Search for Charged-Lepton Flavour Violation in Top-Quark Interactions using $e\mu$ final states with the ATLAS Detector

Alexios Stampekis

A thesis submitted to the University of Birmingham for the degree of Doctor of Philosophy

Particle Physics Group,
School of Physics and Astronomy,
University of Birmingham.

January 16, 2024

To all who helped

and

to all that helped.

Abstract

A search for charged-lepton-flavour violation in top-quark production and decay is presented. The data analysed corresponds to 139 fb^{-1} of proton-proton collisions at a centre-of-mass energy of $\sqrt{s} = 13$ TeV delivered by the LHC and recorded by the ATLAS detector. Two processes are studied; the first is the production of a semileptonically decaying single top quark in association with a pair of opposite-sign different-flavour leptons, the second is the production of a top-quark pair, where one top quark decays into a pair of opposite-sign different-flavour leptons and a light *up*-type quark, while the other decays into a leptonically decaying W boson and a b-quark according to the Standard Model. The signal signature requires the presence of three leptons defined as electrons or muons, and is thus characterised by the presence of one b-jet, up to one light-quark jet, and either two electrons and one muon, or one electron and two muons. In the absence of a signal, an observed (expected) upper limit on the branching ratio of BR($t \to e\mu q$) $< 1.6\ (1.7) \times 10^{-7}$ is set at the 95% confidence level. An EFT interpretation is also discussed.

During Run 3 of LHC, the elevated luminosity and pileup conditions call for advanced electronic systems to efficiently trigger the physics objects. Among the new processing units, of high importance are the Feature Extractor (FEX) modules, responsible for identifying leptons and jets. For the commissioning of these modules, the design, testing and commissioning of a separate module, the FEX Test Module was needed. In this thesis, the process of testing and commissioning this FPGA-based digital system is discussed. Tests concerning the interpretation of incoming electrical signals as localised energy deposits in the ATLAS detector by the electron FEX module firmware are also discussed briefly.

Declaration

The ATLAS experiment relies on the effort of thousands of people who contribute to the detector technology, the development of data analysis tools and frameworks, and to the software distribution. In that spirit, I am heartily obliged to acknowledge the contribution of the ATLAS community to this thesis, without which it would not have been possible. I therefore acknowledge the contributions of:

- the authors, developers and contributors to the ROOT data analysis framework;
- the ATLAS Top-Quark Group for the development of ATLAS software specific for the needs of top-quark analyses;
- the developers of the TRExFitter framework for statistical analysis;
- the theorists/authors of Monte Carlo simulation packages;
- the CERN community in general for the operation, monitoring and upgrades of the LHC;
- the efforts of physicists throughout history, particularly those who led to and formulated the Standard Model of particle physics as well as theories which extend it.

Expanding on the work that I personally conducted and the contributions of other people, I briefly state:

- In chapter 2, important elements of the Standard Model (SM) of Particle Physics are discussed. The description follows well established textbooks, as well as related articles. In this chapter, when describing beyond the SM theories encompassing charged-lepton flavour violation (section 2.3.2) and when introducing the model-independent Effective Field Theory (EFT) approach (section 2.4), some of the description has been taken verbatim from ref. [1].
- In chapter 3, the ATLAS detector, with emphasis on the trigger system, is described. As mentioned in the beginning, ATLAS is a technological achievement of thousands of people around the world, mainly within or attached to the CERN community.
- My work on some of the digital systems of the ATLAS trigger system is presented in chapter 4. My work consisted of developing software for performing certain tests on the FTM (section 4.2) and eFEX (section 4.3) modules. These modules have been designed by members of the ATLAS Level-1 Calorimeter (L1Calo) Trigger group. This group is also responsible for the design, monitoring and maintenance of all hardware and (online/offline) software in the context of the L1Calo Trigger system.
- My work on physics analysis techniques is presented in chapter 5. The basis of the analysis was formed by an earlier team [1, 2], from the University of Bonn (referred to as the "80 fb^{-1} analysis team"). The event selection/analysis region definitions were studied by me. The multi-variate analysis was initially developed by the 80 fb^{-1} analysis team. I performed tests for re-optimising the technique for the new regions and the new data. The top-reconstruction was designed by

the 80 fb^{-1} analysis team. The fit was developed by me, but owes greatly to the TRExFitter team (as mentioned in the beginning) for providing support and a powerful framework. In the same spirit, I acknowledge the support of the ATLAS Top-Quark Group and Top-CLFV team. For the generation of Ntuples of data, the ATLAS software was used, specifically the *AnalysisTop* software of the Top-Quark Group.

In summary, I declare that the work presented in this document is original research conducted by me, and the works of people indispensable for the results presented will be properly and adequately referenced.

Acknowledgements

I think Alexander the Great would agree that when setting out on a campaign, you'd better choose your companions carefully and wisely. Maybe this analogy is not exact, as a lot of my helpers were discovered throughout this journey – a journey of hardships that is, but how else would it have been rewarding – but I hope the well-intended reader will get the picture.

Let me start by acknowledging the organisation which supported me financially throughout this campaign, namely the Science and Technology Facilities Council (STFC) of UK Research and Innovation (UKRI) organisation. Without their support nothing would have happened. I express my gratitude for all your efforts.

Let me start getting personal, by addressing my thanks to my supervisors, Chris Hawkes and Miriam Watson. Their constant support, guidance and, of course, patience were key elements for seeing this through.

I proceed with thanking all the Birmingham Particle Physics Group, guided by Paul, and more specifically the ATLAS group, guided by Dave. I have met valuable members of staff from other experiments too, such as Nigel from LHCb and Cristina from NA62 and LHCb.

In the context of the ATLAS group I had the chance to collaborate with the L1Calo team. I benefited greatly from working with Richard and Juraj for my qualification task. And of course Francesco, Steve, Alan and Paul T.

I don't forget the support from Mark S. concerning computing resources and programming skills. Nor Maria, our secretary, for taking care of so much of the administrative stuff (and being so friendly too).

On a more personal level, I cannot forget the group of PhD students that I met and had the chance to collaborate with, or just go for a walk, eat some humus or just chat casually. Rob W. (next desk), Antonio, Naomi and Tom have been a great company (yes, before the CoViD struck...) as my "same-year" gang. I am not excluding Cameron, because he was with us during the first year (and certainly for some nice bowling and mini-golf).

I have learned a lot from previous PhD students, Rob V., James K., Elliot, Russell, Dan B., Dan L., Nandish and Gov. From LCHb it was Kristian and Jonathan. With some of them we formed a great company when going for football as well as rock-climbing!

I was very fortunate to work with a very supportive Top-group in Birmingham. Rob V. has been a great collaborator, and a great friend. Jacob has offered lots of guidance and inspiration. Will G. was a great and tireless colleague, with a lot of insight in physics. I am glad he joined the Top-group.

Apart from the Birmingham Top-group, the "CERN" associates from other universities, Markus, Carlo and Gabriel, contributed significantly in the coding skills, physics insight and discussions around the analysis.

I separately want to thank Kostas Nikolopoulos, without whom I probably wouldn't have joined the Birmingham community. My warm thanks to him, because, even though maybe not apparent at first,

under the disguise of discipline, he is a an inspiring person of value.

In that spirit, my mind goes to the people who inspired me and guided me during my studies in Athens. Namely Dimitris Fassouliotis and Christine Kourkoumelis. Close to them, it has been Stelios and Panagiotis. I also want to thank all my professors in Athens, and a bit more specifically Eirene Mavrommatis and Theo Mertzimekis. Also the people of the Athens secretariat, friendly and supportive.

Of course, this thesis would not have been possible without the support of my family, relatives and friends. I thank my parents for their constant support and patience. My brothers, Dimitris and Thanos, for helping me see things from a more general perspective; high energy physics is not the only thing in life. My uncle and aunt, Andreas and Rena, for their warmth. My uncle and aunt in America, Perry and Robin. A source of inspiration, but also a warm hug. Close to them, my cousins, Nasos, Nicko and Michael. My grandmother Marika, in her hard times having always a nice word. But also being there throughout my life; my mind cannot help but go to my late grandfather Nasos, also a figure who raised me and my brothers. My late aunt Zozo, with her constant belief in me.

I want to take the chance to thank some friends who were also vital for my well-being and mental balance. I had a great time in the nature with Antonis; I wish you great times in Marathon. George has been a valuable friend for many years, and more to come... Konstantina has been a source of inspiration for me for many years.

I certainly cannot forget the many years that shaped a lot of my character and fortitude close to Athena. Thank you for your love, support and belief in me.

