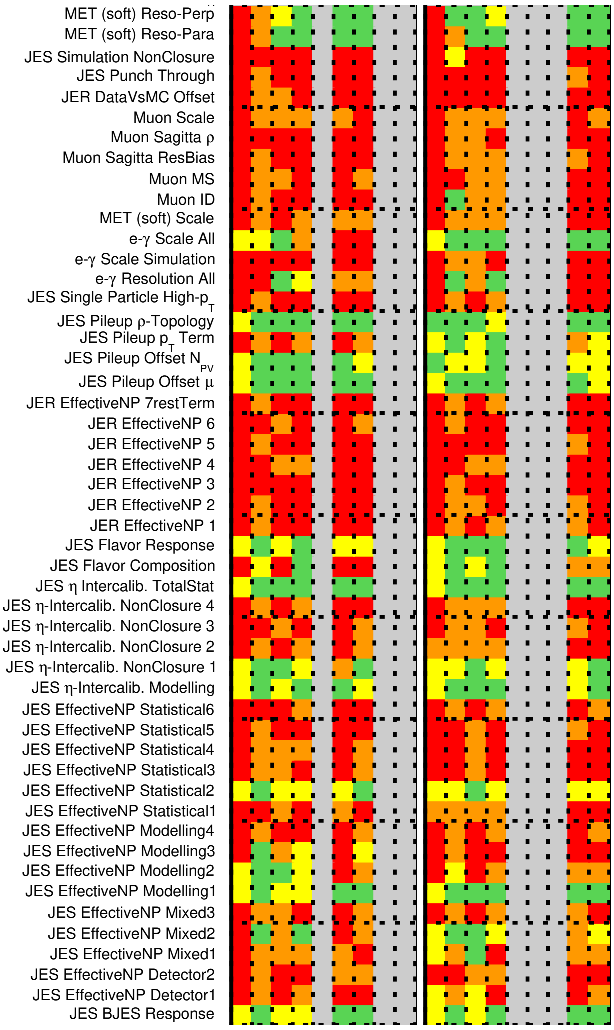


Figure 11.10: The nuisance parameters (4) considered for normalisation and shape uncertainties and whether they are kept or pruned in the fit. The labels of the processes are the same as in Figure 11.7. The left set of columns belong to the $O_{NN} < 0.6$ region and the right set of columns belong to the $O_{NN} \geq 0.6$ region.

11.3 Fit with Asimov data

Before the actual fit is performed on data, the stability of the fit procedure, the behaviour of the systematic uncertainties and the expected sensitivity of the measurement are evaluated using a test fit by creating a pseudo data template called *Asimov*. The pseudo data corresponds to the sum of the predicted events of the signal and all the background processes. Hence, the (pseudo) data is exactly the same as the prediction, i.e. a *perfect fit*.

The fit with Asimov data results with the NPs fitted to the same mean value of 0, i.e. does not get *pulled* away, however, can get constrained, i.e. $\Delta\hat{\theta} < 1$ standard deviation. The fitted γ-factors' uncertainty reflects the MC statistical uncertainty. The pull (shift of $\hat{\theta}$ from 0) and the constraints of all the NPs and the fitted γ-factors are shown in the Figure 11.11 and 11.12, where it can be noted that the green region corresponds to $\Delta\hat{\theta} < 1$ standard deviation and the yellow region corresponds to $\Delta\hat{\theta} < 2$ standard deviation with zero mean value.

Only a few NPs show constraints larger than 5%. The NP corresponding to the normalisation uncertainty of $t\bar{t}\gamma$ decay is constrained to 71% and of $W\gamma$ is constrained to 87%. Among the NPs corresponding to the experimental systematic uncertainties, JES Pileup ρ-Topology and h-fake γ SF error gets constrained to around 94%. The rest of the NPs get very minimal constraints. This means that for the number of expected events in each bin in the fitted distribution – their *a-priori* systematic uncertainties (like the background normalisation uncertainties) were estimated to be larger, or they could be constrained through the (anti-)correlation with other systematic uncertainties.

The γ factors are fitted to be near the value of 1 and their uncertainties are mostly infinitesimally small except for the lepton fake background since it has small statistics.

A fit with Asimov data results in NPs getting correlated between themselves also. The correlation between the NPs among themselves and also with the POIs: charge asymmetry and μ_+ (named in the plot as *Unfolding Unfolded Truth Bin 2*) are determined during the fit and shown in the Figure 11.13. The figure shows only those NPs for which the correlation with at least another NP or POI is more than 10%. The significant correlations are mainly present for the constrained NP of $t\bar{t}\gamma$ decay normalisation uncertainty with various jet-related uncertainties and $W\gamma$ normalisation uncertainty. Although μ_+ have large correlations, A_C does not have such correlations.

11.3.1 Results with Asimov fit

The fit with the pseudo data gives the asymmetry to be

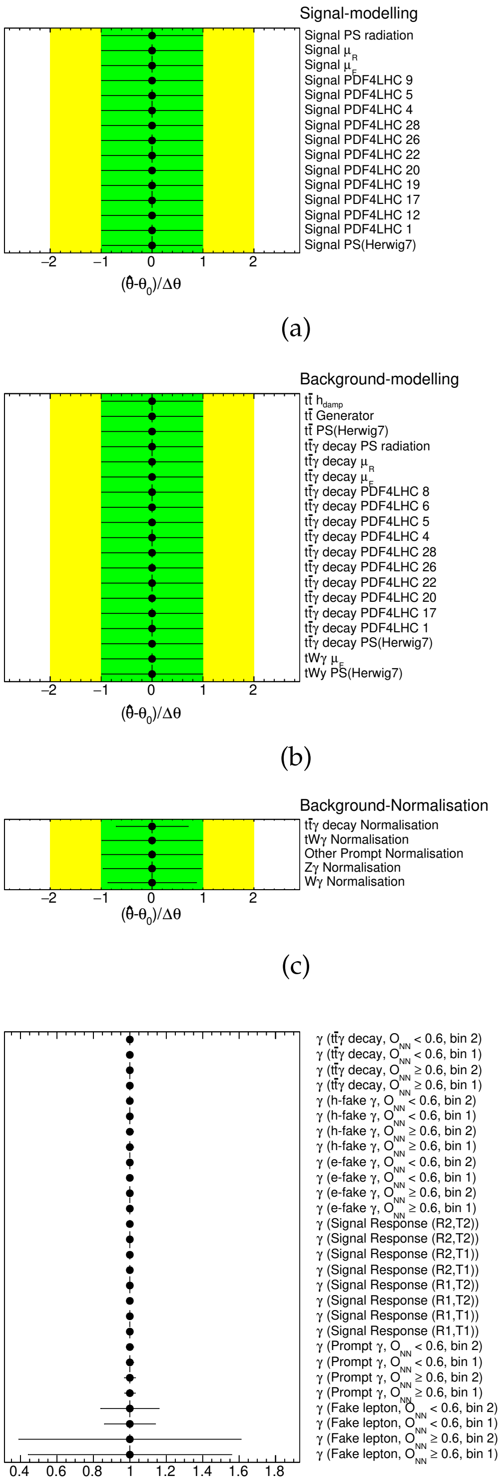
$$\begin{aligned} A_C &= -0.014 \pm 0.033 \\ &= -0.014 \pm 0.027(\text{stat}) \pm 0.018(\text{syst}) \end{aligned}$$

which is expected. The central value is the same as the expected asymmetry from the signal MC at the particle level in the fiducial phase space as described in Chapter 4.5.

To estimate the impact of individual uncertainties on the measurement of the A_C, a series of repeated fits are done individually by keeping the value of the corresponding NP (θ) fixed within the upper and lower range, i.e. ± 1 at pre-fit and the pulled/constrained values following the fit (post-fit) i.e. $\hat{\theta} \pm \Delta\hat{\theta}$. The resulting ranking plot is shown in Figure 11.14. In this way, the difference between these newly fitted A_C values and the nominal fit value is considered as the impact of the corresponding NP on the measurement. However, these are only estimates to assess the impact of each NP, but not their contribution to the final error.

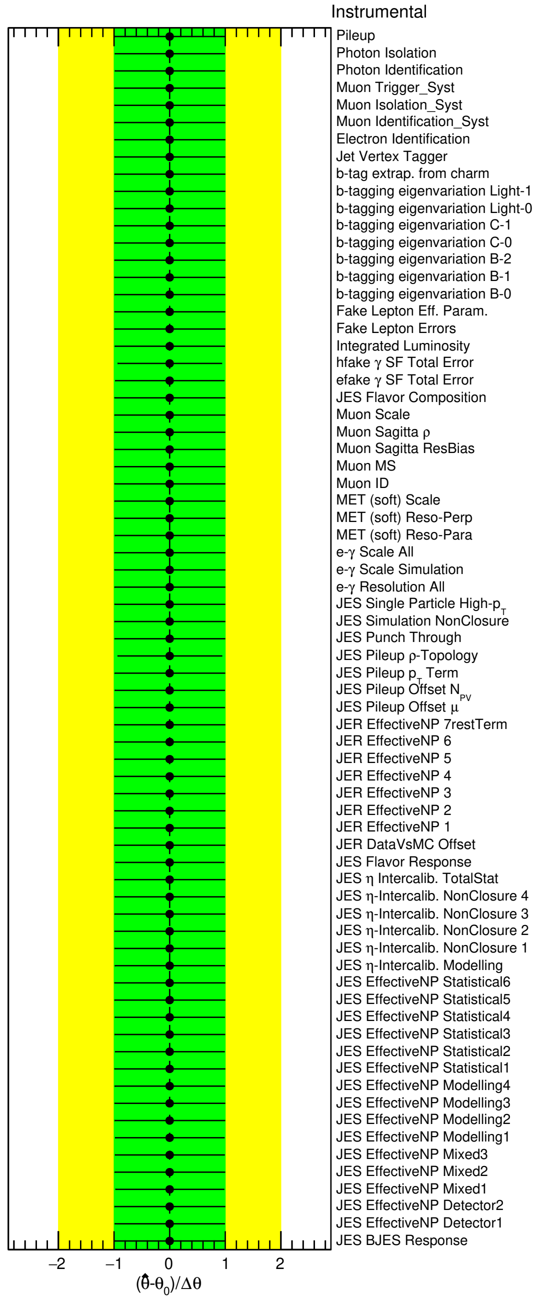
The two most important systematic uncertainties are, one component of MET (soft) resolution uncertainty and the JES flavour response uncertainty. The comparison of these uncertainties with their nominal distributions is shown in Appendix A. The MC statistical uncertainties for the prompt γ and fake lepton backgrounds are also important.

Along with the fit on A_C, the inference on μ_+ (hence also μ_-) can also be achieved that gives the resulting absolute cross-sections of the signal extracted at the particle level fiducial region in the two bins of $|y(t)| - |y(\bar{t})|$ as shown in the Figure 11.15. The unfolded differential cross-sections have around 10% uncertainty on each bin. The unfolded Asimov data is compared with the MC predictions using two different parton shower and hadronisation algorithms: Pythia and Herwig, and they agree well within the uncertainty as per expectation.



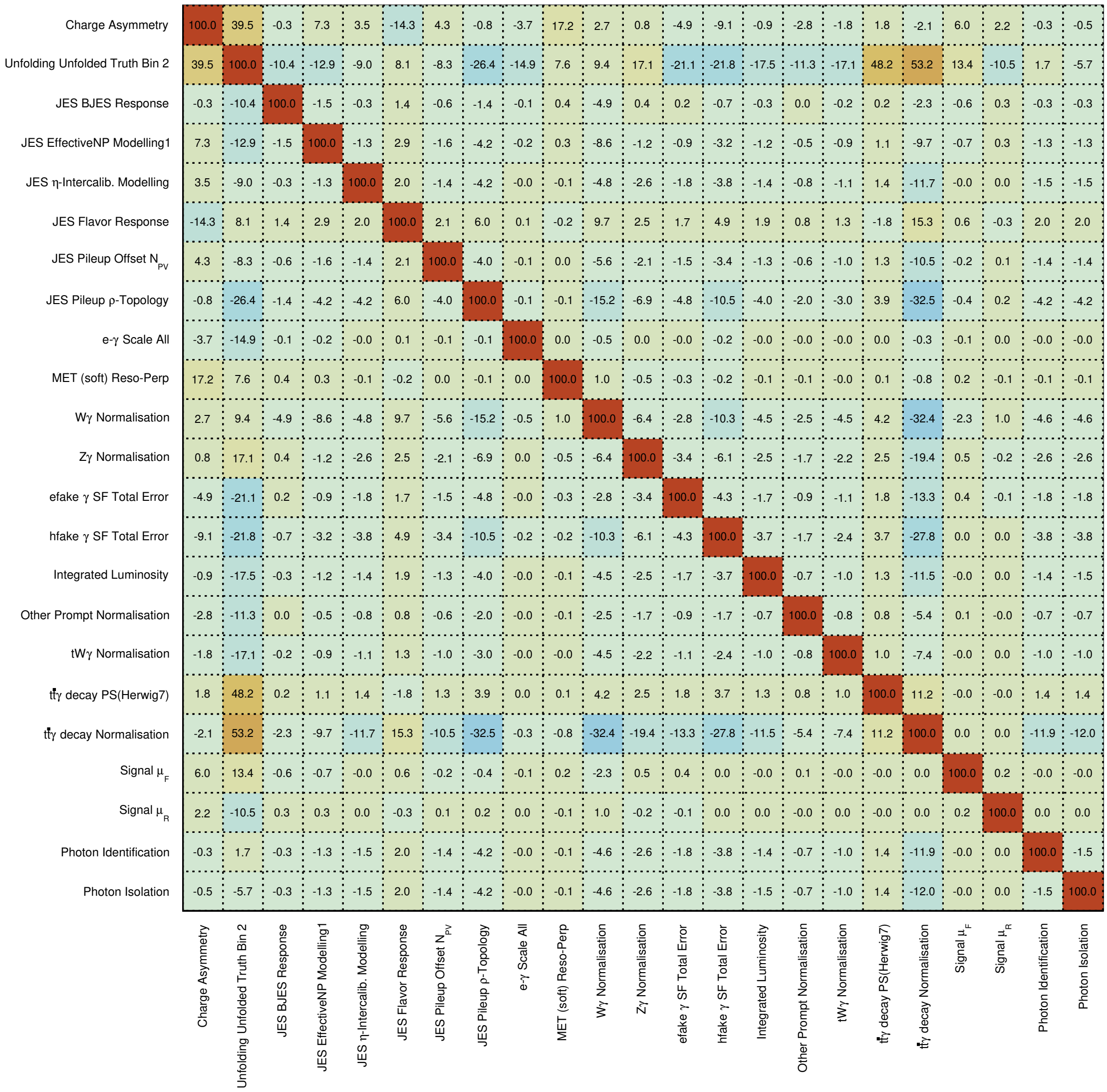
(d)

Figure 11.11: (a,b,c) The fitted NPs ($\hat{\theta}$) (1) with their pulls and constraints with respect to their pre-fit values and (d) Gamma Parameters after the fit with Asimov data.