Approaching conclusion, my thoughts go to a special group of people, who formed for me kind of a “circle of life”, or elsely put, a different window to life, or simply a warm hug. It has helped me throughout the years to overcome adversity and hard times. My warm gratitude goes to Dafni, Chrysa, Eleni, Maria, Nancy, Vaso, Gogo, Aggeliki, Nikos, Pericles, George P. and Panagos. But also past members, Manos, George S., Marianna, Smaragda and Anastasia. I send a separate hug to Pyrros. And of course I address my sincere and loving gratitude to Ersi, for being there for so long.

Probably this has to end at some point. I just want to take the chance to say – and this is coming from a person who spends a lot of time on his own – that reaching out to people is good. And when that is not possible, then try to connect with what I would call “virtual” friends. May it be in history, philosophy, literature, art, cinema, music, sports, medicine, health (thank you McD.), find people trying to make a difference in a world of doom, people who inspire. Internet is there for you. It is a matter of choice.

Honor to those who in their lives
have set and guard Thermopylae
...

C.P. Cavafy, *Thermopylae* [1901,1903]

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The Standard Model (SM) of Particle Physics is a well understood theory that is able to predict with great precision a vast number of observed particle properties, but it is far from being a final theory of the world that surrounds us. Cosmological observations provide compelling evidence of the existence of matter of an unknown nature, the so called Dark Matter, which accounts for around 85 % of all matter in the universe. But, even before resorting to dark matter, the Standard Model does not predict some crucial phenomena, such as the matter/antimatter imbalance in the cosmos.

Minimal extensions to the SM are required in order to accommodate massive neutrinos, prompted by the observation of neutrino oscillations, which have proven that there is flavour mixing in the lepton sector. That, and the established quark flavour mixing within the SM, motivate the search for lepton flavour violation among the charged leptons. Such searches, in the decay of many different particles, have been going on for decades. Searches for charged-lepton flavour violation (CLFV) in top-quark decays have been difficult until recent years, but with the advent of the LHC there is now a good opportunity to perform such searches.

In this thesis, we start in chapter 2 by providing a brief overview of the main principles of the SM, as well as a theoretical description and motivation for CLFV searches. In chapter 3 we provide a description of the ATLAS detector, one of the biggest experiments on Earth, where research in particle physics reaches its furthest frontiers. A crucial component of the ATLAS detector is its triggering system, responsible for selecting events occurring from proton-proton collisions, that are of the greatest interest for new physics searches, among a huge number of ordinary, well understood, collision products. The focus of that chapter is on the trigger system of ATLAS, since a contribution has been made by the author to its upgrade program. The work performed on some of the latest digital systems, which process data from various detector components or simulate them, is presented in chapter 4. In chapter 5, we proceed to the description of the analysis methods followed in the search for CLFV in top-quark production and decay using $e\mu$ final states.

Chapter 2

Theory

2.1 The Standard Model of Particle Physics

Introduction

The following section outlines some of the basic elements of the Standard Model (SM) of particle physics, and is based on well established textbooks [3–6]. Supplementary reading may be provided by [7, 8].

The Standard Model is particle physicists' best theory at this time, and can describe with great precision the majority of the matter that we can observe and the interactions that govern it, at a fundamental level. According to this model, all matter is built from a small number of fundamental spin-$\frac{1}{2}$ particles, or fermions: six quarks and six leptons. For every fermion there is a corresponding antiparticle, identical to the former, but with opposite charge, parity[1], magnetic moment and other intrinsic quantum numbers (e.g. "lepton number" for the leptons and "strangeness" for the quarks).

There exist four kinds of interactions between fermions: the strong, the weak, the electromagnetic and the gravitational[2]. Each interaction is mediated by one or more massive or massless spin-1 particles, the vector gauge bosons. The strong interaction is mediated by the *gluons*, the charged weak by the $W^{\pm}$ bosons, the neutral weak by the Z boson, while the electromagnetic interaction is mediated by the *photon*.

The electromagnetic and the weak interactions are different manifestations of a unified electroweak interaction. In the current list of particles, the Higgs boson has been added most recently. It is a scalar, spin-0, particle which occurs as a consequence of the Higgs mechanism, which is responsible for the masses of the vector bosons and fermions (fig. 2.1).

Symmetry groups of the SM

The SM is best described in the Lagrangian formalism. The **electromagnetic** (EM) interactions arise from the requirement that the Quantum Electrodynamics (QED) Lagrangian be invariant under local gauge transformations of the $U(1)_Q$ symmetry group, Q being the generator of the group and identified with the conserved quantum number which is the charge of the particles, in units of the elementary charge e. The Lagrangian of **Quantum Chromodynamics** (QCD), the theory that describes the strong interactions of quarks and gluons, is constructed in a similar fashion, with the symmetry group now being $SU(3)_C$. The conserved quantum number is now the color charge, as quarks appear in three color states. Gluons are also colored particles, which leads to gluonic self-interactions.

Cumulative experimental evidence led physicists to describe the charged **weak** interactions by the

[1] In contrast, anti-bosons carry the same parity as bosons.
[2] The gravitational force is not described by the Standard Model.

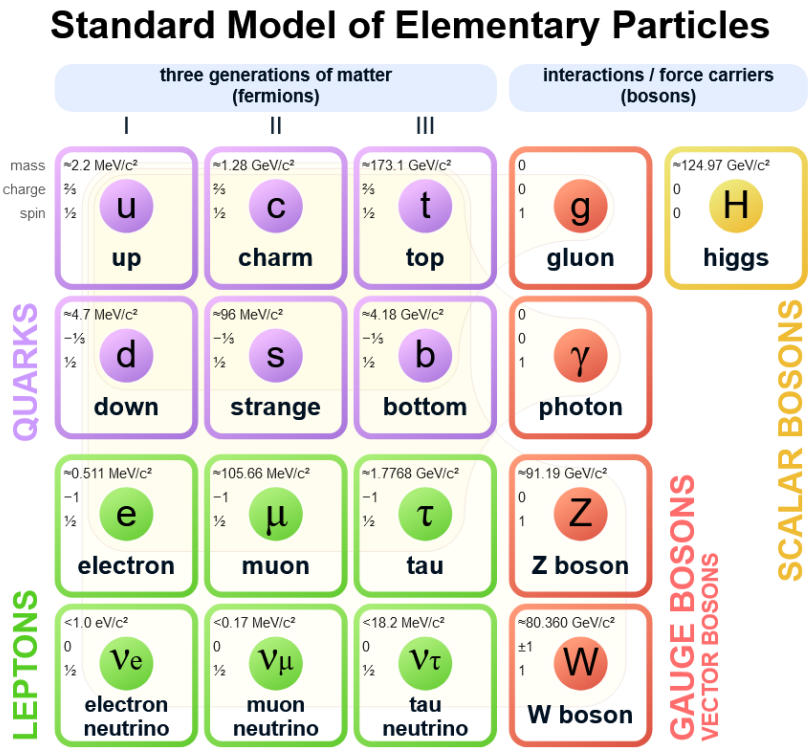


Figure 2.1: The elementary particles of the Standard Model. The three generations of matter particles (quarks and leptons), the force carriers (vector gauge bosons) and the scalar Higgs boson [9].

$SU(2)_L$ symmetry group of weak isospin. Only left-handed fermions and right-handed anti-fermions[3] take part in the charged weak interaction, hence the subscript L. These are arranged in $SU(2)$ doublets and assigned the quantum number called weak isospin (symbolised by T^3). The neutral weak interaction involves right-handed particles (and left-handed anti-particles) as well. The neutral weak and EM interactions combined form a $U(1)_Y$ symmetry group, with the generator being the weak hypercharge operator Y, which obeys the Gell-Mann–Nishijima formula

$$Q = T^3 + \frac{Y}{2} \,. \tag{2.1}$$

Electroweak unification becomes possible through the enlarged $SU(2)_L \times U(1)_Y$ symmetry group.

Ultimately, the SM is a Quantum Field Theory (QFT) that describes the strong and electroweak interactions. The requirement that the Lagrangian of the Standard Model be invariant under local gauge transformations of the group $SU(3)_C \times SU(2)_L \times U(1)_Y$ gives rise to the interactions.

The calculation of interaction rates within the SM relies heavily on the use of quantum mechanical (QM) perturbation techniques. In that context, transitions between different states of particles' wave functions are described by path integrals, based on the description of the mediating particles (*propagators*) with Green's functions. A visual representation of these QM transitions is done by the use of *Feynman diagrams*. Calculations at first (or leading) order (LO) in perturbation theory usually offer a good idea concerning transition amplitudes. In order to obtain more precise estimates though, higher order

[3] Eigenstates of the left and right *chirality* operator. Chirality is the same as helicity for massless particles.

corrections must be added, usually called next-to-leading order (NLO) and next-to-next-to-leading order (NNLO) corrections, when referring to second- and third-order perturbation theory calculations. These corrections involve additional Feynman diagrams, where the propagators form additional interactions, often in closed loops.