(a)

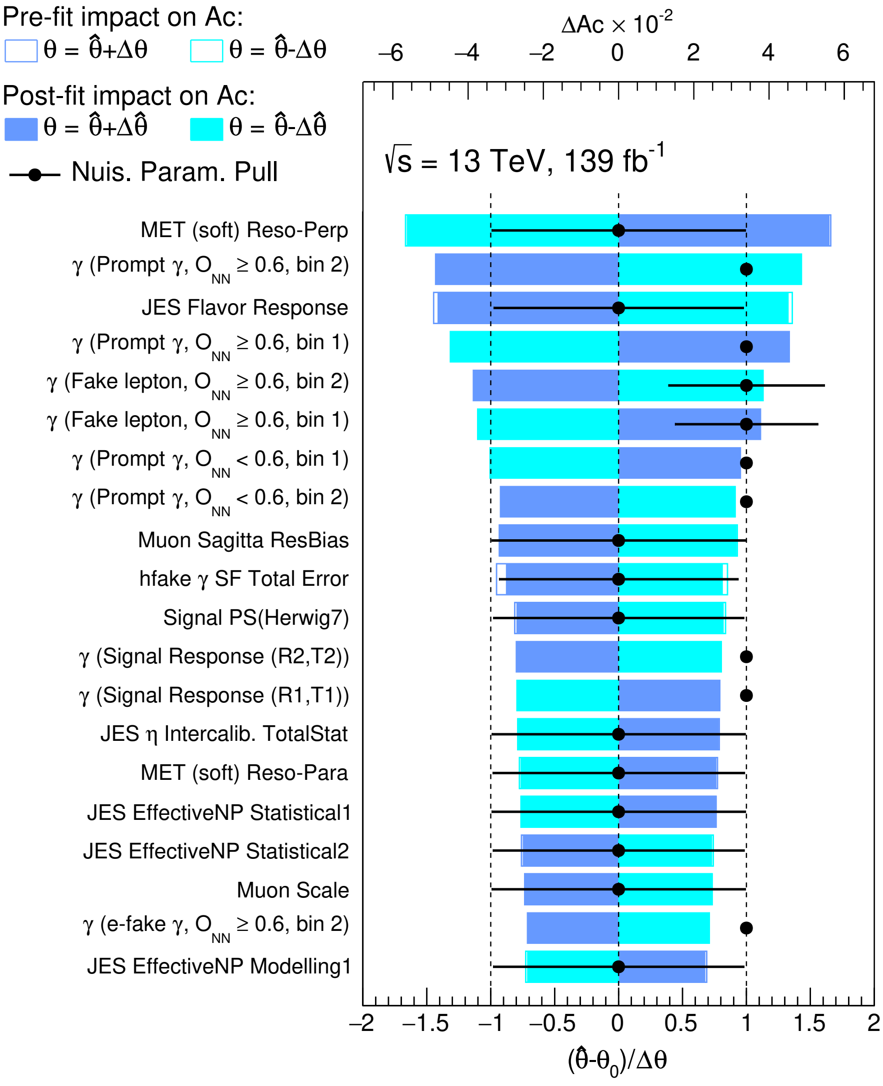
Figure 11.12: The fitted NPs ($\hat{\theta}$) (2) with their pulls and constraints with respect to their pre-fit values after the fit with Asimov data.



	Charge Asymmetry	Unfolding Unfolded Truth Bin 2	JES BJES Response	JES EffectiveNP Modelling1	JES η-Intercalib. Modelling	JES Flavor Response	JES Pileup Offset N_{PV}	JES Pileup ρ-Topology	e-γ Scale All	MET (soft) Reso-Perp	Wγ Normalisation	Zγ Normalisation	efake γ SF Total Error	hfake γ SF Total Error	Integrated Luminosity	Other Prompt Normalisation	tWγ Normalisation	$t\bar{t}\gamma$ decay PS(Herwig7)	$t\bar{t}\gamma$ decay Normalisation	Signal μ_F	Signal μ_R	Photon Identification	Photon Isolation
Charge Asymmetry	100.0	39.5	-0.3	7.3	3.5	-14.3	4.3	-0.8	-3.7	17.2	2.7	0.8	-4.9	-9.1	-0.9	-2.8	-1.8	1.8	-2.1	6.0	2.2	-0.3	-0.5
Unfolding Unfolded Truth Bin 2	39.5	100.0	-10.4	-12.9	-9.0	8.1	-8.3	-26.4	-14.9	7.6	9.4	17.1	-21.1	-21.8	-17.5	-11.3	-17.1	48.2	53.2	13.4	-10.5	1.7	-5.7
JES BJES Response	-0.3	-10.4	100.0	-1.5	-0.3	1.4	-0.6	-1.4	-0.1	0.4	-4.9	0.4	0.2	-0.7	-0.3	0.0	-0.2	0.2	-2.3	-0.6	0.3	-0.3	-0.3
JES EffectiveNP Modelling1	7.3	-12.9	-1.5	100.0	-1.3	2.9	-1.6	-4.2	-0.2	0.3	-8.6	-1.2	-0.9	-3.2	-1.2	-0.5	-0.9	1.1	-9.7	-0.7	0.3	-1.3	-1.3
JES η-Intercalib. Modelling	3.5	-9.0	-0.3	-1.3	100.0	2.0	-1.4	-4.2	-0.0	-0.1	-4.8	-2.6	-1.8	-3.8	-1.4	-0.8	-1.1	1.4	-11.7	-0.0	0.0	-1.5	-1.5
JES Flavor Response	-14.3	8.1	1.4	2.9	2.0	100.0	2.1	6.0	0.1	-0.2	9.7	2.5	1.7	4.9	1.9	0.8	1.3	-1.8	15.3	0.6	-0.3	2.0	2.0
JES Pileup Offset N_{PV}	4.3	-8.3	-0.6	-1.6	-1.4	2.1	100.0	-4.0	-0.1	0.0	-5.6	-2.1	-1.5	-3.4	-1.3	-0.6	-1.0	1.3	-10.5	-0.2	0.1	-1.4	-1.4
JES Pileup ρ-Topology	-0.8	-26.4	-1.4	-4.2	-4.2	6.0	-4.0	100.0	-0.1	-0.1	-15.2	-6.9	-4.8	-10.5	-4.0	-2.0	-3.0	3.9	-32.5	-0.4	0.2	-4.2	-4.2
e-γ Scale All	-3.7	-14.9	-0.1	-0.2	-0.0	0.1	-0.1	-0.1	100.0	0.0	-0.5	0.0	-0.0	-0.2	-0.0	-0.0	-0.0	0.0	-0.3	-0.1	0.0	-0.0	-0.0
MET (soft) Reso-Perp	17.2	7.6	0.4	0.3	-0.1	-0.2	0.0	-0.1	0.0	100.0	1.0	-0.5	-0.3	-0.2	-0.1	-0.1	-0.0	0.1	-0.8	0.2	-0.1	-0.1	-0.1
Wγ Normalisation	2.7	9.4	-4.9	-8.6	-4.8	9.7	-5.6	-15.2	-0.5	1.0	100.0	-6.4	-2.8	-10.3	-4.5	-2.5	-4.5	4.2	-32.4	-2.3	1.0	-4.6	-4.6
Zγ Normalisation	0.8	17.1	0.4	-1.2	-2.6	2.5	-2.1	-6.9	0.0	-0.5	-6.4	100.0	-3.4	-6.1	-2.5	-1.7	-2.2	2.5	-19.4	0.5	-0.2	-2.6	-2.6
efake γ SF Total Error	-4.9	-21.1	0.2	-0.9	-1.8	1.7	-1.5	-4.8	-0.0	-0.3	-2.8	-3.4	100.0	-4.3	-1.7	-0.9	-1.1	1.8	-13.3	0.4	-0.1	-1.8	-1.8
hfake γ SF Total Error	-9.1	-21.8	-0.7	-3.2	-3.8	4.9	-3.4	-10.5	-0.2	-0.2	-10.3	-6.1	-4.3	100.0	-3.7	-1.7	-2.4	3.7	-27.8	0.0	0.0	-3.8	-3.8
Integrated Luminosity	-0.9	-17.5	-0.3	-1.2	-1.4	1.9	-1.3	-4.0	-0.0	-0.1	-4.5	-2.5	-1.7	-3.7	100.0	-0.7	-1.0	1.3	-11.5	-0.0	0.0	-1.4	-1.5
Other Prompt Normalisation	-2.8	-11.3	0.0	-0.5	-0.8	0.8	-0.6	-2.0	-0.0	-0.1	-2.5	-1.7	-0.9	-1.7	-0.7	100.0	-0.8	0.8	-5.4	0.1	-0.0	-0.7	-0.7
tWγ Normalisation	-1.8	-17.1	-0.2	-0.9	-1.1	1.3	-1.0	-3.0	-0.0	-0.0	-4.5	-2.2	-1.1	-2.4	-1.0	-0.8	100.0	1.0	-7.4	-0.0	0.0	-1.0	-1.0
$t\bar{t}\gamma$ decay PS(Herwig7)	1.8	48.2	0.2	1.1	1.4	-1.8	1.3	3.9	0.0	0.1	4.2	2.5	1.8	3.7	1.3	0.8	1.0	100.0	11.2	-0.0	-0.0	1.4	1.4
$t\bar{t}\gamma$ decay Normalisation	-2.1	53.2	-2.3	-9.7	-11.7	15.3	-10.5	-32.5	-0.3	-0.8	-32.4	-19.4	-13.3	-27.8	-11.5	-5.4	-7.4	11.2	100.0	0.0	0.0	-11.9	-12.0
Signal μ_F	6.0	13.4	-0.6	-0.7	-0.0	0.6	-0.2	-0.4	-0.1	0.2	-2.3	0.5	0.4	0.0	-0.0	0.1	-0.0	-0.0	0.0	100.0	0.2	-0.0	-0.0
Signal μ_R	2.2	-10.5	0.3	0.3	0.0	-0.3	0.1	0.2	0.0	-0.1	1.0	-0.2	-0.1	0.0	0.0	-0.0	0.0	-0.0	0.0	0.2	100.0	0.0	0.0
Photon Identification	-0.3	1.7	-0.3	-1.3	-1.5	2.0	-1.4	-4.2	-0.0	-0.1	-4.6	-2.6	-1.8	-3.8	-1.4	-0.7	-1.0	1.4	-11.9	-0.0	0.0	100.0	-1.5
Photon Isolation	-0.5	-5.7	-0.3	-1.3	-1.5	2.0	-1.4	-4.2	-0.0	-0.1	-4.6	-2.6	-1.8	-3.8	-1.5	-0.7	-1.0	1.4	-12.0	-0.0	0.0	-1.5	100.0

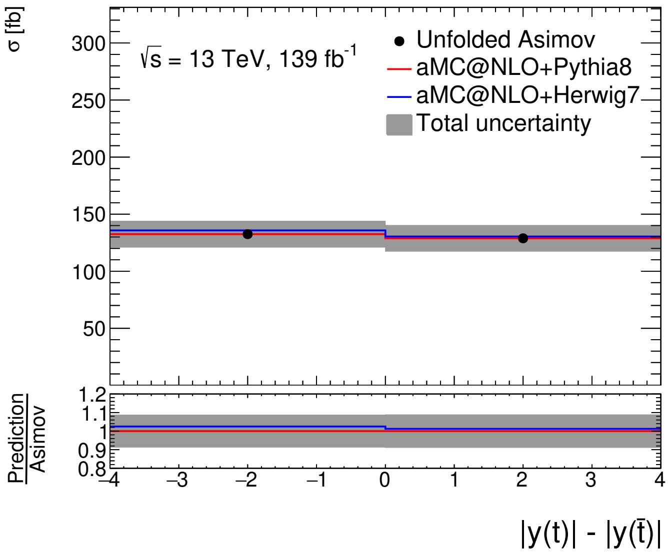
(a)

Figure 11.13: Correlations between different fitted NPs and the parameters of interest for the Asimov fit. The plot shows only those NPs for which the correlation with at least another NP or POI is more than 10%. The POIs A_C and μ_+ are named here as *Charge Asymmetry* and *Unfolding Unfolded Truth Bin 2* respectively.



(a)

Figure 11.14: The ranking of the different NPs according to their impact on the measurement of the charge asymmetry, A_C for the fit with Asimov data. The top scale corresponds to the impact of the NPs on the measurement of A_C, whereas the bottom scale corresponds to the fitted NPs with their pulls and constraints same as in the Figures 11.11 and 11.12.



(a)

Figure 11.15: The unfolded distribution of $|y(t)| - |y(\bar{t})|$ at the particle level with its uncertainty corresponding to the Asimov data and comparison to the MC predicted distributions with different parton showers

11.4 Linearity test

A linearity test is performed to probe the robustness of the unfolding method such that the setup does not bias the result towards the asymmetry of the signal MC and for the fit performed with data the unfolded asymmetry is the same as its true value. For the test, the unfolding procedure is repeated by using a varied signal distribution to form a pseudo-data. The unfolded asymmetry obtained with a perfectly unbiased and uncorrelated setup should be exactly the same as its true value. The test is performed with multiple different asymmetry values. A linear relation is expected between the true asymmetry and the unfolded asymmetry values (with slope **m**, offset **c**) such that

$$\text{unfolded asymmetry} = m \times \text{true asymmetry} + c \tag{11.10}$$

To mimic different true asymmetry values, the $|y(t)| - |y(\bar{t})|$ distributions from nominal signal MC are varied using linear reweighting method, similar to the method used for $t\bar{t}\gamma$ decay background reweighting as discussed in Chapter 9.5. The equations 9.13-9.20 are used again in this case replacing A_C^{LO} by the nominal $t\bar{t}\gamma$ production truth (particle level) asymmetry and A_C^{NLO} by the modified truth asymmetry. Here the migration matrices **M** correspond to the same matrices shown in Figure 11.3. As a result the newly reconstructed distributions, with yields R_1 and R_2, of the corresponding modified truth asymmetry are used for creating the pseudo data for the test fits.

11.4.1 Test results

Three different linearity tests are performed using the linear reweighting method.

Test 1: closure test

This test is done using the full statistics of the signal MC sample as pseudo data. The test values for the varied truth asymmetry are chosen uniformly in the range of [-0.060,+0.030] with 19 values around the expected A_C covering approximately a range of $\pm 1.3\sigma$ around it, where σ corresponds to the expected uncertainty. The test values are fitted well linearly with the unfolded values and they are shown in Figure 11.16 (a). From the linear fit, the slope $m = 0.9824$ and offset $c = -0.0002$.

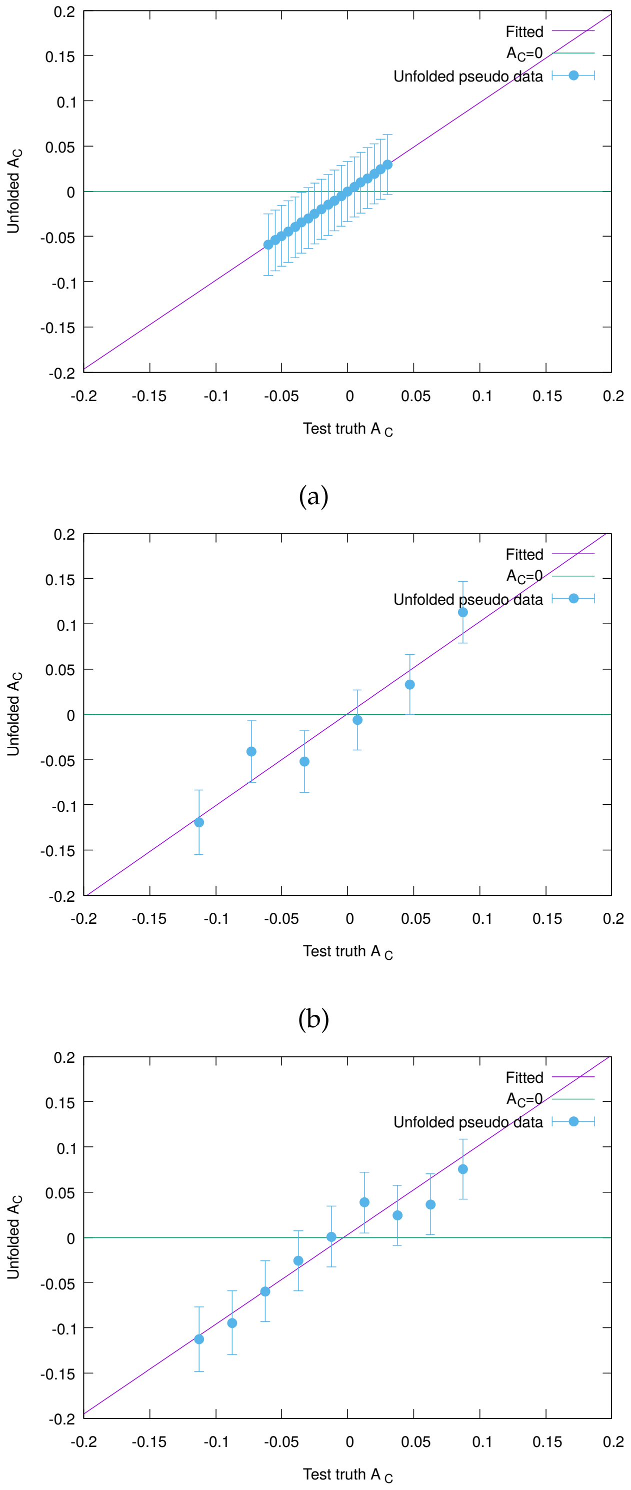
Test 2: closure test with partial dataset

The closure test is repeated using a partially available dataset for creating the pseudo data. The 6 test values cover approximately the range of $\pm 3\sigma$ around the expected A_C value: -0.1128, -0.0728, -0.0328, 0.0072, 0.0472, 0.0872. The test values are fitted well linearly with the unfolded values and they are shown in Figure 11.16 (c). From the linear fit, the slope $m = 1.0168 \pm 0.1498$ and offset $c = 0.0007 \pm 0.0103$.

Test 3: independent datasets

A third test is done by creating 9 independent MC samples and using them independently for corresponding 9 test values of asymmetry. The 9 test values cover approximately the range of $\pm 3\sigma$ around the expected A_C value: -0.1125, -0.0875, -0.0625, -0.0375, -0.0125, 0.0125, 0.0375, 0.0625, 0.0875. The unfolded asymmetry agrees well with their corresponding test values and they fit well linearly as shown in Figure 11.16 (b) with the slope $m = 0.9918 \pm 0.0765$ and offset $c = 0.0031 \pm 0.0051$.

All three tests show a strong linear relation between the truth asymmetry value of the pseudo data and the corresponding unfolded asymmetry value. The slope and offset values show a negligible bias within their uncertainties.



(a)

(b)

(c)

Figure 11.16: Linearity test of the unfolding setup using (a) full dataset of signal MC (b) partial dataset and (c) independent individual datasets

11.5 Measurement of the charge asymmetry

After the test fit with Asimov data, the experimental data is used instead, keeping everything else of the fit setup the same for the final measurement of A_C. The fitted NPs for this fit are shown in Figures 11.18 and 11.19 with their pulls and constraints along with the γ parameters. From the fit the signal strength μ_+ is also measured along with A_C and found to be $1.20\,^{+10\%}_{-9\%}$.

Few NPs, e.g. the normalisation uncertainties of Wγ and t$\bar{\text{t}}\gamma$ decay are slightly pulled away from 0 and also get constrained as expected from the fit with the Asimov data. All NPs are within one standard deviation of their pre-fit values.

The correlation matrix for the fit is shown in Figure 11.20. The correlations are quite similar to the same for the fit with Asimov data (Figure 11.13).

11.5.1 Results

The asymmetry A_C is measured to be [3]

$$\begin{aligned}\mathbf{A_C} &= \mathbf{-0.003 \pm 0.029}\\ &= -0.003 \pm 0.024(\text{stat}) \pm 0.017(\text{syst})\end{aligned}$$

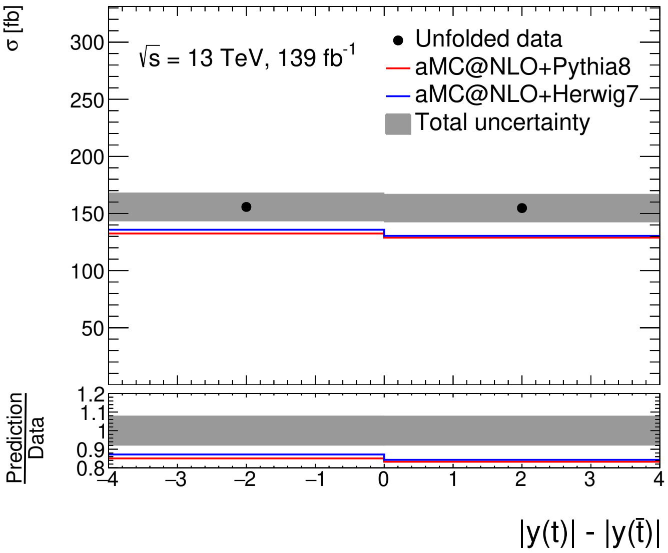
in agreement with the SM prediction at the NLO QCD (from Chapter 4.5): $\mathbf{-0.014 \pm 0.001}$ (scale uncertainty).

Along with the asymmetry A_C, the differential cross-section measurement as a function of $|y(t)| - |y(\bar{t})|$ at the particle level fiducial region is also performed by inferring on μ_+ and μ_-. By using the Equation 11.4, for $n_{b,c}$ in the likelihood definition, the normalised differential cross-section measurement as a function of $|y(t)| - |y(\bar{t})|$ is also performed. They are shown in Figure 11.17 and compared with the MC predictions.

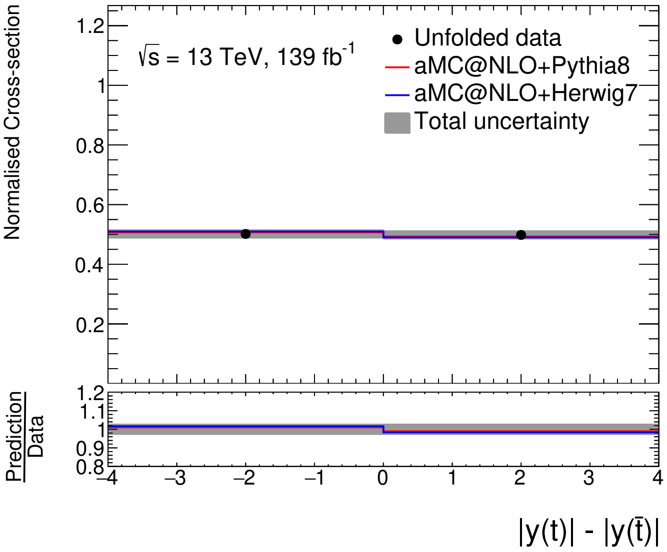
The ranking of the individual NPs in terms of their impact on the final measurement (A_C) is shown in Figure 11.21. As expected, the important uncertainties are very similar to the fit with the Asimov data.

The systematic uncertainties can also be grouped into several categories to better illustrate their impact on the A_C measurement. To assess the impact, the set of NPs corresponding to an individual group is fixed to the fitted values and the fit is repeated. The resulting (quadratic) difference in total uncertainty is taken as the impact. The total uncertainty is not the quadratic sum of the grouped impacts and they are described in Table 11.3.

The measurement is dominated by statistical uncertainty. Also, the MC statistical uncertainties are significant. Among the other uncertainties, the experimental uncertainties in the jets and $E_{\mathrm{T}}^{\mathrm{miss}}$ are found to be important.

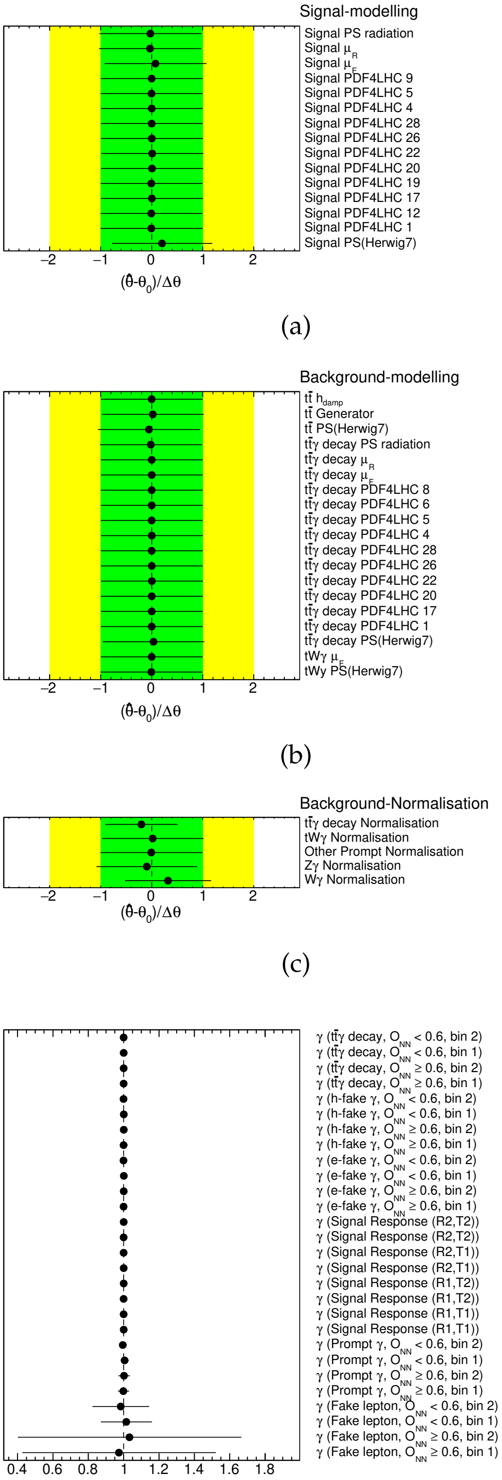


(a)



(b)

Figure 11.17: The unfolded distribution as (a) absolute cross-section and (b) normalised cross-section of $|y(t)| - |y(\bar{t})|$ at the particle level fiducial region, with their uncertainty and comparison to the MC predicted distributions with different parton showers



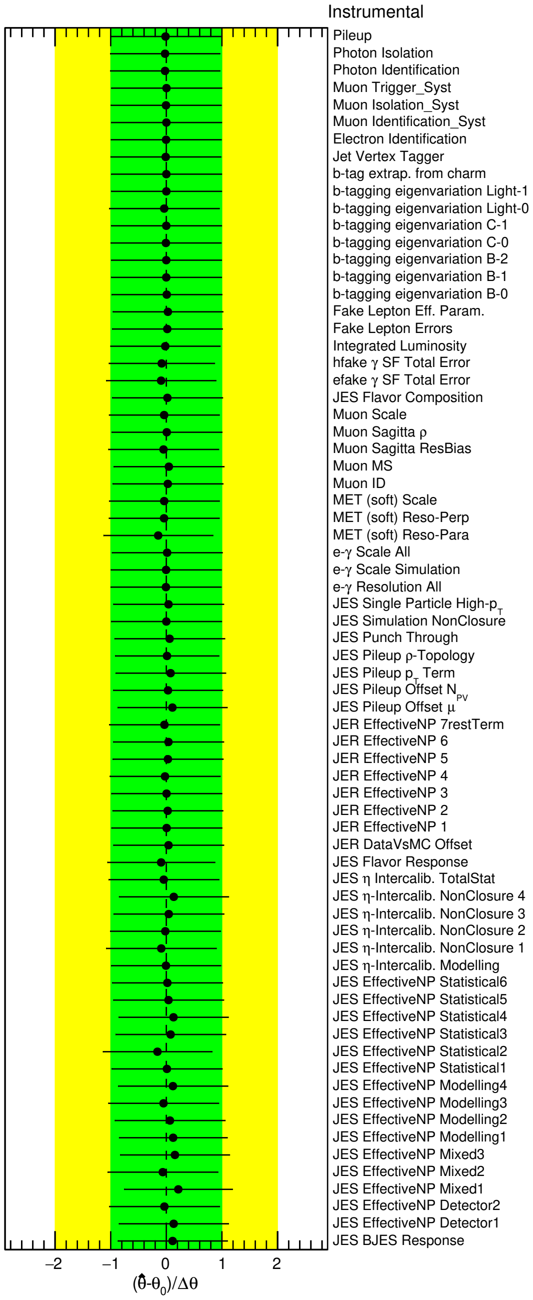
(a)

(b)

(c)

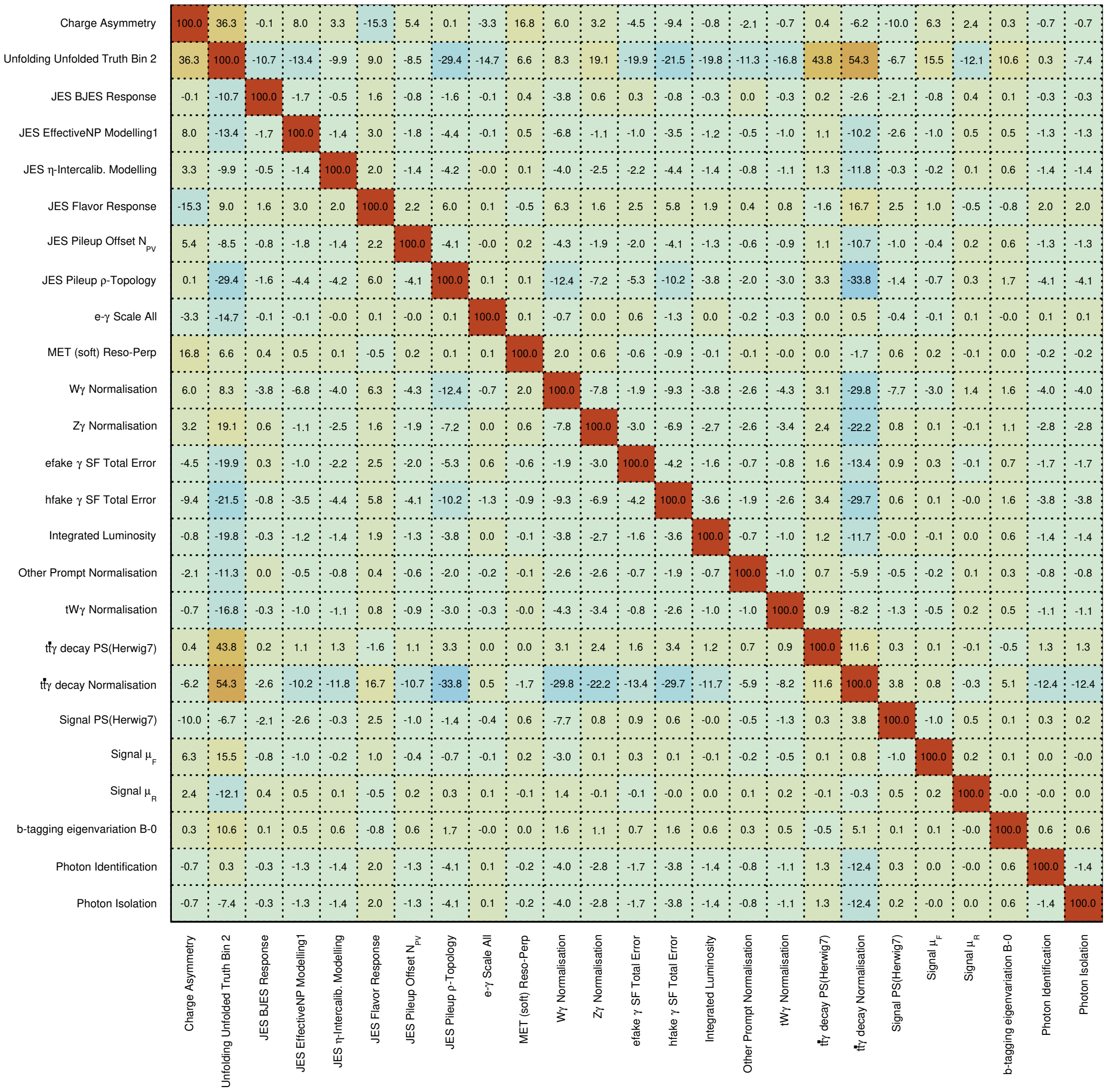
(d)

Figure 11.18: (a,b,c) The fitted NPs ($\hat{\theta}$) (1) with their pulls and constraints with respect to their pre-fit values and (d) Gamma Parameters after the fit.