Renormalisation

The SM has the feature of being *renormalisable*. Renormalisation is a process of treating infinities that arise in the calculation of higher order (or "loop") transitions in perturbation theory. A key feature of renormalisation, is the fact that the coupling strengths of the interactions become functions of the energy transferred between the particles (hence often referred to as "running coupling constants"). A major difference when transitioning from an Abelian theory (such as EM) to a non-Abelian theory, such as QCD, is the fact that the self-interaction of the mediator bosons results in a different energy dependence of the coupling. The value of the coupling strength of QCD, α_s, decreases as the energy increases[4] (or equivalently as the distance becomes smaller). This behaviour is the source of *asymptotic freedom*, i.e. the fact that we can treat quarks as free particles in the high energy regime, when α_s is sufficiently small, so that perturbative techniques can be applied. At the other end of the spectrum, as distances between the quarks increase, the quarks and gluons are forced to be confined inside hadrons (a process observed in collider experiments by the detection of hadronic jets, which are formed by the *fragmentation* of the quarks initially produced into hadrons, via the splitting of the *gluon-strings* that connect two quarks). The top quark, further described in section 2.2, has a unique position among the other quarks as it decays before undergoing the hadronisation process described previously. Figure 2.2 shows the energy dependence of the strong coupling strength parameter, α_s, where the theoretical calculations of QCD are compared to experimental results. In fig. 2.3, the energy dependence of the coupling strengths of the SM is given.

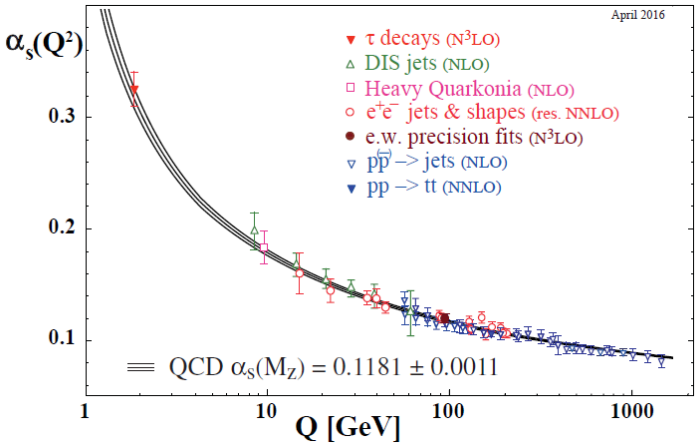


Figure 2.2: Summary of measurements of α_s as a function of the energy scale Q. The respective degree of QCD perturbation theory used in the extraction of α_s is indicated in brackets [NLO: next-to-leading order; NNLO: next-to-next-to leading order; res. NNLO: NNLO matched with resummed next-to-leading logs, accounting for soft and colinear gluon emissions (see also section 5.2.2); N^{3}LO: next-to-NNLO [10]].

[4]The weak coupling exhibits similar behaviour, though with a slower variation.

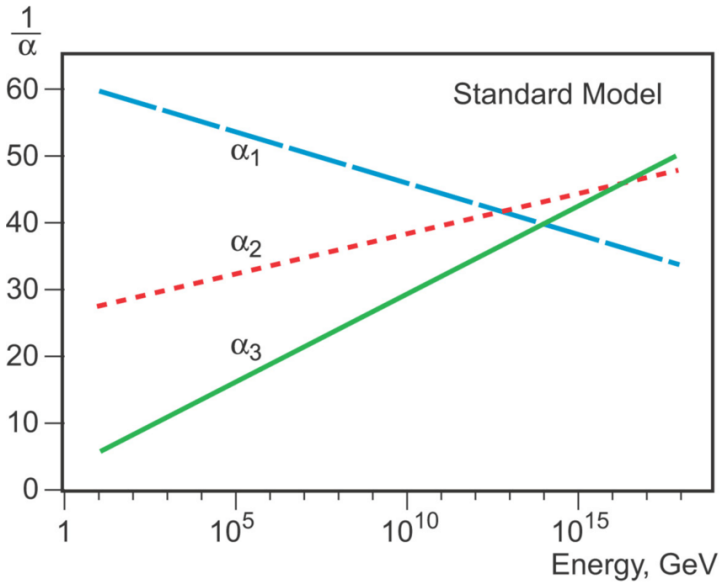


Figure 2.3: Reciprocal of the running coupling constants of the Standard Model interactions as a function of energy. The electromagnetic coupling (α_1) increases with energy scale, while the weak force (α_2) and the strong force (α_3) couplings decrease, reflecting the asymptotically free nature of the latter interactions. Taken from [11]. An attractive feature of certain beyond the Standard Model theories attempting to unify the electroweak and strong interactions, particularly those encompassing supersymmetry, is the convergence of the coupling strengths in one value at some high energy regime.

Higgs mechanism

The imposition of local gauge symmetry implies the existence of massless vector particles. For QED and QCD, we know that these bosons (the photons and gluons) are actually massless, but the weak force mediators, W and Z, are massive. To obtain massive vector bosons, the electroweak gauge symmetry must be broken somehow. If we introduce explicit symmetry breaking terms, in the form of arbitrary gauge boson masses, we alter the high-energy behaviour of the theory in such a way that the renormalisability of the theory is lost. This contradiction is resolved via spontaneous breaking of the symmetry. This is achieved by the *Higgs mechanism*: a single Higgs weak isospin doublet is introduced, with its vacuum state chosen in such a way that, after breaking the symmetry, we can acquire the $U(1)$ of electromagnetism (with a massless photon as the mediator) and be left with three massive gauge bosons, which are the mediators of the weak force. These are the $W^{\pm}$ and Z, whose masses obey the relation $M_W/M_Z = \cos\theta_W$, where θ_W is the weak mixing angle (which has to be determined experimentally).

An attractive and economical feature of the SM is that we can use the same Higgs doublet to give masses to the fermions, by introducing Yukawa couplings between the fermions and the Higgs doublet. Yukawa couplings can be described as 3×3 matrices in the fermion generation (or family) space. Since in the SM neutrinos are conceived as being massless, the diagonalisation of the Yukawa matrices for leptons is very simple, leaving no room for any cross-generational mixing between the leptons; hence the lepton family number, ascribed to the leptons as in table 2.1, is conserved [12].

On the other hand, it is well established that the coupling of the W to quarks is not as simple as in the leptonic case. Although the generation structure is similar, the charged weak interactions of quarks

do not strictly respect their separate identities. The Cabibbo-Kobayashi-Maskawa (CKM) matrix [10] relates the weak eigenstates to the mass eigenstates in the quark sector. Hence there is cross-generation mixing for the quarks and for charged currents[5].

Hence, in the quark sector, diagonalising the Yukawa matrices involves the CKM matrix, which allows for quark flavour violation (although the total baryon number is conserved in total).

Generation	I	II	III
Charged Leptons	e	μ	τ
Neutral Leptons	ν_e	ν_μ	ν_τ
L_e	1	0	0
L_μ	0	1	0
L_τ	0	0	1
L	1	1	1

Table 2.1: The lepton family quantum number assigned to the leptons of the SM. A "total" lepton number $L = 1$ is assigned to all leptons. For anti-leptons these numbers have the opposite value.

2.2 Elements of top quark physics

The top quark was the last quark to be discovered in 1995 by the CDS [13] and DØ [14] experiments at Fermilab. The top quark is the heaviest quark, with a measured mass of $m_t = 172.69 \pm 0.30$ GeV [15].

Top quarks are copiously produced in $t\bar{t}$ events, in proton-proton (pp) collisions at CERN's Large Hadron Collider (LHC), with a cross-section, at pp centre-of-mass energy of 13 TeV, measured by the ATLAS experiment to be $836 \pm 1(\text{stat.}) \pm 20(\text{syst.})$ pb [16], in excellent agreement with the theoretical prediction of 832^{+40}_{-45} pb at NNLO in QCD [16]. The cross-sections of $t\bar{t}$ and other top-related processes produced at the LHC are seen in fig. 2.4. The $t\bar{t}$ production Feynman diagrams at LO, in pp collisions, are shown in fig. 2.5.

Top quarks are also produced in single-top processes in the LHC. The main contribution comes from single-top production in the t-channel ($\sim 73\%$), followed by single-top production in association with a W boson, tW ($\sim 23\%$). Feynman diagrams at LO for these processes, as well as for single-top produced in association with a Z boson (tZq), and single-top production in the s-channel, are shown in figs. 2.6 and 2.7. Finally, example Feynman diagrams, at LO, for $t\bar{t}$ production in association with a vector boson ($t\bar{t}W$, $t\bar{t}Z$) are shown in fig. 2.8.