(a)

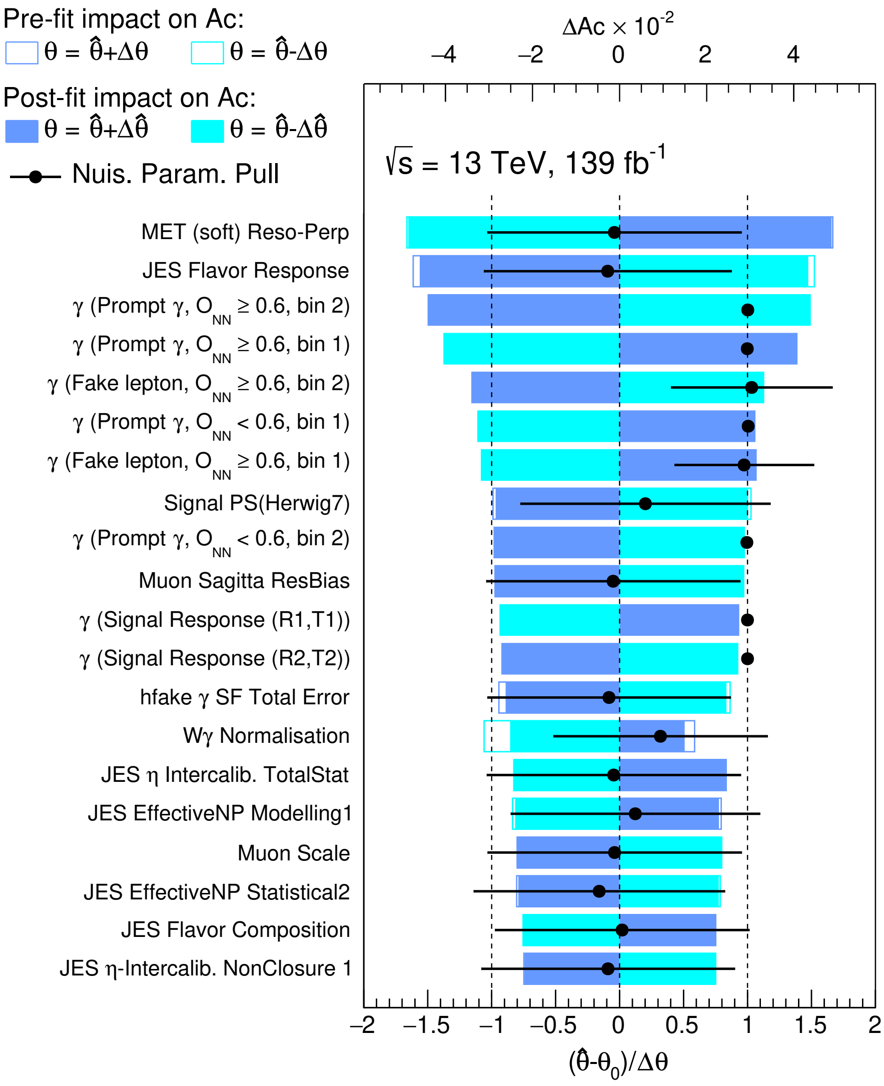
Figure 11.19: The fitted NPs $(\hat{\theta})$ (2) with their pulls and constraints with respect to their pre-fit values after the fit.



	Charge Asymmetry	Unfolding Unfolded Truth Bin 2	JES BJES Response	JES EffectiveNP Modelling1	JES η-Intercalib. Modelling	JES Flavor Response	JES Pileup Offset N_{PV}	JES Pileup ρ-Topology	e-γ Scale All	MET (soft) Reso-Perp	Wγ Normalisation	Zγ Normalisation	efake γ SF Total Error	hfake γ SF Total Error	Integrated Luminosity	Other Prompt Normalisation	tWγ Normalisation	$t\bar{t}\gamma$ decay PS(Herwig7)	$t\bar{t}\gamma$ decay Normalisation	Signal PS(Herwig7)	Signal μ_F	Signal μ_R	b-tagging eigenvariation B-0	Photon Identification	Photon Isolation
Charge Asymmetry	100.0	36.3	-0.1	8.0	3.3	-15.3	5.4	0.1	-3.3	16.8	6.0	3.2	-4.5	-9.4	-0.8	-2.1	-0.7	0.4	-6.2	-10.0	6.3	2.4	0.3	-0.7	-0.7
Unfolding Unfolded Truth Bin 2	36.3	100.0	-10.7	-13.4	-9.9	9.0	-8.5	-29.4	-14.7	6.6	8.3	19.1	-19.9	-21.5	-19.8	-11.3	-16.8	43.8	54.3	-6.7	15.5	-12.1	10.6	0.3	-7.4
JES BJES Response	-0.1	-10.7	100.0	-1.7	-0.5	1.6	-0.8	-1.6	-0.1	0.4	-3.8	0.6	0.3	-0.8	-0.3	0.0	-0.3	0.2	-2.6	-2.1	-0.8	0.4	0.1	-0.3	-0.3
JES EffectiveNP Modelling1	8.0	-13.4	-1.7	100.0	-1.4	3.0	-1.8	-4.4	-0.1	0.5	-6.8	-1.1	-1.0	-3.5	-1.2	-0.5	-1.0	1.1	-10.2	-2.6	-1.0	0.5	0.5	-1.3	-1.3
JES η-Intercalib. Modelling	3.3	-9.9	-0.5	-1.4	100.0	2.0	-1.4	-4.2	-0.0	0.1	-4.0	-2.5	-2.2	-4.4	-1.4	-0.8	-1.1	1.3	-11.8	-0.3	-0.2	0.1	0.6	-1.4	-1.4
JES Flavor Response	-15.3	9.0	1.6	3.0	2.0	100.0	2.2	6.0	0.1	-0.5	6.3	1.6	2.5	5.8	1.9	0.4	0.8	-1.6	16.7	2.5	1.0	-0.5	-0.8	2.0	2.0
JES Pileup Offset N_{PV}	5.4	-8.5	-0.8	-1.8	-1.4	2.2	100.0	-4.1	-0.0	0.2	-4.3	-1.9	-2.0	-4.1	-1.3	-0.6	-0.9	1.1	-10.7	-1.0	-0.4	0.2	0.6	-1.3	-1.3
JES Pileup ρ-Topology	0.1	-29.4	-1.6	-4.4	-4.2	6.0	-4.1	100.0	0.1	0.1	-12.4	-7.2	-5.3	-10.2	-3.8	-2.0	-3.0	3.3	-33.8	-1.4	-0.7	0.3	1.7	-4.1	-4.1
e-γ Scale All	-3.3	-14.7	-0.1	-0.1	-0.0	0.1	-0.0	0.1	100.0	0.1	-0.7	0.0	0.6	-1.3	0.0	-0.2	-0.3	0.0	0.5	-0.4	-0.1	0.1	-0.0	0.1	0.1
MET (soft) Reso-Perp	16.8	6.6	0.4	0.5	0.1	-0.5	0.2	0.1	0.1	100.0	2.0	0.6	-0.6	-0.9	-0.1	-0.1	-0.0	0.0	-1.7	0.6	0.2	-0.1	0.0	-0.2	-0.2
Wγ Normalisation	6.0	8.3	-3.8	-6.8	-4.0	6.3	-4.3	-12.4	-0.7	2.0	100.0	-7.8	-1.9	-9.3	-3.8	-2.6	-4.3	3.1	-29.8	-7.7	-3.0	1.4	1.6	-4.0	-4.0
Zγ Normalisation	3.2	19.1	0.6	-1.1	-2.5	1.6	-1.9	-7.2	0.0	0.6	-7.8	100.0	-3.0	-6.9	-2.7	-2.6	-3.4	2.4	-22.2	0.8	0.1	-0.1	1.1	-2.8	-2.8
efake γ SF Total Error	-4.5	-19.9	0.3	-1.0	-2.2	2.5	-2.0	-5.3	0.6	-0.6	-1.9	-3.0	100.0	-4.2	-1.6	-0.7	-0.8	1.6	-13.4	0.9	0.3	-0.1	0.7	-1.7	-1.7
hfake γ SF Total Error	-9.4	-21.5	-0.8	-3.5	-4.4	5.8	-4.1	-10.2	-1.3	-0.9	-9.3	-6.9	-4.2	100.0	-3.6	-1.9	-2.6	3.4	-29.7	0.6	0.1	-0.0	1.6	-3.8	-3.8
Integrated Luminosity	-0.8	-19.8	-0.3	-1.2	-1.4	1.9	-1.3	-3.8	0.0	-0.1	-3.8	-2.7	-1.6	-3.6	100.0	-0.7	-1.0	1.2	-11.7	-0.0	-0.1	0.0	0.6	-1.4	-1.4
Other Prompt Normalisation	-2.1	-11.3	0.0	-0.5	-0.8	0.4	-0.6	-2.0	-0.2	-0.1	-2.6	-2.6	-0.7	-1.9	-0.7	100.0	-1.0	0.7	-5.9	-0.5	-0.2	0.1	0.3	-0.8	-0.8
tWγ Normalisation	-0.7	-16.8	-0.3	-1.0	-1.1	0.8	-0.9	-3.0	-0.3	-0.0	-4.3	-3.4	-0.8	-2.6	-1.0	-1.0	100.0	0.9	-8.2	-1.3	-0.5	0.2	0.5	-1.1	-1.1
$t\bar{t}\gamma$ decay PS(Herwig7)	0.4	43.8	0.2	1.1	1.3	-1.6	1.1	3.3	0.0	0.0	3.1	2.4	1.6	3.4	1.2	0.7	0.9	100.0	11.6	0.3	0.1	-0.1	-0.5	1.3	1.3
$t\bar{t}\gamma$ decay Normalisation	-6.2	54.3	-2.6	-10.2	-11.8	16.7	-10.7	-33.8	0.5	-1.7	-29.8	-22.2	-13.4	-29.7	-11.7	-5.9	-8.2	11.6	100.0	3.8	0.8	-0.3	5.1	-12.4	-12.4
Signal PS(Herwig7)	-10.0	-6.7	-2.1	-2.6	-0.3	2.5	-1.0	-1.4	-0.4	0.6	-7.7	0.8	0.9	0.6	-0.0	-0.5	-1.3	0.3	3.8	100.0	-1.0	0.5	0.1	0.3	0.2
Signal μ_F	6.3	15.5	-0.8	-1.0	-0.2	1.0	-0.4	-0.7	-0.1	0.2	-3.0	0.1	0.3	0.1	-0.1	-0.2	-0.5	0.1	0.8	-1.0	100.0	0.2	0.1	0.0	-0.0
Signal μ_R	2.4	-12.1	0.4	0.5	0.1	-0.5	0.2	0.3	0.1	-0.1	1.4	-0.1	-0.1	-0.0	0.0	0.1	0.2	-0.1	-0.3	0.5	0.2	100.0	-0.0	-0.0	0.0
b-tagging eigenvariation B-0	0.3	10.6	0.1	0.5	0.6	-0.8	0.6	1.7	-0.0	0.0	1.6	1.1	0.7	1.6	0.6	0.3	0.5	-0.5	5.1	0.1	0.1	-0.0	100.0	0.6	0.6
Photon Identification	-0.7	0.3	-0.3	-1.3	-1.4	2.0	-1.3	-4.1	0.1	-0.2	-4.0	-2.8	-1.7	-3.8	-1.4	-0.8	-1.1	1.3	-12.4	0.3	0.0	-0.0	0.6	100.0	-1.4
Photon Isolation	-0.7	-7.4	-0.3	-1.3	-1.4	2.0	-1.3	-4.1	0.1	-0.2	-4.0	-2.8	-1.7	-3.8	-1.4	-0.8	-1.1	1.3	-12.4	0.2	-0.0	0.0	0.6	-1.4	100.0

(a)

Figure 11.20: The correlations between different NPs and the parameters of interests (POI). The plot shows only those NPs for which the correlation with at least another NP or POI is more than 10%. The POIs A_C and μ_+ are named here as *Charge Asymmetry* and *Unfolding Unfolded Truth Bin 2* respectively.



(a)

Figure 11.21: The ranking of the different NPs according to their impact on the measurement of the charge asymmetry, A_C. The top scale corresponds to the impact of the NPs on the measurement of A_C, whereas the bottom scale corresponds to the fitted NPs with their pulls and constraints same as in the Figures 11.18 and 11.19.

Table 11.3: Summary of the impact of the systematic uncertainties on A_C grouped into different categories. The quoted uncertainties are obtained by repeating the fit with a group of nuisance parameters fixed to their fitted values and subtracting in quadrature the resulting total uncertainty from the uncertainty of the complete fit. However, the total uncertainty is not the quadratic sum of the grouped impacts, as this approach neglects the correlation among the different groups. The category *Other experimental* includes uncertainties associated with leptons, pile-up and luminosity

	Grouped Impact
Total uncertainty	0.029
Statistical uncertainty	0.024
MC statistical uncertainties	
Background processes	0.008
Signal	0.004
Modelling uncertainties	
Signal modelling	0.003
Background modelling	0.002
All prompt γ bkg. normalisation	0.002
Experimental uncertainties	
Jet	0.009
$E_{\mathrm{T}}^{\mathrm{miss}}$	0.005
Fake lepton background	0.005
Fake photon background	0.003
Photon	0.001
b-tagging	0.001
Other experimental	0.004

CHAPTER 12

Conclusion and outlook

This is the first measurement of the top quark pair charge asymmetry in $t\bar{t}\gamma$ production in agreement with the SM expectation [3].

A set of novel and crucial methods are considered for the analysis. For the first time, the reconstructed top quark kinematic observables are studied for the $t\bar{t}\gamma$ production process. A folding approach in the likelihood fit enables direct inference on the unfolded distribution and hence reduces the uncertainty compared to unfolding an already fitted distribution. This is also the first $t\bar{t}\gamma$ production related measurement where the radiative top decays are not considered as signal to enhance the sensitivity of the underlying physics. Using a neural network discriminator a good separation is achieved between the $t\bar{t}\gamma$ production and the $t\bar{t}\gamma$ decay background.

The measurement is expected to be sensitive to potential BSM physics as the process is more exposed to the quark-initiated production than in $t\bar{t}$ production. This is still to be tested either via the relevant operators in the Effective Field Theory approach or by constraining the parameters of axigluon models.