Top quarks decay via the weak interaction, producing almost always a W boson and a b-quark. Decays comprising the other "down-type" (down and strange) quarks are heavily suppressed by the relevant CKM matrix elements. Due to its high mass, the top quark decays rapidly, with a mean lifetime of around 5×10^{-25} s, before it has the chance to hadronise. The decay modes are categorised according to the decay channel of the W boson, which decays leptonically with a branching ratio (BR) of around $\frac{1}{3}$, and hadronically, to two quarks, with a BR of around $\frac{2}{3}$. A Feynman diagram summarising the possible decay channels of $t\bar{t}$ is shown in fig. 2.9; their relative contributions are summarised in fig. 2.10.

[5] In neutral weak interactions, mediated by the Z, such flavour changing is forbidden at tree level, and strongly suppressed at loop level (by the GIM mechanism, named after Glashow, Iliopoulos and Mayani who proposed it in 1970 [10]). Searches for flavour-changing neutral currents (FCNC) can probe extensions to the SM.

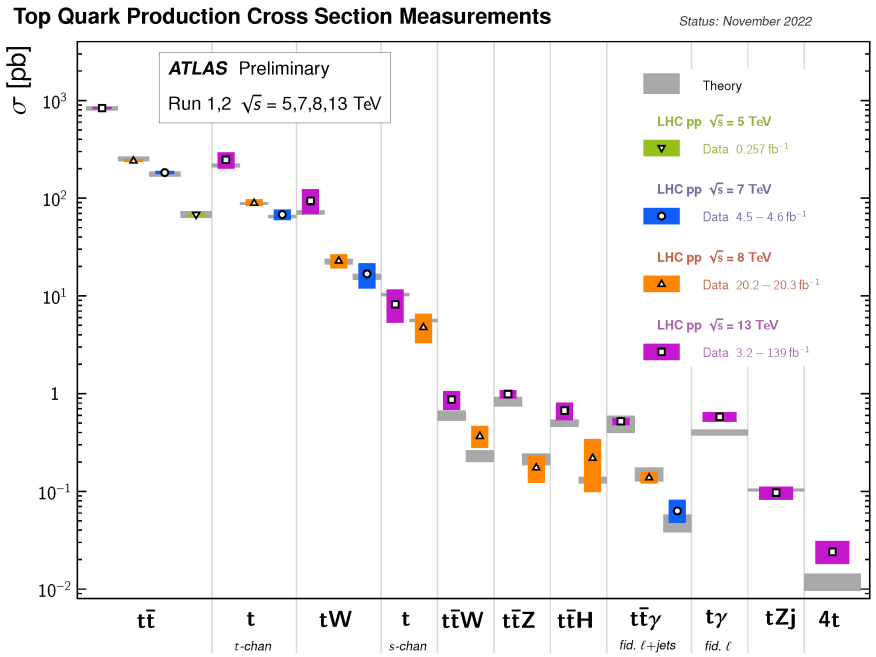


Figure 2.4: Summary of several top-quark related production cross-section measurements, compared to the corresponding theoretical expectations. All theoretical expectations were calculated at NLO or higher. Fiducial (fid.) cross section measurements are derived by considering the "final-state" particles which are produced by the top-quark decay, usually in a restricted phase space. From [16].

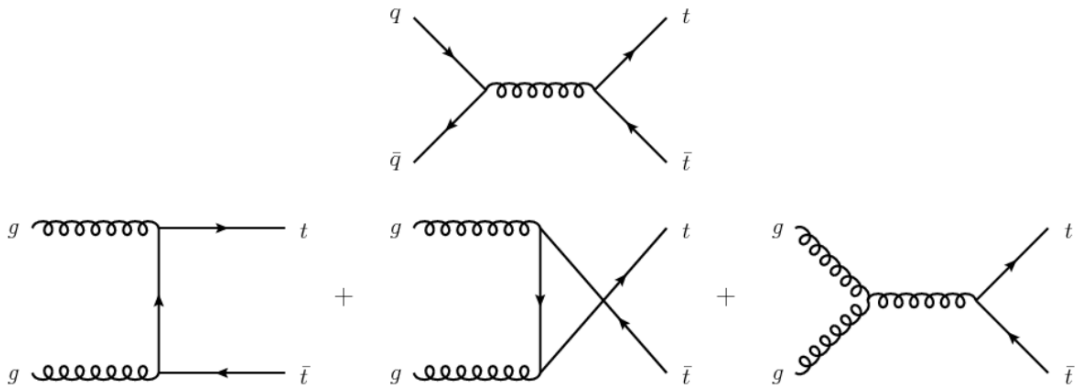


Figure 2.5: Leading-order $t\bar{t}$ production Feynman diagrams. The gluonic (bottom diagrams) are the dominant production modes (90%). Adapted from [17].

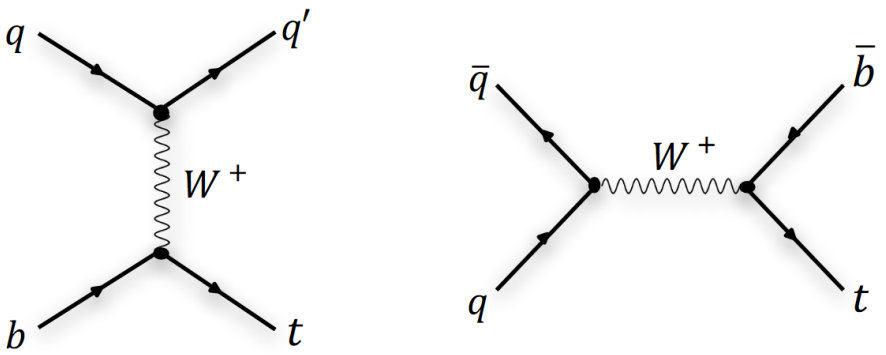


Figure 2.6: Leading-order example Feynman diagrams for single-top production in t-channel and s-channel. From [17].

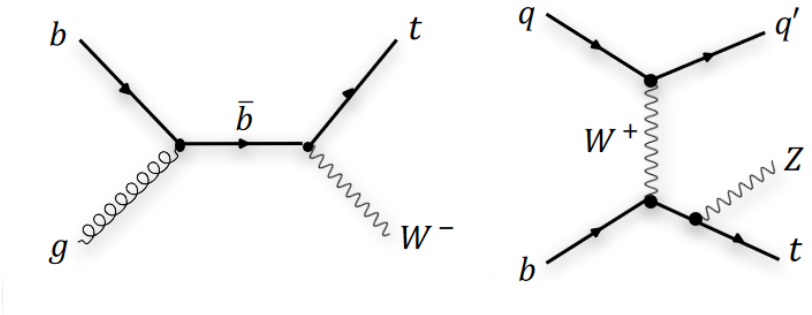


Figure 2.7: Leading-order example Feynman diagrams for tW and tZq production. From [17].

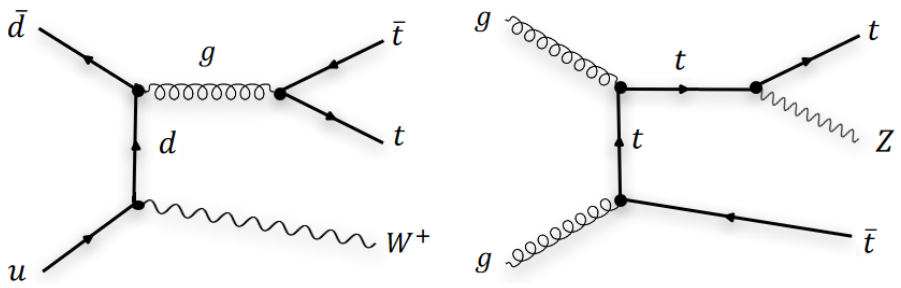


Figure 2.8: Leading-order example Feynman diagrams for $t\bar{t}V$ production. From [17].

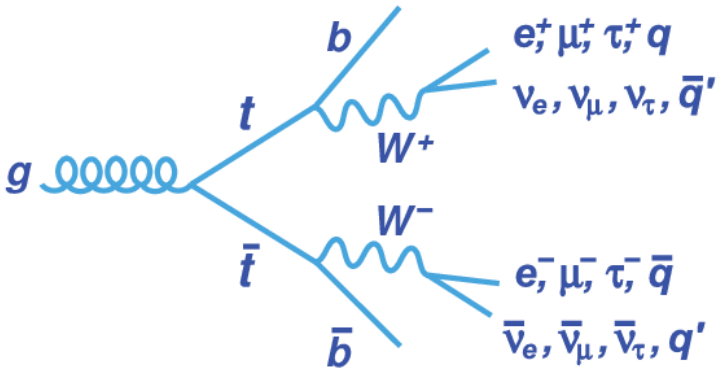


Figure 2.9: Decay modes of $t\bar{t} \rightarrow W^+bW^-\bar{b}$. Each of the W bosons can decay hadronically or leptonically. From [18].