With the tiny cross-section of the $t\bar{t}\gamma$ production process compared to the $t\bar{t}$ production and the limited ATLAS Run 2 dataset, it was expected that the measurement would be limited by the statistics. The precision of this measurement is therefore expected to improve significantly with the HL-LHC data to be collected in the future with 20 times more integrated luminosity, reducing the statistical uncertainty. Theoretical predictions with higher-order QCD and EW corrections would make interesting comparisons with such future measurements. Given this large uncertainty on the measurement, the hypothesis of zero asymmetry in $t\bar{t}\gamma$ production cannot be excluded. However, since the nullification of oppositely signed multiple asymmetric effects are present in this process, the measured asymmetry being zero would still be physics relevant.

As an alternative measurement, can be performed especially with a much larger dataset, a different observable of the same kind, the leptonic charge asymmetry ($A_C^{\ell\ell}$) can be defined on the charged leptons from the top quark pair decays in the dilepton channels:

$$A_C^{\ell\ell} = \frac{N(\Delta|\eta| > 0) - N(\Delta|\eta| < 0)}{N(\Delta|\eta| > 0) + N(\Delta|\eta| < 0)} \tag{12.1}$$

where $\Delta|\eta| = |\eta_{\ell+}| - |\eta_{\ell-}|$. The charge asymmetry between the top quarks would be propagated to the corresponding charged leptons. This approach could avoid the kinematic reconstruction of top quarks and one would only need the reconstructed leptons. This could potentially reduce the systematic uncertainties that originate from the jets.

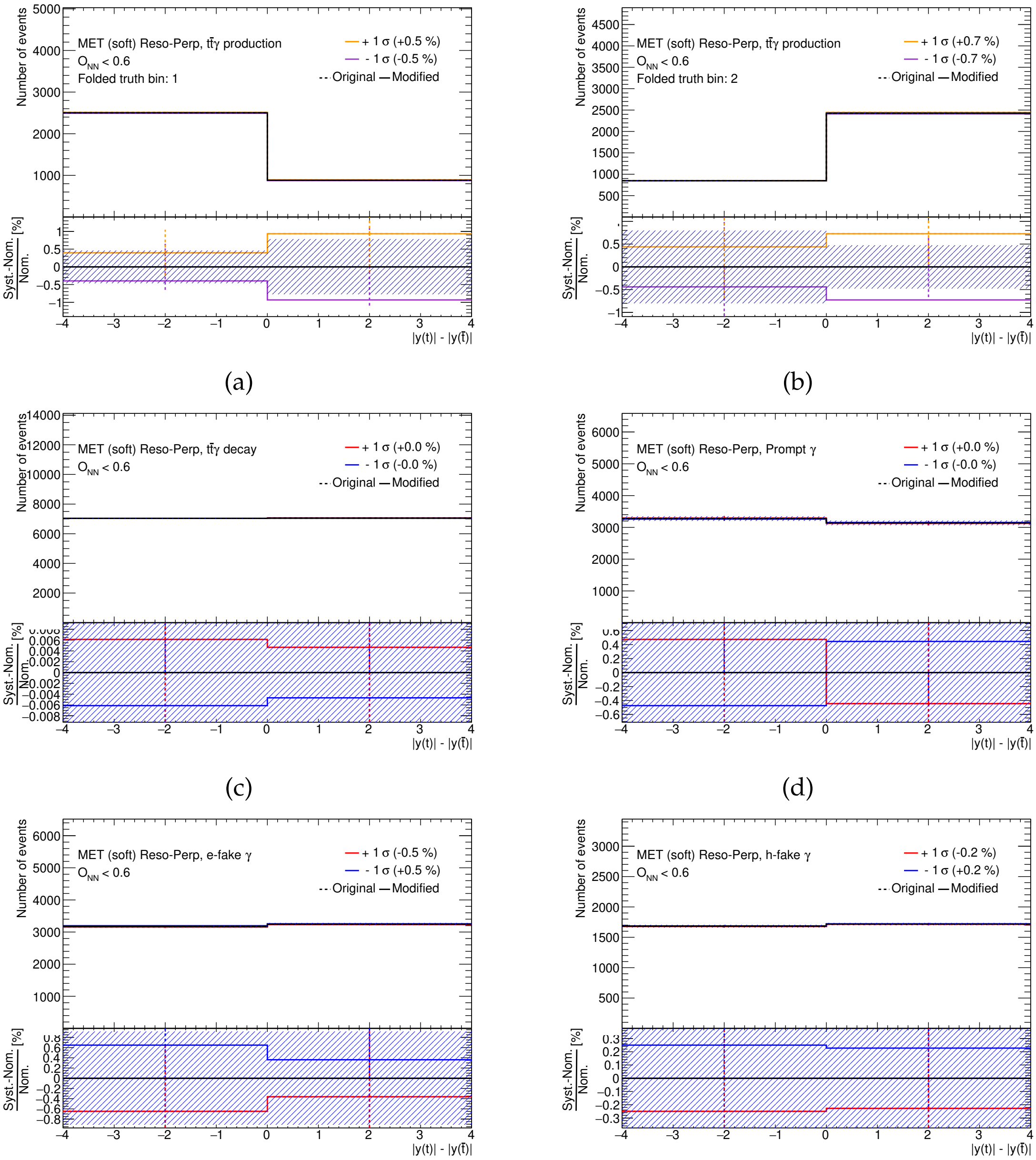
Other approaches for improving the experimental systematic uncertainties could be done mainly by improving the method of reconstruction of the top quarks. Advanced machine learning based methods like SPA-NET [105] and a better b-tagging algorithm (already a reality in Run 3) could improve the top quark reconstruction efficiency and hence the related uncertainties.

It is also vital to look for a more asymmetric phase space of the $\mathrm{t\bar{t}}\gamma$ production process. From theoretical studies [23], it is known that the asymmetry will get enhanced with photon radiation towards the forward region (higher $|\eta(\gamma)|$) in the $\mathrm{t\bar{t}}\gamma$ production, similar to the boosted $\mathrm{t\bar{t}}$ system in $\mathrm{t\bar{t}}$ production. A differential charge asymmetry measurement as a function of the $|\eta(\gamma)|$ would give a more in-depth knowledge of the underlying physics. This could also be explored in the future with more data.

This new measurement discussed in this thesis has its limitations at the current stage. However, with its novel approaches, it has the potential to serve as a stepping stone for related theoretical and experimental physics research in the future.

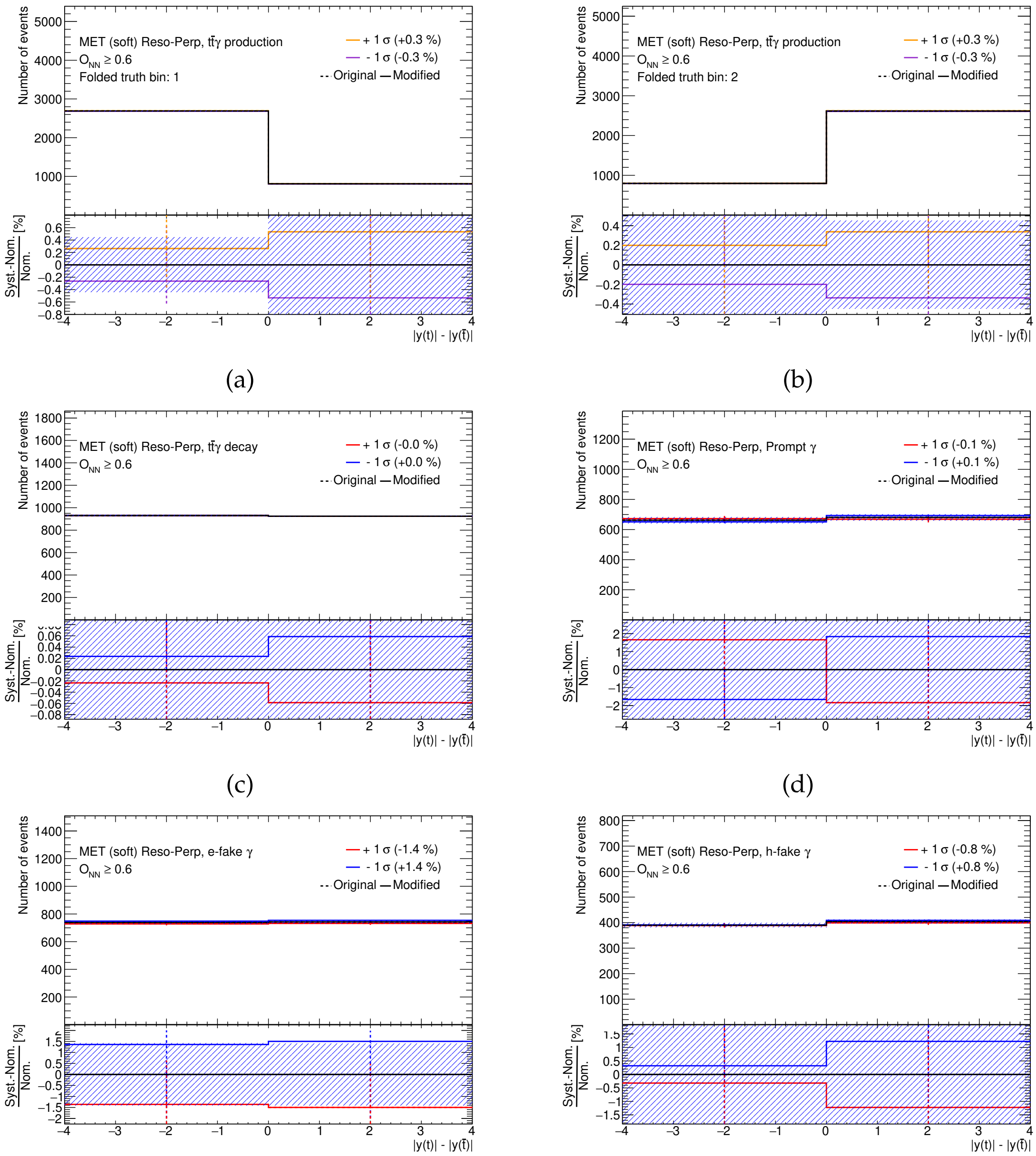
APPENDIX A

Comparisons of distributions for important uncertainties



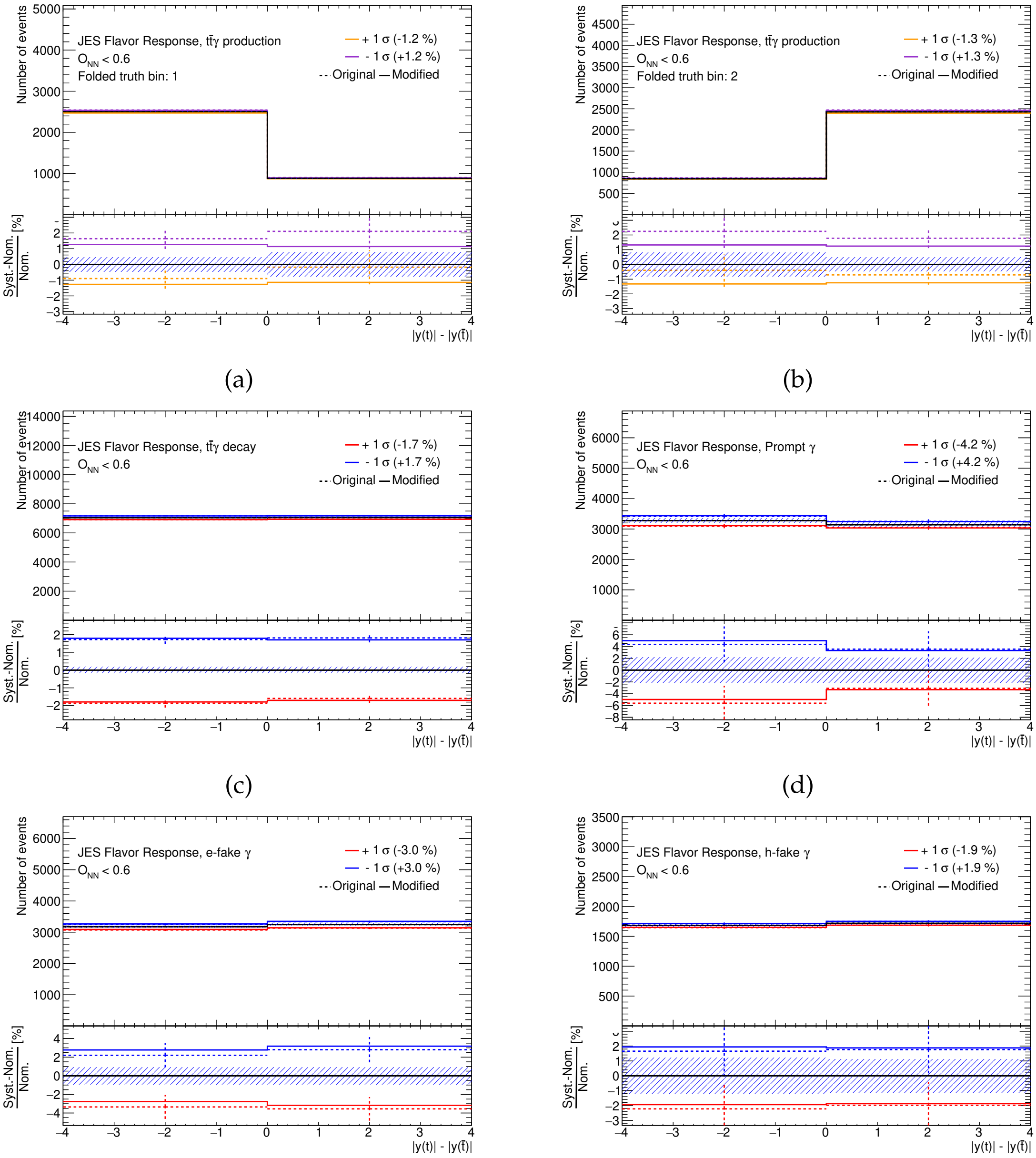
(a) (b) (c) (d) (e) (f)

Figure A.1: The variations (up and down) on $|y(t)| - |y(\bar{t})|$ for systematic uncertainty of MET (soft) resolution component are shown and compared to the nominal distribution for all the processes in the signal region with $O_{NN} < 0.6$. The original variations are shown with dotted lines and the symmetrised variations are shown with solid lines. The plot corresponds to the processes (a) and (b) signal, (c) $t\bar{t}\gamma$ decay, (d) prompt γ, (e) e-fake γ and (f) h-fake γ.



(a) (b) (c) (d) (e) (f)

Figure A.2: The variations (up and down) on $|y(t)| - |y(\bar{t})|$ for systematic uncertainty of MET (soft) resolution component are shown and compared to the nominal distribution for all the processes in the signal region with $O_{NN} \geq 0.6$. The original variations are shown with dotted lines and the symmetrised variations are shown with solid lines. The plot corresponds to the processes (a) and (b) signal, (c) tt̄γ decay, (d) prompt γ, (e) e-fake γ and (f) h-fake γ.