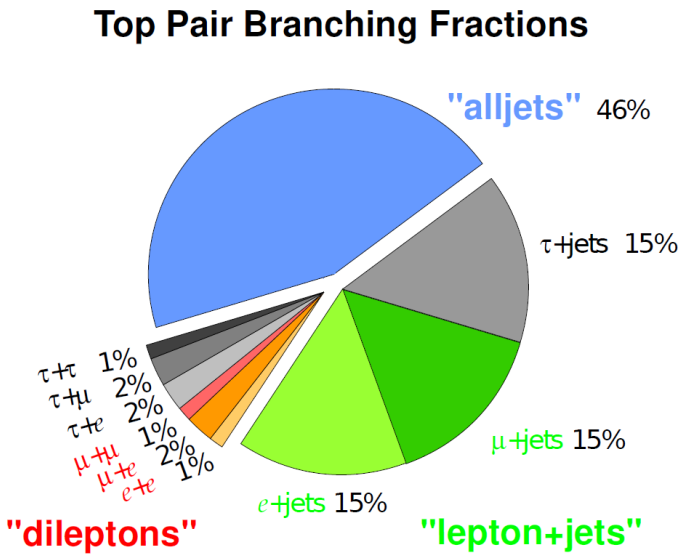


Figure 2.10: Relative contributions of the different decay modes of $t\bar{t}$. From [17].

2.3 Theoretical introduction to charged-lepton flavour violation

2.3.1 Massive neutrinos and charged-lepton flavour violation

Lepton family (or flavour, or generation) number is a global symmetry of the SM, but it should be regarded as an *accidental symmetry*, arising from the gauge principles, the simple Higgs model and the fixed particle content of the theory [5]. Yet, the observation of neutrino oscillations is a manifestation of lepton flavour violation in the neutral lepton sector. Neutrinos can convert from one flavour to another (for instance $\nu_e \leftrightarrow \nu_\mu$), which means that neutrinos have non-zero mass and that the lepton family numbers are not separately conserved. Neutrino oscillations resolve the solar neutrino problem [19] and suggest that modest changes are required to the Standard Model by introducing massive Dirac neutrinos (and hence the presence of right-handed neutrinos). Neutrinos interact as flavour eigenstates but they propagate as eigenstates of the free Hamiltonian, i.e. the mass eigenstates. The neutrino flavour and mass eigenstates are related by the PMNS matrix [20], named after Pontecorvo, Maki, Nakagawa and Sakata (analogous to the CKM quark flavour mixing matrix). Mass terms for Majorana neutrinos (i.e. a particle which is its own antiparticle), can easily be accommodated by the SM. In that scenario though, the total lepton number L would be violated as well [21].

Neutrino mixing allows charged lepton flavour violation (CLFV) processes, such as the muon decay $\mu \to e\gamma$, at loop level, as in fig. 2.11 . The resulting branching fraction for this process is however very small, of the order 10^{-55}, about 40 orders of magnitude lower than the sensitivity of present-day experiments (for a calculation see either [5] or [12]). The CLFV $\mu \to e\gamma$ decay is strongly suppressed by the tiny values of neutrino masses compared to the W boson mass, even for large PMNS matrix elements. This means that an experimental observation of CLFV would provide an unambiguous sign of new physics beyond the simplest extension to the SM.

Experimental searches for CLFV processes have been performed in numerous decay modes of various particles, e.g. $\mu \to e\gamma$, $\tau \to e\gamma$, $\tau \to \mu\gamma$, $\mu \to 3e$, $\tau \to 3\mu$ and many more in meson (π, K, D, B etc.)

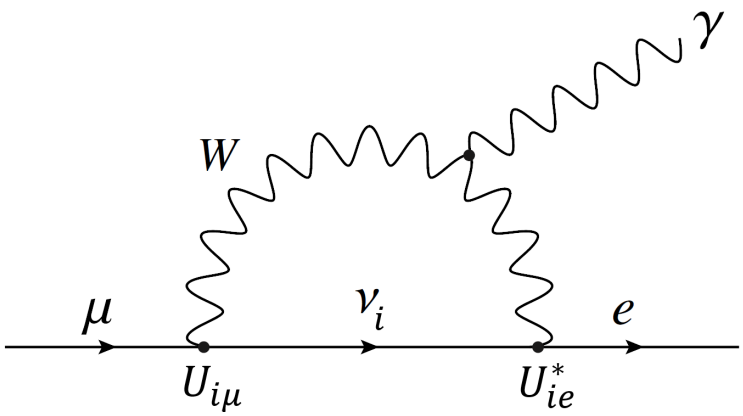


Figure 2.11: One-loop Feynman diagram that contributes to $\mu \to e\gamma$ in the SM extended with right-handed (Dirac) neutrinos. $U_{i\mu}, U_{ie}^*$ are PMNS matrix elements, accounting for neutrino flavour mixing. In addition to this diagram, one also has to consider 1-loop Feynman diagrams with the photon line attached to the external leptons. Adapted from [5].

decays [10, 12]. Processes involving the top quark have not yet been studied extensively, apart from searches with the HERA ep collider in single-top production [22].

ATLAS is contributing to the experimental effort of searching for CLFV in the decays of the τ lepton [23], the Z boson [24, 25], the Higgs boson [26], the top quark [2, 27], and other exotic processes [28, 29].

2.3.2 CLFV in beyond SM theories and the leptoquark model

In this section and the following one (section 2.4), some of the description concerning the leptoquark and Effective Field Theory approach has been taken verbatim from ref. [1].

A number of proposed theoretical models beyond the SM allow CLFV, including supersymmetric versions of the see-saw mechanism [30–32], Higgs-doublet models [33–35] and leptoquark models. We give a brief description of the leptoquark (LQ) model, which is of particular interest for the experimental search under study in this thesis (see section 2.4).

Leptoquarks are hypothetical gauge bosons, which can turn quarks into leptons and vice versa, of either a scalar (spin-0) or vector (spin-1) nature. Leptoquarks arise naturally in models that unify quarks and leptons (in gauge symmetry groups such as $SU(5)$, $SO(10)$ and $SU(5) \times U(1)$, which lead to a Grand Unified Theory), in supersymmetric models, technicolor models, composite scenarios and even superstring theories [36–43].

Independently from the underlying model, the maximum number of leptoquark states is limited to twelve, if it is assumed that the interactions between SM fermions and leptoquarks are invariant under the SM gauge group and that the corresponding couplings are dimensionless. As a result of respecting the symmetries of the SM, some LQs can couple to quark pairs, but none can couple to lepton pairs. Also, the leptoquarks can be classified according to their intrinsic quantum numbers (e.g. weak isospin, hypercharge, spin, baryon number) as dictated by the SM symmetries [10, 44].

Both flavour-changing neutral currents (FCNCs) and CLFV processes, which are strongly suppressed in the SM, can be induced at tree-level (i.e. at leading-order) by leptoquarks. Interactions involving three quarks and a lepton, which violate the conservation of both baryon number and lepton number,

are allowed by scalar LQ exchange, and can lead to proton decay (e.g. $p \to \pi^0 e^+$) [45]. Interactions involving two quarks and two leptons, can allow the transition $u_i \to \ell^- \ell'^+ u_j$, as depicted in fig. 2.12.

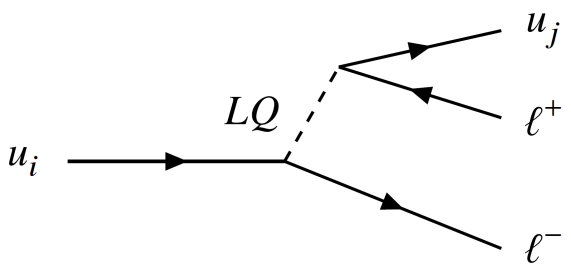


Figure 2.12: Decay of a heavy *up*-type quark, mediated by a scalar LQ.

Leptoquarks can also mediate various other processes, such as leptonic meson decays (such as π, K, D, D_S, $B_S^0 \to \ell\ell'$) and also semi-leptonic decays. Most measurements of these decays are compatible with the SM, placing further constraints on the LQ parameters. Exceptions have been seen in the decays $B^0 \to D^*\ell\nu$, $B^0 \to K^{*0}\ell\ell$ and $B \to K\ell\ell$, which showed deviations from the SM predictions at levels up to 3.8 standard deviations (3.8σ) [46–51]. Further tests of *lepton flavour universality* using Λ_b^0 decays have shown compatibility with the SM to within 1σ [52], while the latest results from LHCb for $B \to K^{(*)}\ell\ell$ are also consistent with the SM [53, 54]. In fact, S_1 and U_1 (scalar and vector LQs respectively, singlet representations of the $SU(2)_L$ group of weak isospin [44], hence the subscript 1) leptoquarks could explain both the aforementioned *B-anomalies* as well as potential FCNC in top quark processes [55, 56]. The top FCNC processes mediated by S_1 are closely related to the CLFV process that is studied here.