(a)

(b)

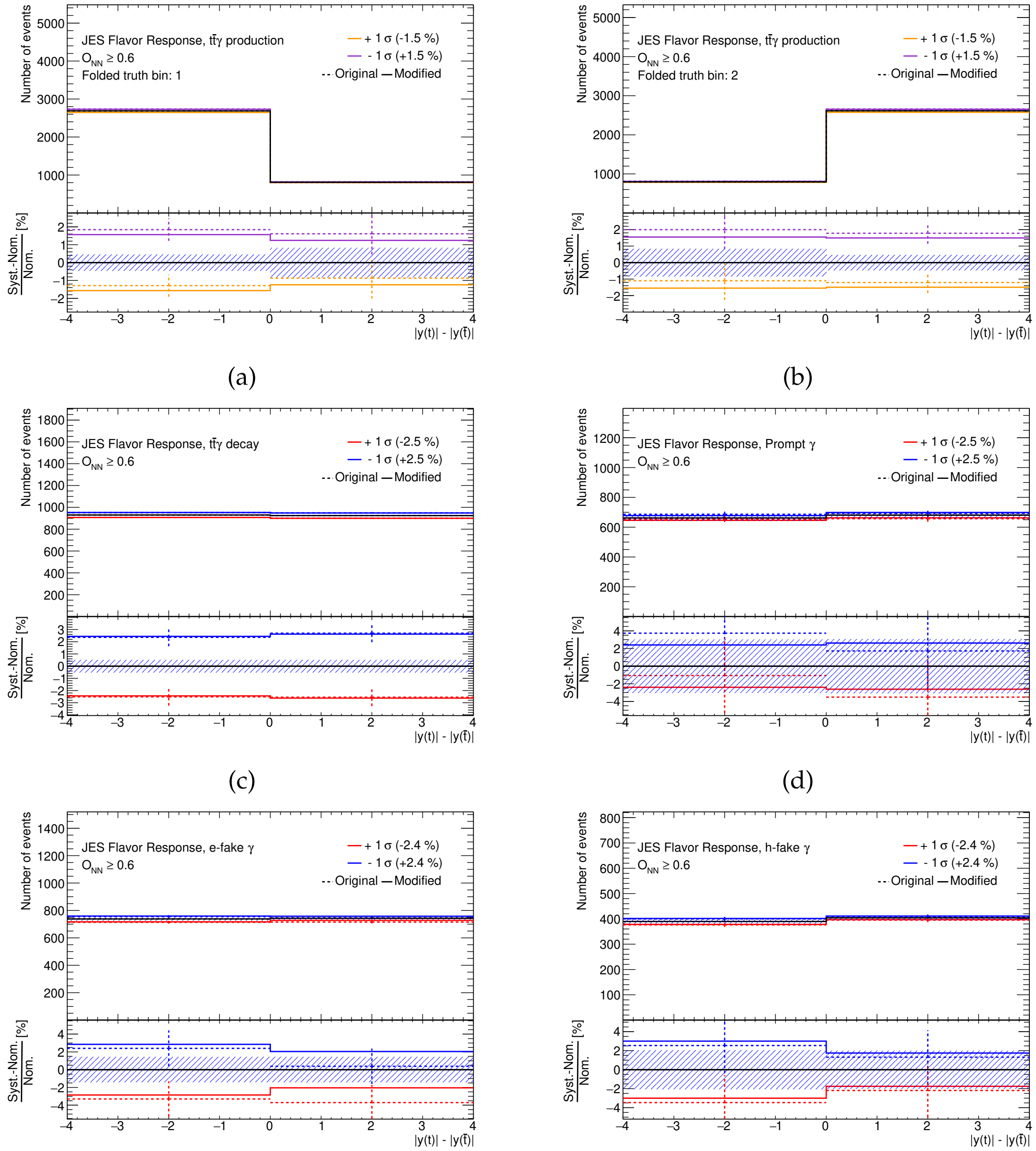
(c)

(d)

(e)

(f)

Figure A.3: The variations (up and down) on $|y(t)| - |y(\bar{t})|$ for systematic uncertainty of flavour response for jet energy scale are shown and compared to the nominal distribution for all the processes in the signal region with $O_{NN} < 0.6$. The original variations are shown with dotted lines and the symmetrised variations are shown with solid lines. The plot corresponds to the processes (a) and (b) signal, (c) $t\bar{t}\gamma$ decay, (d) prompt γ, (e) e-fake γ and (f) h-fake γ.



(a) (b)

(c) (d)

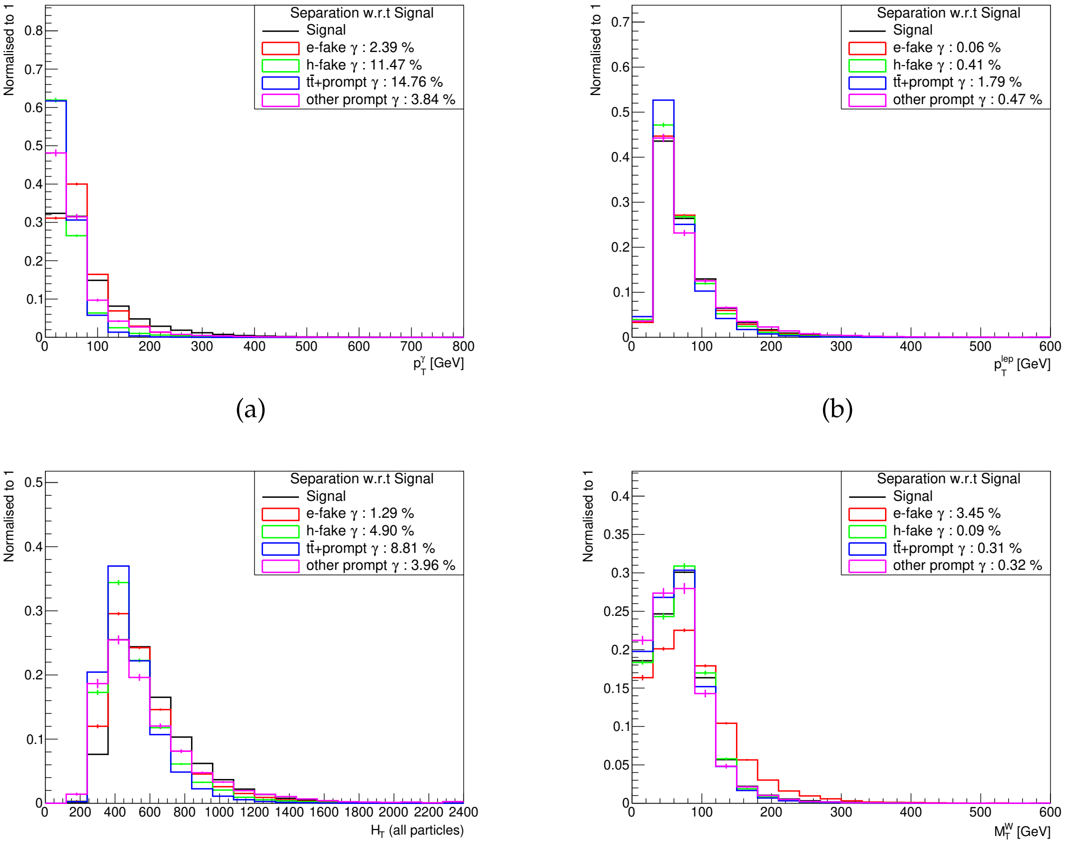
(e) (f)

Figure A.4: The variations (up and down) on $|y(t)| - |y(\bar{t})|$ for systematic uncertainty of flavour response for jet energy scale are shown and compared to the nominal distribution for all the processes in the signal region with $O_{NN} \geq 0.6$. The original variations are shown with dotted lines and the symmetrised variations are shown with solid lines. The plot corresponds to the processes (a) and (b) signal, (c) $t\bar{t}\gamma$ decay, (d) prompt γ, (e) e-fake γ and (f) h-fake γ.

APPENDIX B

Supporting materials: Neural Network

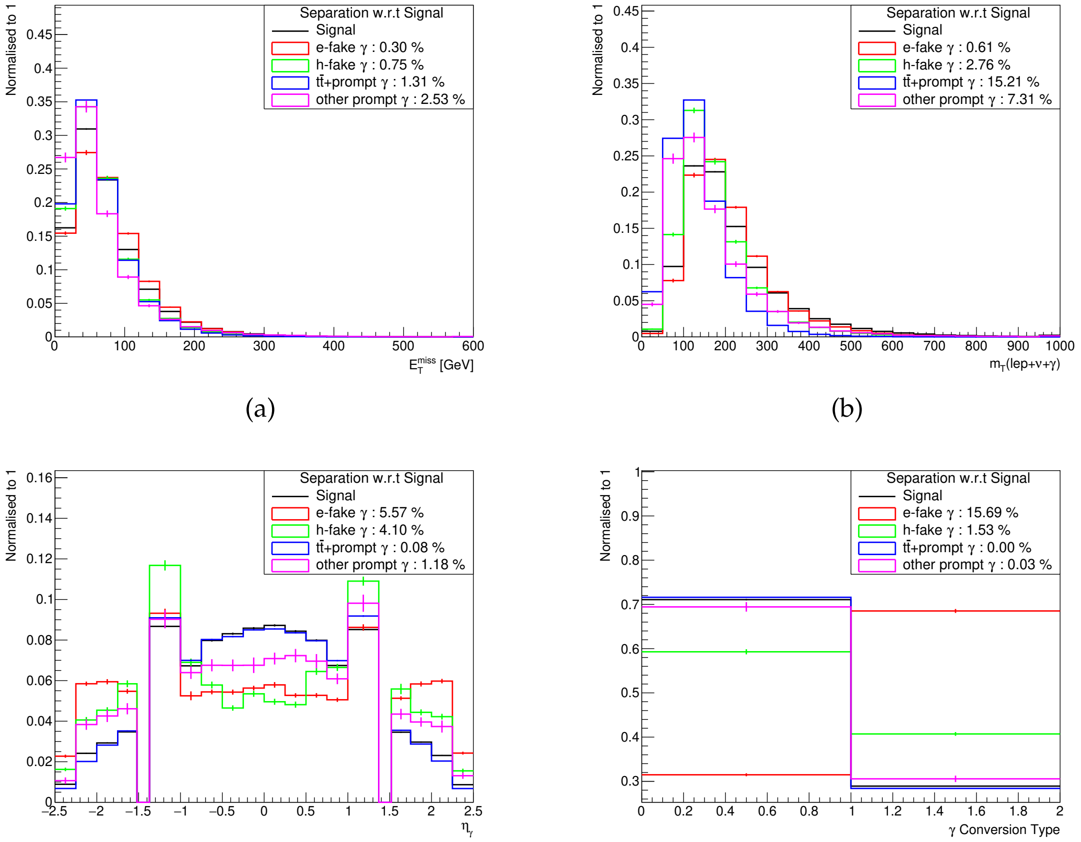
B.1 Input variables and Separation calculation



(a) (b)

(c) (d)

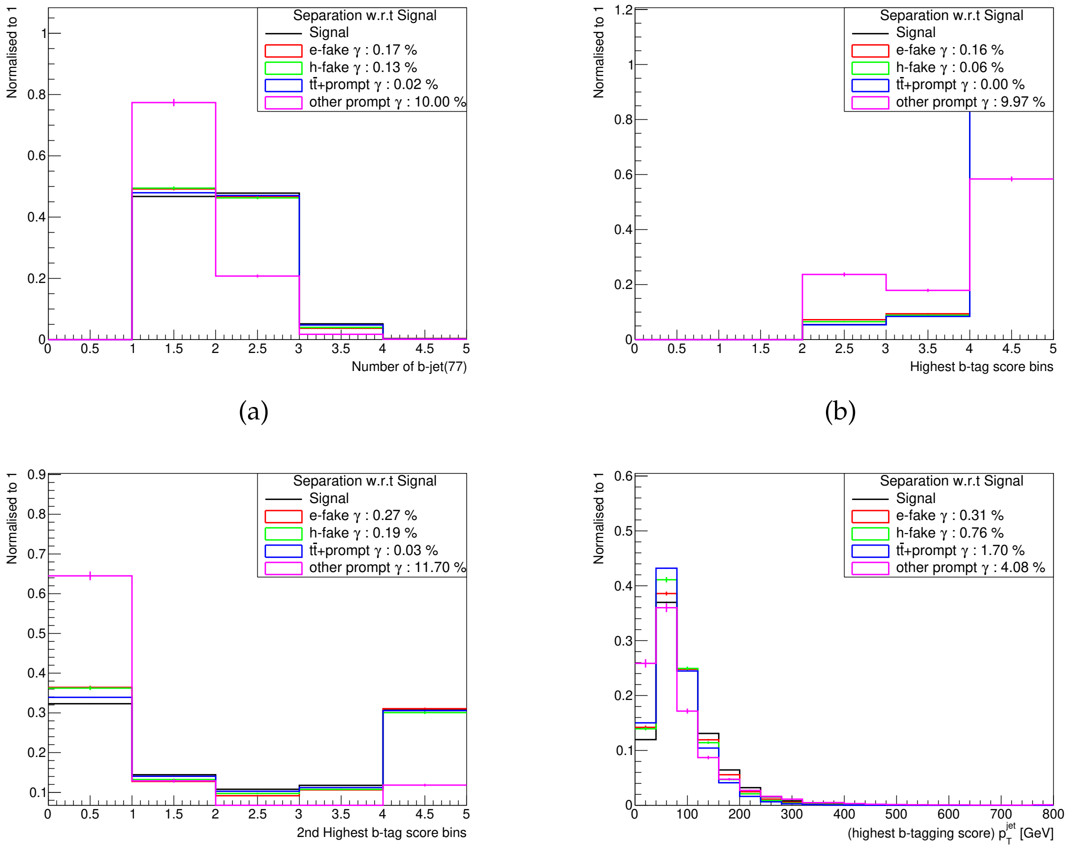
Figure B.1: The normalised comparison of the different background categories with respect to signal to calculate the separation power for the input variables



(a) (b)

(c) (d)

Figure B.2: The normalised comparison of the different background categories with respect to signal to calculate the separation power for the input variables



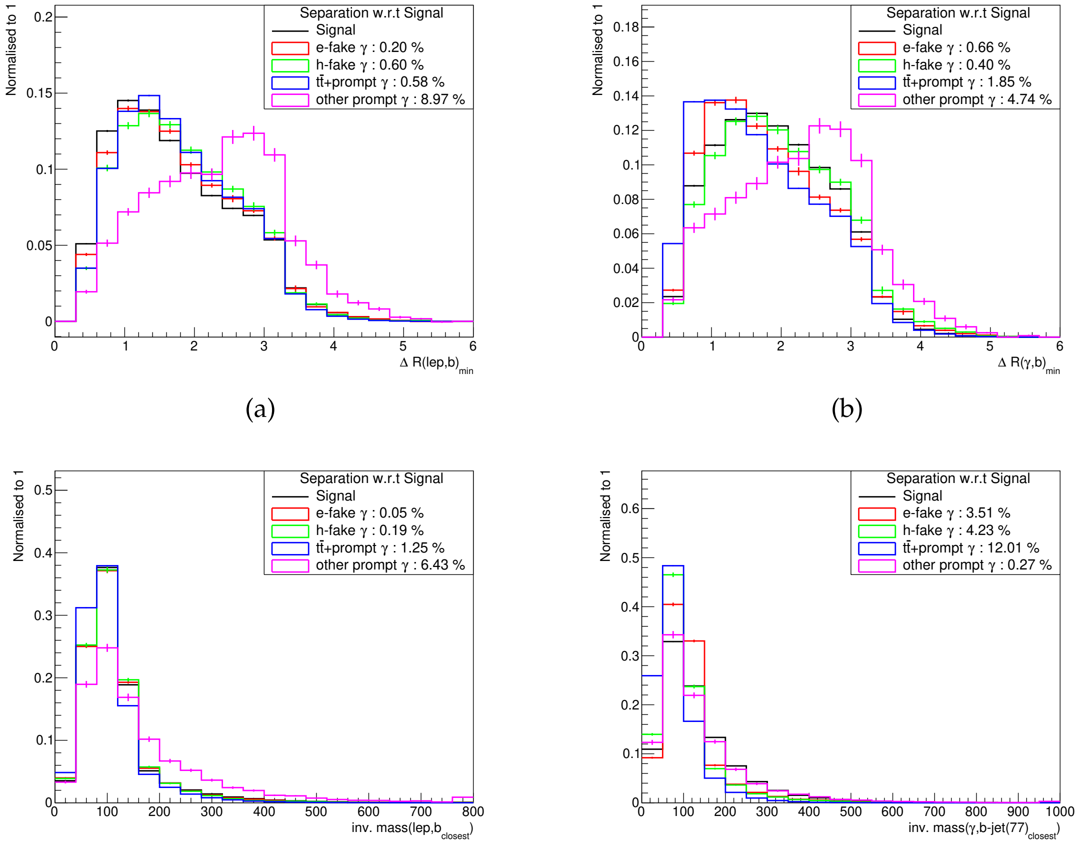
(a)

(b)

(c)

(d)

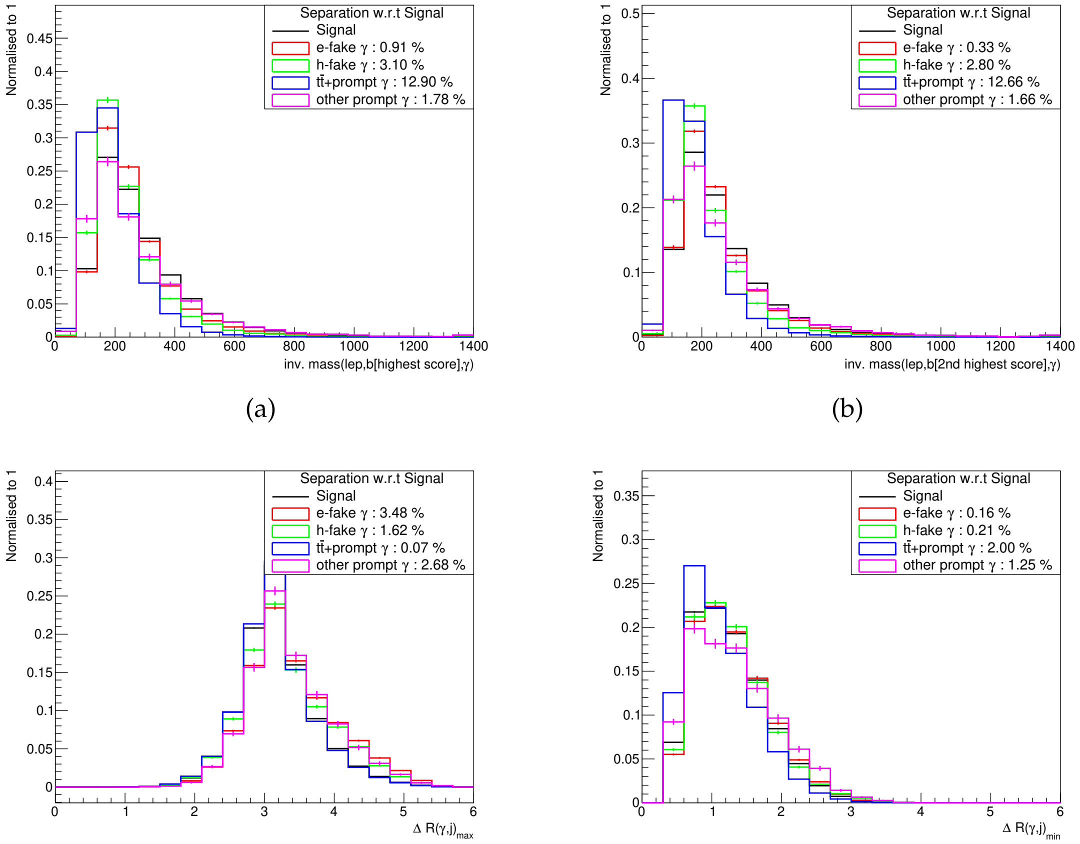
Figure B.3: The normalised comparison of the different background categories with respect to signal to calculate the separation power for the input variables



(a) (b)

(c) (d)

Figure B.4: The normalised comparison of the different background categories with respect to signal to calculate the separation power for the input variables



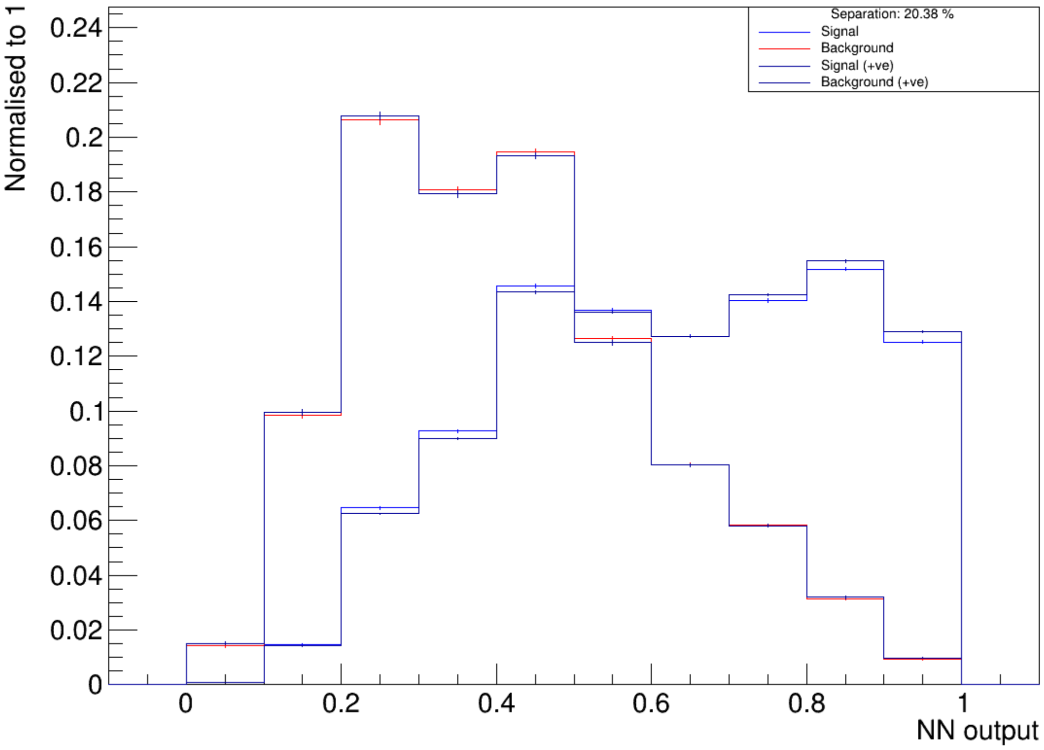
(a) (b)

(c) (d)

Figure B.5: The normalised comparison of the different background categories with respect to signal to calculate the separation power for the input variables

B.2 Check of negative weights for NN training

To check if the negative weighted events of signal and background are significantly different from the positive weighted events, a comparison is done with all events and positive weighted events for the neural network output variable for both signal and backgrounds with their normalised distributions. The result shows that there is no significant difference. Hence, it is safe to use the positive weighted events only for the training.



(a)

Figure B.6: The comparison of signal and background normalised distribution of NN output with all events and only positive weighted events

Bibliography

[1] R. Lysák, *Charge Asymmetry in Top Quark Pair Production*, Symmetry **12** (2020), ISSN: 2073-8994,
URL: `https://www.mdpi.com/2073-8994/12/8/1278`
(cit. on pp. 2, 10).

[2] ATLAS Collaboration, *Evidence for the charge asymmetry in $pp \to t\bar{t}$ production at $\sqrt{s} = 13\,TeV$ with the ATLAS detector*, (2022),
arXiv: `2208.12095 [hep-ph]` (cit. on pp. 2, 12, 78).

[3] ATLAS Collaboration, *Measurement of the charge asymmetry in top-quark pair production in association with a photon with the ATLAS experiment*,
Physics Letters B **843** (2023) 137848, ISSN: 0370-2693,
arXiv: `2212.10552 [hep-ex]`, URL: `https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PAPERS/TOPQ-2021-29/` (cit. on pp. 2, 14, 96, 126, 133).

[4] *Standard Model diagram, taken from this source*, URL: `https://commons.wikimedia.org/w/index.php?curid=59970304#/media/File:Standard_Model_of_Elementary_Particles.svg` (cit. on p. 6).

[5] R. L. Workman et al., *Review of Particle Physics*, PTEP **2022** (2022) 083C01
(cit. on pp. 6, 7, 13).

[6] *Standard Model Wikipedia*,
URL: `https://en.wikipedia.org/wiki/Standard_Model`
(cit. on p. 5).

[7] P. W. Anderson, *Plasmons, Gauge Invariance, and Mass*,
Phys. Rev. **130** (1 1963) 439,
URL: `https://link.aps.org/doi/10.1103/PhysRev.130.439`
(cit. on p. 6).

[8] P. Higgs, *Broken symmetries, massless particles and gauge fields*,
Physics Letters **12** (1964) 132, ISSN: 0031-9163, URL: `https://www.sciencedirect.com/science/article/pii/0031916364911369`
(cit. on p. 6).

[9] F. Englert and R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons*, Phys. Rev. Lett. **13** (9 1964) 321, URL: `https://link.aps.org/doi/10.1103/PhysRevLett.13.321` (cit. on p. 6).

[10] G. S. Guralnik, C. R. Hagen and T. W. B. Kibble, *Global Conservation Laws and Massless Particles*, Phys. Rev. Lett. **13** (20 1964) 585, URL: `https://link.aps.org/doi/10.1103/PhysRevLett.13.585` (cit. on p. 6).

[11] *Tevatron - Wikipedia*, URL: `https://en.wikipedia.org/wiki/Tevatron` (cit. on p. 7).

[12] *Large Hadron Collider*, URL: `https://home.cern/science/accelerators/large-hadron-collider` (cit. on p. 7).

[13] *Branching fractions of $t\bar{t}$ production*, URL: `https://www-d0.fnal.gov/Run2Physics/top/top_public_web_pages/top_feynman_diagrams.html` (cit. on p. 8).

[14] M. Thomson, *Modern Particle Physics*, Cambridge, United Kingdom: Cambridge University Press, 2013 (cit. on pp. 8, 9, 52).

[15] *Handedness*, URL: `https://atlas.cern/glossary/handedness` (cit. on p. 9).

[16] *First Measurement of the Charge Asymmetry in Beauty-Quark Pair Production*, Phys. Rev. Lett. **113** (8 2014) 082003, URL: `https://link.aps.org/doi/10.1103/PhysRevLett.113.082003` (cit. on p. 10).

[17] *First measurement of the forward-backward asymmetry in bottom-quark pair production at high mass*, Phys. Rev. D **92** (3 2015) 032006, URL: `https://link.aps.org/doi/10.1103/PhysRevD.92.032006` (cit. on p. 10).

[18] H. Sagawa et al., *Measurement of $e^+e^- \to bb^-$ forward-backward charge asymmetry between $\sqrt{s}$ =52 and 57 GeV*, Phys. Rev. Lett. **63** (21 1989) 2341, URL: `https://link.aps.org/doi/10.1103/PhysRevLett.63.2341` (cit. on p. 10).

[19] *Axigluon Wikipedia*, URL: `https://en.wikipedia.org/wiki/Chiral_color` (cit. on p. 10).

[20] M. Czakon et al., *Top-quark charge asymmetry at the LHC and Tevatron through NNLO QCD and NLO EW*, Phys. Rev. D **98** (1 2018) 014003, URL: `https://link.aps.org/doi/10.1103/PhysRevD.98.014003` (cit. on pp. 12, 26, 78).

[21] *Combined Forward-Backward Asymmetry Measurements in Top-Antitop Quark Production at the Tevatron*, Phys. Rev. Lett. **120** (4 2018) 042001, URL: `https://link.aps.org/doi/10.1103/PhysRevLett.120.042001` (cit. on p. 12).

[22] J. Aguilar-Saavedra, E. Álvarez and A. Juste, *Shedding light on the $t\bar{t}$ charge asymmetry: the photon handle*, JHEP **04** (2014) 188, arXiv: `1402.3598 [hep-ph]` (cit. on p. 12).

[23] J. Bergner and M. Schulze, *The top quark charge asymmetry in $t\bar{t}\gamma$ production at the LHC*, Eur. Phys. J. C **79** (2019) 189, arXiv: `1812.10535 [hep-ph]` (cit. on pp. 12, 26, 29, 134).

[24] A. Collaboration, *Search for leptonic charge asymmetry in $t\bar{t}W$ production in final states with three leptons at $\sqrt{s} = 13$ TeV*, 2023, arXiv: `2301.04245 [hep-ex]` (cit. on p. 12).

[25] *Evidence for $t\bar{t}\gamma$ production and measurement of $\sigma_{t\bar{t}\gamma}/\sigma_{t\bar{t}}$*, Phys. Rev. D **84** (3 2011) 031104, URL: `https://link.aps.org/doi/10.1103/PhysRevD.84.031104` (cit. on p. 13).

[26] *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** (7 2015) 072007, URL: `https://link.aps.org/doi/10.1103/PhysRevD.91.072007` (cit. on p. 13).

[27] ATLAS Collaboration, *Measurements of inclusive and differential fiducial cross-sections of $t\bar{t}\gamma$ production in leptonic final states at $\sqrt{s} = 13$ TeV in ATLAS*, Eur. Phys. J. C **79** (2019) 382, arXiv: `1812.01697 [hep-ex]` (cit. on p. 13).

[28] ATLAS Collaboration, *Measurements of inclusive and differential cross-sections of combined $t\bar{t}\gamma$ and $tW\gamma$ production in the $e\mu$ channel at 13 TeV with the ATLAS detector*, JHEP **09** (2020) 049, arXiv: `2007.06946 [hep-ex]` (cit. on p. 13).

[29] The CMS Collaboration, *Measurement of the inclusive and differential $t\bar{t}\gamma$ cross sections in the single-lepton channel and EFT interpretation at $\sqrt{s} = 13$ TeV*, JHEP **12** (2021) 180, arXiv: `2107.01508 [hep-ex]` (cit. on p. 13).

[30] The 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 **05** (2022) 091, arXiv: `2201.07301 [hep-ex]` (cit. on p. 13).

[31] D. Pagani, H.-S. Shao, I. Tsinikos and M. Zaro, *Automated EW corrections with isolated photons: $t\bar{t}\gamma$, $t\bar{t}\gamma\gamma$ and $t\gamma j$ as case studies*, (), arXiv: `2106.02059 [hep-ph]` (cit. on p. 13).

[32] *The ATLAS Experiment at the CERN Large Hadron Collider*, Journal of Instrumentation **3** (2008) S08003, URL: `https://dx.doi.org/10.1088/1748-0221/3/08/S08003` (cit. on p. 15).

[33] *The Large Hadron Collider*, URL: `https://home.cern/science/accelerators/large-hadron-collider` (cit. on p. 15).

[34] *Philippe Mouche. Overall view of the LHC. General Photo. June 2014*, URL: `http://cds.cern.ch/record/1708847` (cit. on pp. 15, 18).