Finally, it is worth mentioning that the LQ interpretation is involved in other exciting tests of the SM predictions, such as the anomalous magnetic moment of the muon. The latest measurement from Fermilab [57] is in tension with the SM, and agrees with a previous result from the Brookhaven National Laboratory [58].

In order to interpret the various experimental results, a collection of different leptoquark models has to be applied, though all having in common the heavy mass of leptoquarks (over a few TeV), which would make any direct production of LQs out of reach of current accelerators. Thus, it is convenient to employ an effective field theory (EFT) approach, which has the advantage of being independent from the details of the underlying model.

2.4 Effective Field Theory approach

In general, an effective theory can be seen as an abstract version of the "complete" theory; an effective field theory captures the essential features of a field theory for a particular problem. It involves the presumed relevant degrees of freedom, and respects desired symmetries of the theory [59]. The degrees of freedom correspond to the light particles' fields, which are used to build the operators responsible for the experimental observables. The known or assumed symmetries regulate the nature of the allowed operators: all operators must be invariant under the EFT symmetries. Bearing in mind that the SM is valid up to a certain energy scale Λ of possible new physics, we can treat it as an EFT of a more

general and unknown theory. The effective Lagrangian of the whole theory can be written as

$$\mathcal{L}_{eff} = \mathcal{L}^{(4)}_{SM} + \mathcal{L}^{(5)} + \mathcal{L}^{(6)} + ... = \mathcal{L}^{(4)}_{SM} + \frac{1}{\Lambda}\sum_k c_k \mathcal{O}^{(5)}_k + \frac{1}{\Lambda^2}\sum_k c_k \mathcal{O}^{(6)}_k + ... \tag{2.2}$$

The higher the order of the expansion in negative powers of Λ, the more precise is the EFT. The effects of the physics introduced by higher order terms appear only as changes to the couplings and are suppressed by powers of the factor $1/\Lambda$. For each term of the expansion, labelled by the dimension[6] of the operators (in parentheses), several operators $\mathcal{O}$ can be present. Each is preceded by a coefficient c, called a *Wilson coefficient* [60] (WC), which is not predicted by the theory. For dimension 5, only one operator is allowed by the symmetries. It violates overall lepton number conservation, by generating massive Majorana neutrinos, and is not relevant to this analysis. The operators of dimension 6 are numerous. The set of 59 independent operators currently in use is referred to as the *Warsaw basis* [61]. Out of these 59, the 7 operators reported in table 2.2 describe four-fermion interactions involving two leptons and two quarks. If no restriction is applied by hand, these operators allow for charged lepton-flavour violation and/or flavour-changing neutral currents. Bounds on the corresponding Wilson coefficients have been placed from searches for rare muon and B, K meson decays and also from single top searches at the HERA collider (a compilation is provided in [62, 63]).

A previous study by ATLAS [2] targeted the decay $t \to \ell^{\pm}\ell'^{\mp}u_k$, with $u_k = \{u, c\}$ and $\ell = \{e, \mu\}$ in top-quark pair production, as depicted in fig. 2.13a, and used data corresponding to 80 fb^{-1} of pp collisions, collected by ATLAS during 2015-2017. The intention of the present analysis is to improve, finalise and extend the result of [2] by considering an additional signal process, the production of a single-top quark with a CLFV vertex $gu_k \to \ell^{\pm}\ell'^{\mp}t$, depicted in figs. 2.13b and 2.13c, by extending the data used to 139 fb^{-1} and by optimising the analysis techniques, mainly those concerning the background estimation.

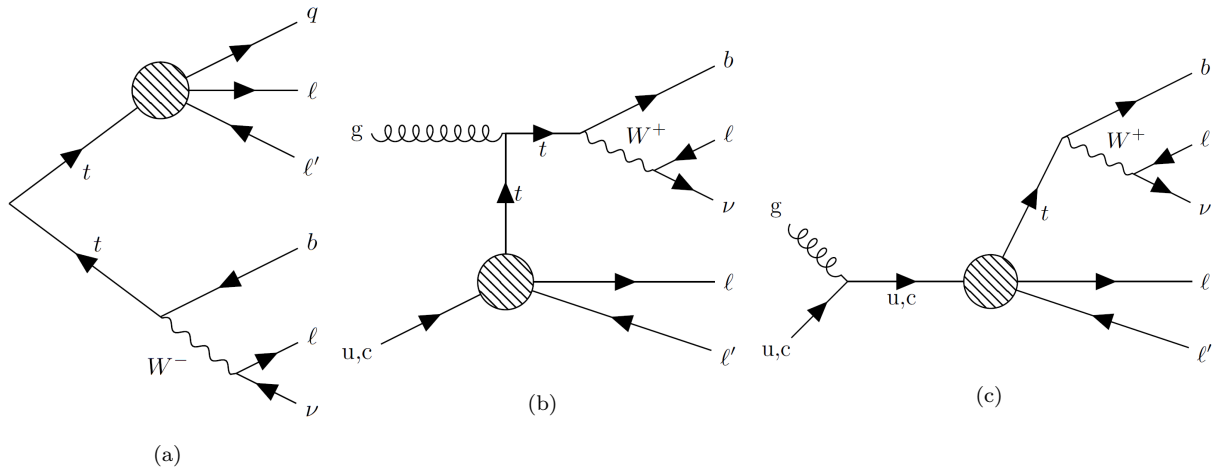


Figure 2.13: Example Feynman diagrams for the processes under study, with $\{\ell, \ell'\} = \{e, \mu\}$ or $\{\mu, e\}$. Top-quark pair production and CLFV decay process (fig. 2.13a), single-top production and CLFV process in t-channel (fig. 2.13b), and single-top production and CLFV process in s-channel (fig. 2.13c).

[6] In QFT formalism, it is customary to use the natural unit system, where $\hbar = c = 1$. In this unit system, all quantities are measured in units of energy to some power (the *dimension*). So, for example, all fermion fields are of dimension 3/2. The derivatives and particle masses are of dimension 1. The SM Lagrangian itself is of dimension 4. The dimension of each operator in the EFT Lagrangian is thus related to the energy dimension.

The $t \to \ell\ell' u_k$ decay and $g u_k \to t\ell\ell'$ production processes are described by the $SU(3)_C \times SU(2)_L \times U(1)_Y$ dimension-6 operators listed in table 2.2.

$\mathcal{O}_{lq}^{1(ijkl)}$	$= (\bar{l}_i \gamma^\mu l_j)(\bar{q}_k \gamma_\mu q_l)$
$\mathcal{O}_{lq}^{3(ijkl)}$	$= (\bar{l}_i \gamma^\mu \sigma^I l_j)(\bar{q}_k \gamma_\mu \sigma^I q_l)$
$\mathcal{O}_{eq}^{(ijkl)}$	$= (\bar{e}_i \gamma^\mu e_j)(\bar{q}_k \gamma_\mu q_l)$
$\mathcal{O}_{lu}^{(ijkl)}$	$= (\bar{l}_i \gamma^\mu l_j)(\bar{u}_k \gamma_\mu u_l)$
$\mathcal{O}_{eu}^{(ijkl)}$	$= (\bar{e}_i \gamma^\mu e_j)(\bar{u}_k \gamma_\mu u_l)$
${}^\ddagger\mathcal{O}_{lequ}^{1(ijkl)}$	$= (\bar{l}_i e_j)\varepsilon(\bar{q}_k u_l)$
${}^\ddagger\mathcal{O}_{lequ}^{3(ijkl)}$	$= (\bar{l}_i \sigma^{\mu\nu} e_j)\varepsilon(\bar{q}_k \sigma_{\mu\nu} u_l)$

Table 2.2: EFT operator basis and degrees of freedom. In the convention used, l and q are the left-handed lepton and quark doublets, while u and e are the right-handed up-type quark and charged lepton singlets. The indices ij and kl represent the lepton and quark generations respectively, hence there are numerous operators corresponding to the flavour of the fermions involved. The superscripts 1 and 3 do *not* refer to the dimension of the operators (all operators are of dimension 6), but are simply notations, corresponding to operators that could be combined as a linear sum (see e.g. [61]). σ^I are the Pauli matrices and $\varepsilon = i\sigma^2$, expressed by the 2$^{\text{nd}}$ Pauli matrix, is the isospin antisymmetric contractor ($\varepsilon_{12} = 1$). The ‡ symbol marks non-hermitian operators.