[35] ATLAS Collaboration, *Luminosity determination in pp collisions at $\sqrt{s} = 13\,TeV$ using the ATLAS detector at the LHC*, ATLAS-CONF-2019-021, 2019, URL: `https://cds.cern.ch/record/2677054` (cit. on pp. 16, 56).

[36] *Public ATLAS Luminosity Results for Run-2 of the LHC*, URL: `https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2` (cit. on p. 17).

[37] *Pseudorapidity Wikipedia*, URL: `https://en.wikipedia.org/wiki/Pseudorapidity#/media/File:Pseudorapidity2.png` (cit. on p. 18).

[38] G. Aad et al., *Electron and photon performance measurements with the ATLAS detector using the 2015–2017 LHC proton-proton collision data*, JINST **14** (2019) P12006, arXiv: `1908.00005 [hep-ex]` (cit. on pp. 19, 34, 35, 38–40, 42).

[39] *Measurement of the Inelastic Proton-Proton Cross Section at $\sqrt{s} = 13$ TeV with the ATLAS Detector at the LHC*, Phys. Rev. Lett. **117** (18 2016) 182002, URL: `https://link.aps.org/doi/10.1103/PhysRevLett.117.182002` (cit. on p. 20).

[40] GEANT4 Collaboration, S. Agostinelli et al., GEANT4 – *a simulation toolkit*, Nucl. Instrum. Meth. A **506** (2003) 250 (cit. on p. 24).

[41] W. Lukas, *Fast Simulation for ATLAS: Atlfast-II and ISF*, Journal of Physics: Conference Series **396** (2012) 022031, URL: `https://dx.doi.org/10.1088/1742-6596/396/2/022031` (cit. on p. 24).

[42] ATLAS Collaboration, *The ATLAS Collaboration Software and Firmware*, ATL-SOFT-PUB-2021-001, 2021, URL: `https://cds.cern.ch/record/2767187` (cit. on p. 25).

[43] 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 **07** (2014) 079, arXiv: `1405.0301 [hep-ph]` (cit. on p. 25).

[44] R. D. Ball et al., *Parton distributions for the LHC run II*, JHEP **04** (2015) 040, arXiv: `1410.8849 [hep-ph]` (cit. on p. 25).

[45] S. Frixione, E. Laenen, P. Motylinski and B. R. Webber, *Angular correlations of lepton pairs from vector boson and top quark decays in Monte Carlo simulations*, JHEP **04** (2007) 081, arXiv: `hep-ph/0702198` (cit. on p. 25).

[46] P. Artoisenet, R. Frederix, O. Mattelaer and R. Rietkerk, *Automatic spin-entangled decays of heavy resonances in Monte Carlo simulations*, JHEP **03** (2013) 015, arXiv: `1212.3460 [hep-ph]` (cit. on p. 25).

[47] S. Frixione, *Isolated photons in perturbative QCD*, Phys. Lett. B **429** (1998) 369, arXiv: `hep-ph/9801442` (cit. on p. 25).

[48] T. Sjöstrand et al., *An introduction to PYTHIA 8.2*, Comput. Phys. Commun. **191** (2015) 159, arXiv: `1410.3012 [hep-ph]` (cit. on pp. 25, 52).

[49] D. J. Lange, *The EvtGen particle decay simulation package*, Nucl. Instrum. Meth. A **462** (2001) 152 (cit. on p. 25).

[50] P. Nason, *A new method for combining NLO QCD with shower Monte Carlo algorithms*, JHEP **11** (2004) 040, arXiv: `hep-ph/0409146` (cit. on p. 26).

[51] S. Frixione, P. Nason and C. Oleari, *Matching NLO QCD computations with parton shower simulations: the POWHEG method*, JHEP **11** (2007) 070, arXiv: `0709.2092 [hep-ph]` (cit. on p. 26).

[52] S. Alioli, P. Nason, C. Oleari and E. Re, *A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX*, JHEP **06** (2010) 043, arXiv: `1002.2581 [hep-ph]` (cit. on p. 26).

[53] M. Czakon and A. Mitov, *Top++: A program for the calculation of the top-pair cross-section at hadron colliders*, Comput. Phys. Commun. **185** (2014) 2930, arXiv: `1112.5675 [hep-ph]` (cit. on p. 26).

[54] N. Kidonakis, *Next-to-next-to-leading logarithm resummation for s-channel single top quark production*, Phys. Rev. **D81** (2010) 054028, arXiv: `1001.5034 [hep-ph]` (cit. on p. 27).

[55] N. Kidonakis, *Two-loop soft anomalous dimensions for single top quark associated production with a* W^- *or* H^-, Phys. Rev. D **82** (2010) 054018, arXiv: `1005.4451 [hep-ph]` (cit. on p. 27).

[56] N. Kidonakis, *Next-to-next-to-leading-order collinear and soft gluon corrections for t-channel single top quark production*, Phys. Rev. **D83** (2011) 091503, arXiv: `1103.2792 [hep-ph]` (cit. on p. 27).

[57] S. Frixione, E. Laenen, P. Motylinski, C. White and B. R. Webber, *Single-top hadroproduction in association with a W boson*, JHEP **07** (2008) 029 (cit. on p. 27).

[58] T. Gleisberg et al., *Event generation with SHERPA 1.1*, JHEP **02** (2009) 007, arXiv: `0811.4622 [hep-ph]` (cit. on p. 27).

[59] S. Höche, F. Krauss, S. Schumann and F. Siegert, *QCD matrix elements and truncated showers*, JHEP **05** (2009) 053, arXiv: `0903.1219 [hep-ph]` (cit. on p. 27).

[60] T. Gleisberg and S. Höche, *Comix, a new matrix element generator*, JHEP **12** (2008) 039, arXiv: `0808.3674 [hep-ph]` (cit. on p. 27).

[61] S. Schumann and F. Krauss, *A parton shower algorithm based on Catani–Seymour dipole factorisation*, JHEP **03** (2008) 038, arXiv: `0709.1027 [hep-ph]` (cit. on p. 27).

[62] S. Höche, F. Krauss, M. Schönherr and F. Siegert, *A critical appraisal of NLO+PS matching methods*, JHEP **09** (2012) 049, arXiv: `1111.1220 [hep-ph]` (cit. on p. 27).

[63] S. Catani, F. Krauss, R. Kuhn and B. R. Webber, *QCD Matrix Elements + Parton Showers*, JHEP **11** (2001) 063, arXiv: `hep-ph/0109231` (cit. on p. 27).

[64] S. Höche, F. Krauss, M. Schönherr and F. Siegert, *QCD matrix elements + parton showers. The NLO case*, JHEP **04** (2013) 027, arXiv: `1207.5030 [hep-ph]` (cit. on p. 27).

[65] F. Cascioli, P. Maierhöfer and S. Pozzorini,
Scattering Amplitudes with Open Loops, Phys. Rev. Lett. **108** (2012) 111601,
arXiv: `1111.5206 [hep-ph]` (cit. on p. 27).

[66] A. Denner, S. Dittmaier and L. Hofer,
COLLIER*: A fortran-based complex one-loop library in extended regularizations*,
Comput. Phys. Commun. **212** (2017) 220, arXiv: `1604.06792 [hep-ph]`
(cit. on p. 27).

[67] *Measurement of W± and Z-boson production cross sections in pp collisions at s=13 TeV with the ATLAS detector*, Physics Letters B **759** (2016) 601, ISSN: 0370-2693,
URL: `https://www.sciencedirect.com/science/article/pii/S0370269316302763` (cit. on p. 28).

[68] J. M. Campbell and R. K. Ellis,
Update on vector boson pair production at hadron colliders,
Phys. Rev. D **60** (1999) 113006, arXiv: `hep-ph/9905386` (cit. on p. 28).

[69] D. de Florian et al., *Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector*, (2016), arXiv: `1610.07922 [hep-ph]`
(cit. on p. 28).

[70] L. Devroye, *Non-Uniform Random Variate Generation*,
New York: Springer, 1986 (cit. on p. 37).

[71] L. Lyons, D. Gibaut and P. Clifford,
How to combine correlated estimates of a single physical quantity,
Nucl. Instrum. Meth. A **270** (1988) 110 (cit. on p. 38).

[72] ATLAS Collaboration, *Electron efficiency measurements with the ATLAS detector using 2012 LHC proton–proton collision data*, Eur. Phys. J. **C77** (2017) 195,
arXiv: `1612.01456 [hep-ex]` (cit. on pp. 42, 73).

[73] ATLAS Collaboration, *Muon reconstruction and identification efficiency in ATLAS using the full Run 2 pp collision data set at* $\sqrt{s} = 13\,\text{TeV}$,
Eur. Phys. J. C **81** (2021) 578 (cit. on p. 43).

[74] M. Cacciari, G. P. Salam and G. Soyez, *The anti-*k_t *jet clustering algorithm*,
JHEP **04** (2008) 063, arXiv: `0802.1189 [hep-ph]` (cit. on p. 43).

[75] M. Cacciari, G. P. Salam and G. Soyez, *FastJet user manual*,
Eur. Phys. J. C **72** (2012) 1896, arXiv: `1111.6097 [hep-ph]` (cit. on p. 43).

[76] ATLAS Collaboration,
Jet reconstruction and performance using particle flow with the ATLAS Detector,
Eur. Phys. J. C **77** (2017) 466, arXiv: `1703.10485 [hep-ex]` (cit. on p. 43).

[77] 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, arXiv: `1703.09665 [hep-ex]` (cit. on p. 43).

[78] ATLAS Collaboration,
Tagging and suppression of pileup jets with the ATLAS detector,
ATLAS-CONF-2014-018, 2014,
URL: `https://cds.cern.ch/record/1700870` (cit. on p. 43).

[79] *ATLAS flavour-tagging algorithms for the LHC Run 2 pp collision dataset*, (2022), arXiv: `2211.16345 [physics.data-an]` (cit. on pp. 43, 44).

[80] G. Aad et al.,
Performance of electron and photon triggers in ATLAS during LHC Run 2,
Eur. Phys. J. C **80** (2020) 47, arXiv: `1909.00761 [hep-ex]` (cit. on p. 47).

[81] G. Aad et al., *Performance of the ATLAS muon triggers in Run 2*,
JINST **15** (2020) P09015, arXiv: `2004.13447 [physics.ins-det]`
(cit. on p. 47).

[82] J. Butterworth et al., *PDF4LHC recommendations for LHC Run II*,
J. Phys. G **43** (2016) 023001, arXiv: `1510.03865 [hep-ph]` (cit. on p. 52).

[83] J. Bellm et al., *Herwig 7.1 Release Note*, (2017), arXiv: `1705.06919 [hep-ph]`
(cit. on p. 52).

[84] B. Andersson, G. Gustafson, G. Ingelman and T. Sjöstrand,
Parton fragmentation and string dynamics, Physics Reports **97** (1983) 31,
ISSN: 0370-1573, URL: `https://www.sciencedirect.com/science/article/pii/0370157383900807` (cit. on p. 52).

[85] B. Batool et al., *Measurement of the* $t\bar{t}\gamma$ *production cross-section in* pp *collision at* $\sqrt{s} = 13$ *TeV with the ATLAS detector*, tech. rep., CERN, 2017,
URL: `https://cds.cern.ch/record/2266485` (cit. on p. 53).

[86] *Muon reconstruction and identification efficiency in ATLAS using the full Run 2 pp collision data set at* $\sqrt{s} = 13$ *TeV*, Eur. Phys. J. C **81** (2021) 578,
arXiv: `2012.00578 [hep-ex]` (cit. on p. 53).

[87] ATLAS Collaboration, *Jet energy scale and resolution measured in proton–proton collisions at* $\sqrt{s} = 13$ *TeV with the ATLAS detector*, Eur. Phys. J. **C81** (2021) 689,
arXiv: `2007.02645 [hep-ex]` (cit. on p. 54).

[88] *ATLAS b-jet identification performance and efficiency measurement with* $t\bar{t}$ *events in pp collisions at* $\sqrt{s}$*=13 TeV*, Eur. Phys. J. C **79** (2019) 970,
arXiv: `1907.05120 [hep-ex]` (cit. on p. 55).

[89] ATLAS Collaboration,
Performance of algorithms that reconstruct missing transverse momentum in $\sqrt{s} = 8\,\text{TeV}$ proton–proton collisions in the ATLAS detector,
Eur. Phys. J. **C77** (2017) 241, arXiv: `1609.09324 [hep-ex]` (cit. on p. 56).

[90] J. Erdmann et al., *A likelihood-based reconstruction algorithm for top-quark pairs and the KLFitter framework*, Nucl. Instrum. Meth. A **748** (2014) 18,
arXiv: `1312.5595 [hep-ex]` (cit. on p. 58).

[91] T. Chwalek, *Messung der W-Boson-Helizitätsanteile in Top-Quark-Zerfällen mit dem CDF II Experiment und Studien zu einer frühen Messung des ttbar-Wirkungsquerschnitts mit dem CMS Experiment*, Presented 12 Feb 2010, 2010, URL: `https://cds.cern.ch/record/1416031` (cit. on p. 62).

[92] *Scale factors of hadron fake photon, obtained via private communication from PD Dr. Carmen Diez Pardos, University of Siegen* (cit. on p. 72).

[93] *Neural Network*,
URL: `https://www.ibm.com/topics/neural-networks`
(cit. on p. 81).

[94] F. Pedregosa et al., *Scikit-learn: Machine Learning in Python*,
Journal of Machine Learning Research **12** (2011) 2825 (cit. on p. 92).

[95] K. He, X. Zhang, S. Ren and J. Sun, *Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification*, (2015),
arXiv: `1502.01852 [cs.CV]` (cit. on p. 92).

[96] S. Ioffe and C. Szegedy, *Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift*, (2015),
arXiv: `1502.03167 [cs.LG]` (cit. on p. 93).

[97] *L2 Regularization*, URL:
`https://developers.google.com/machine-learning/crash-course/regularization-for-simplicity/l2-regularization`
(cit. on p. 93).

[98] F. Chollet et al., *Keras*, `https://keras.io`, 2015 (cit. on p. 93).

[99] Martín Abadi et al.,
TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems,
Software available from tensorflow.org, 2015,
URL: `https://www.tensorflow.org/` (cit. on p. 93).

[100] D. P. Kingma and J. Ba, *Adam: A Method for Stochastic Optimization*, (2014),
arXiv: `1412.6980 [cs.LG]` (cit. on p. 93).

[101] S. J. Reddi, S. Kale and S. Kumar, *On the Convergence of Adam and Beyond*, (2019), arXiv: `1904.09237 [cs.LG]` (cit. on p. 93).

[102] *ROC curve and AUC definitions*, URL: `https://en.wikipedia.org/wiki/Receiver_operating_characteristic` (cit. on p. 95).

[103] *EarlyStopping function*,
URL: `https://keras.io/api/callbacks/early_stopping/`
(cit. on p. 95).

[104] Olaf Behnke et. al. (Editors),
Data Analysis in High Energy Physics: A Practical Guide to Statistical Methods,
Weinheim, Germany: WILEY-VCH, 2013 37, 362 (cit. on pp. 102, 105).

[105] A. Shmakov et al., *SPANet: Generalized permutationless set assignment for particle physics using symmetry preserving attention*, SciPost Phys. **12** (2022) 178,
URL: `https://scipost.org/10.21468/SciPostPhys.12.5.178`
(cit. on p. 134).

List of Figures

List of Tables