According to a paper by Chala et al. [64], the operators $\mathcal{O}_{lq}^1$ and $\mathcal{O}_{lq}^3$ can be combined into one single operator:

$$\mathcal{O}_{lq}^{-(ijkl)} \equiv \mathcal{O}_{lq}^{1(ijkl)} - \mathcal{O}_{lq}^{3(ijkl)} \,. \tag{2.3}$$

The basis of independent operators, a linear combination of which gives rise to the interactions between the physical fields, is referred to as degrees of freedom. Assigning a Wilson coefficient c to each of these operators, by adopting the convention used in ref. [64], the decay width of $t \to \ell\ell' u_k$ can be expressed in terms of 8 degrees of freedom as follows:

$$\begin{aligned}\Gamma(t \to \ell_i^+ \ell_j^- u_k) = \frac{m_t}{6144\pi^3}\left(\frac{m_t}{\Lambda}\right)^4 \Big\{ & 4|c_{lq}^{-(jik3)}|^2 + 4|c_{eq}^{(jik3)}|^2 + 4|c_{lu}^{(jik3)}|^2 + 4|c_{eu}^{(jik3)}|^2 \\ & + |c_{lequ}^{1(jik3)}|^2 + |c_{lequ}^{1(ij3k)}|^2 + 48|c_{lequ}^{3(jik3)}|^2 + 48|c_{lequ}^{3(ij3k)}|^2 \Big\} \,.\end{aligned} \tag{2.4}$$

According to the same ref. [64], the Wilson coefficients are assumed to be real, which implies that:

$$\begin{gathered} c_{lq}^{-(jilk)} = c_{lq}^{-(ijkl)},\ c_{eq}^{(jilk)} = c_{eq}^{(ijkl)},\ c_{lu}^{(jilk)} = c_{lu}^{(ijkl)},\ c_{eu}^{(jilk)} = c_{eu}^{(ijkl)}, \\ c_{lequ1}^{1(jilk)} = c_{lequ}^{1(ijkl)},\ c_{lequ}^{3(jilk)} = c_{lequ}^{3(ijkl)} \,. \end{gathered}$$

Hence we are essentially left with 6 degrees of freedom, which correspond to 6 different "EFT signal samples" used in this analysis (eu, eq, lu, lq, $lequ1$, $lequ3$).

A recent set of bounds on the Wilson coefficient for the $e\mu$ channel has been published in ref. [64] and can be seen in table 2.3. The bounds are obtained by reinterpreting the ATLAS $t \to qZ$ flavour-changing

neutral current search [65].

u_k	$c_{lq}^{-(ijk3)}$	$c_{eq}^{(ijk3)}$	$c_{lu}^{(ijk3)}$	$c_{eu}^{(ijk3)}$	$c_{lequ}^{1(ijk3)}$	$c_{lequ}^{1(ij3k)}$	$c_{lequ}^{3(jik3)}$	$c_{lequ}^{3(ij3k)}$
u	5.6	5.6	5.6	5.6	12.0	12.0	1.5	1.5
c	6.3	6.3	6.3	6.3	13.0	13.0	1.7	1.7

Table 2.3: Bounds on the degrees of freedom of the EFT operators of table 2.2 from Chala et al. [64] for the decay $t \to e\mu u_k$.

The CMS Collaboration has recently performed an analysis involving the same processes studied in this thesis, and using hadronic [66] and leptonic decays of the top quark [67], the latter having the same final state as that considered here. CMS uses the following formulae to define three different categories of interactions when setting limits. They are denoted as *scalar*, *vector* and *tensor* according to the Lorentz structure of the operators involved (listed already in table 2.2).

$$\mathcal{O}_{\text{scalar}} = \mathcal{O}_{lequ}^{(1)} \,, \tag{2.5}$$

$$\mathcal{O}_{\text{vector}} = \mathcal{O}_{lq} + \mathcal{O}_{lu} + \mathcal{O}_{eq} + \mathcal{O}_{eu} \,, \tag{2.6}$$

$$\mathcal{O}_{\text{tensor}} = \mathcal{O}_{lequ}^{(3)} \,. \tag{2.7}$$

In table 2.4, the upper limits on the respective Wilson coefficients, set by the CMS search [67], are presented.

CLFV coupling	Lorentz structure	$C_{e\mu tq}/\Lambda^2$ [TeV^{-2}] Exp.	Obs.
$e\mu\tau u$	tensor	0.019	0.020
	vector	0.037	0.041
	scalar	0.077	0.084
$e\mu\tau c$	tensor	0.061	0.068
	vector	0.130	0.144
	scalar	0.269	0.295

Table 2.4: Upper limits (at 95 % Confidence Level) on the Wilson coefficients of the tensor, vector and scalar operators, resulting from the CMS search [67].

The results from the analysis presented in this thesis can be transformed into upper limits on Wilson coefficients, which could be compared with those presented in tables 2.3 and 2.4. This work has not yet been done, but the procedure is explained in section 5.10.

Finally, it is worth noting that a concurrent analysis by ATLAS presents an extension of the lepton selection to include a final state with two muons and one hadronically decaying tau lepton [27]. The search for the $t\mu\tau u_k$ interaction is referred to as the $\boldsymbol{\mu\tau}$ **channel**, in contrast to the search for the $te\mu u_k$ interaction, referred to as the $\mathbf{e\mu}$ **channel**, discussed in this document. Since the EFT operators introducing CLFV interactions in top processes are flavour-dependent, these two channels target different operators.

Cross-sections of the signal processes

In tables 2.5 and 2.6, the cross-sections of the different signal processes considered in this thesis, as

given by Monte Carlo (MC) simulation (further discussed in section 5.2.2), are presented. The CLFV in top-quark decay (decay-signal) and in top-quark production (production-signal) processes are considered separately, and are further broken down by the EFT operator generating the interaction. The total cross section for each signal process (decay and production) is obtained by simply adding the separate contributions shown. For the decay-signal, the cross-section value matches a calculation done in [64], which uses eq. (2.4) and multiplying by a factor 3 to account for any lepton flavour in the CLFV decay. For the production-signal, no such formula is available, so we rely solely on MC simulation.

It is apparent that the decay-signal has a much smaller contribution than the production-signal. For the decay-signal, the up and charm quark contributions are similar (and can be derived by the MC generator information of the produced particles). For the production-signal, separate samples for the processes involving up and charm quarks have been produced. The up quark contribution is dominant, due to the prevalence of up quarks in the parton content of the colliding protons. For both production and decay channels, the tensor operator generates the dominant contribution.

	Cross section $\sigma^{+\text{scale}}_{-\text{scale}} \pm$ PDF (fb)			
$i,j = e,\mu,\tau$	$c_{lq}^{-(ijk3)}$	$c_{eq}^{(ijk3)}$	$c_{lu}^{(ijk3)}$	$c_{eu}^{(ijk3)}$
Production $\ell\ell' ut$	$118.2^{+23.7}_{-18.9} \pm 1.4$	$118.0^{+23.6}_{-18.9} \pm 1.4$	$118.3^{+23.7}_{-18.9} \pm 1.4$	$118.4^{+23.7}_{-18.9} \pm 1.4$
Production $\ell\ell' ct$	$7.9^{+1.2}_{-1.0} \pm 1.6$	$7.9^{+1.2}_{-1.0} \pm 1.6$	$7.9^{+1.2}_{-1.0} \pm 1.6$	$7.9^{+1.2}_{-1.0} \pm 1.7$
Decay $\ell\ell' qt$	$6.90^{+1.79}_{-1.31} \pm 0.07$	$6.90^{+1.79}_{-1.31} \pm 0.07$	$6.90^{+1.79}_{-1.31} \pm 0.06$	$6.90^{+1.79}_{-1.31} \pm 0.06$

Table 2.5: Theoretical cross sections, in fb, for single top-quark production and top-quark decays through CLFV interactions utilising vector EFT WCs. These cross sections assume a top-quark mass of 172.5 GeV, a top-quark decay width of 1.33 GeV, a new physics scale of $\Lambda = 1$ TeV and a WC value of 1.0. The scale and PDF uncertainties are given (discussed in section 5.2.2).

	Cross section $\sigma^{+\text{scale}}_{-\text{scale}} \pm$ PDF (fb)	
$i,j = e,\mu,\tau$	$c_{lequ}^{1(ijk3)}$	$c_{lequ}^{3(ijk3)}$
Production $\ell\ell' ut$	$101.2^{+21.3}_{-16.2} \pm 1.2$	$2145^{+408}_{-322} \pm 24$
Production $\ell\ell' ct$	$6.1^{+1.0}_{-0.8} \pm 1.5$	$153^{+21}_{-18} \pm 29$
Decay $\ell\ell' qt$	$3.46^{+0.90}_{-0.66} \pm 0.03$	$165.8^{+43.1}_{-31.5} \pm 1.5$

Table 2.6: Theoretical cross sections, in fb, for single top-quark production and top-quark decays through CLFV interactions utilising scalar and tensor EFT WCs. These cross sections assume a top-quark mass of 172.5 GeV, a top-quark decay width of 1.33 GeV, a new physics scale of $\Lambda = 1$ TeV and a WC value of 1.0. The scale and PDF uncertainties are given (discussed in section 5.2.2).

Chapter 3

The ATLAS Detector

In this chapter we provide a description of the ATLAS detector and its main components, after introducing some key features and concepts of CERN's Large Hadron Collider. Our focus will be the ATLAS Trigger system, which is described in section 3.3, since contributions were made to its upgrade during the ATLAS Phase-1 Upgrade Program.

3.1 The LHC

The Large Hadron Collider [68] (LHC) is a particle accelerator and storage ring, built and operated by the European Organisation for Nuclear Research (CERN), and located in the vicinity of Geneva, Switzerland. The collider is housed in a circular tunnel of 27 km in circumference, dug 50 m to 175 m deep underground, across the Franco-Swiss border. In the LHC, two beams of particles, protons or heavy ions, travel in vacuum along opposite directions inside two separate beam pipes. Proton-proton (pp) collisions are the main focus (with heavy ion collisions also), occurring at four points around the ring where detectors are situated (ALICE [69], CMS [70], LHCb [71] and ATLAS [72]). Figure 3.1 shows a representation of the LHC in its surrounding area. The aims are to undertake precision measurements of Standard Model parameters and search for phenomena beyond the Standard Model, such as supersymmetry and dark matter. This report focuses on the ATLAS (A Toroidal LHC ApparatuS) experiment.

Figure 3.1: Photograph of the location of the LHC, along with an artistic representation of the two oppositely directed beams of particles, which are accelerated around the same circular tunnel. The locations of the detectors are shown. The smaller ring on the left, showing blue accelerating particles, is the SPS. This is the last among many acceleration stages before the particles reach the LHC [73].

By design, the centre-of-mass energy for pp collisions at the LHC is $\sqrt{s} = 14$ TeV, corresponding to an energy of 7 TeV for each proton beam. The likelihood of collisions depends on the instantaneous luminosity $\mathcal{L}$ of the collider, which in turn can be related to the interaction rate R and the cross section

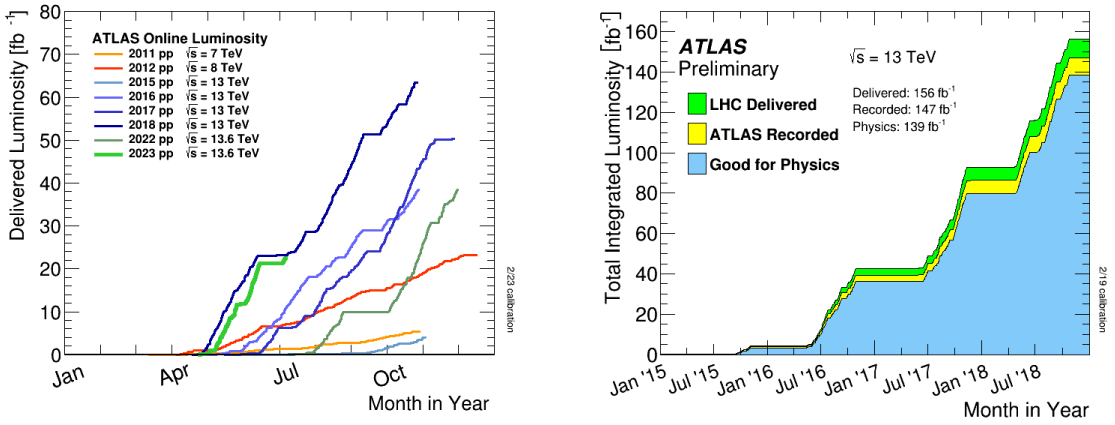
σ of pp interactions by $R = \sigma\mathcal{L}$. The LHC beams are circulated as "bunches" of protons; this gives rise to the following expression for the instantaneous luminosity:

$$\mathcal{L} = \frac{n_p^2 n_b f}{4\pi\sigma_x\sigma_y} ,$$

where n_b and n_p are the number of bunches and the number of protons per bunch, respectively, f is the revolution frequency and σ_x, σ_y correspond to the transverse beam dimensions at the interaction point [10]. Another common term used concerning particle collisions is the average number of interactions per bunch crossing, $\langle\mu\rangle$, called pileup [74].

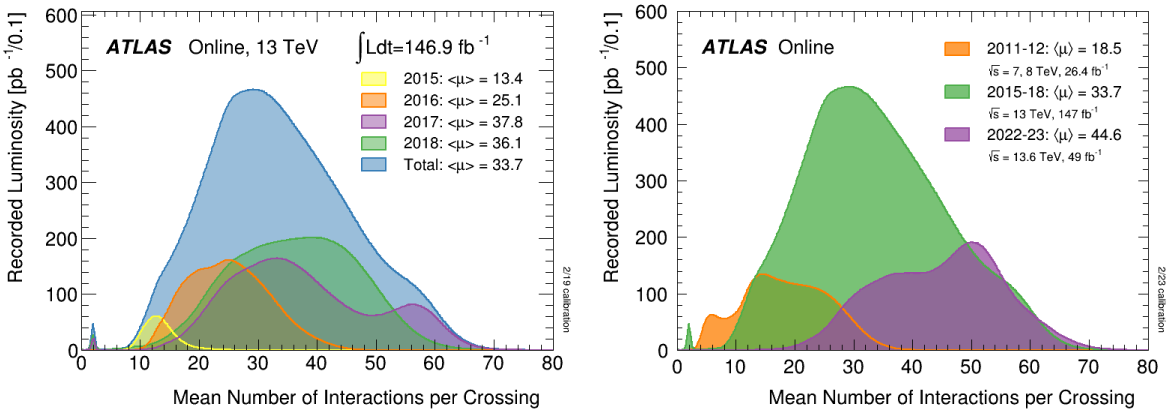
The LHC started its physics operation in March 2010, delivering pp collisions at $\sqrt{s} = 7$ TeV, later increased to $\sqrt{s} = 8$ TeV. The period of data-taking operation from 2010-2013 is called Run 1, and was followed by a Long Shutdown period (LS1) during 2013-2014, for accelerator and detector maintenance and upgrades. The second period of operation started in April 2015, lasted until the end of 2018, and is called Run 2. During Run 2 the pp collisions happened at $\sqrt{s} = 13$ TeV. The highest instantaneous luminosity recorded at ATLAS was around 2×10^{34} cm^{-2}s^{-1}, twice the design value. Proton bunches were separated by a 25 ns time interval. A second long shutdown (LS2) of the LHC lasted from the start of 2019 until the summer of 2022, during which a major upgrade program for the LHC and the detectors (called the Phase-1 Upgrade for ATLAS) took place. The operation of LHC resumed in July 2022, reaching a record $\sqrt{s} = 13.6$ TeV, and marking the start of Run 3, which is expected to last for close to four years. This will be followed by LS3, during which the Phase-2 Upgrade program is scheduled to be implemented, paving the way for the *High-Luminosity LHC* operation.

Figure 3.2 shows the time-integrated luminosity recorded by ATLAS, as it increased over time, while fig. 3.3 shows the pileup distributions during the different data-taking periods of ATLAS.



(a) (b)

Figure 3.2: (a) Cumulative integrated luminosity versus day delivered to ATLAS during periods with stable beams and for high energy pp collisions. (b) Cumulative integrated 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. From [74, 75], including the caption.



(a) (b)

Figure 3.3: (a) Luminosity-weighted distribution of the mean number of interactions per bunch crossing for the 2015-2018 pp collision data at 13 TeV centre-of-mass energy. In (b) the Run 1 and Run 3 (up to summer 2023) values are also included. From [74, 75].

3.2 The ATLAS Detector

3.2.1 Overview

The main references for the following are [76] and [72], with further resources mentioned as inline citations. The ATLAS detector is a multipurpose particle detector with approximately forward-backward symmetric cylindrical geometry. It consists of 4 main parts: the inner detector, the electromagnetic and hadronic calorimeters, the muon spectrometer and the magnet system. Moreover, a trigger system, necessary to cope with the high luminosity of the LHC, used for filtering the data to be kept for further processing, is integrated into the inner detector, calorimeter and muon systems. The detector has a total length of about 44 metres, a diameter of about 25 metres and weighs around 7,000 tonnes. An overall representation of ATLAS is given in fig. 3.4.

The coordinate system

A right-handed Cartesian coordinate system is used in ATLAS, with its origin at the nominal interaction point. The x-axis points to the centre of the LHC ring, the y-axis points upwards and the z-axis is oriented along the beam pipe. Spherical polar coordinates are commonly used, where the azimuthal angle ϕ (ranging between $-\pi$ and π) is defined in the x-y transverse plane and the polar angle θ is measured from the z-axis. It is convenient to use the pseudorapidity $\eta = -\ln(\tan(\theta/2))$; when considering massless particles, this is equal to the rapidity y

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

where E is the energy of the particle and p_z is its longitudinal momentum. Differences in rapidity y are invariant under Lorentz boosts along the z-axis, unlike the polar angle θ. The pseudorapidity η gives a good approximation to y for an ultra-relativistic particle with a momentum that is much larger than

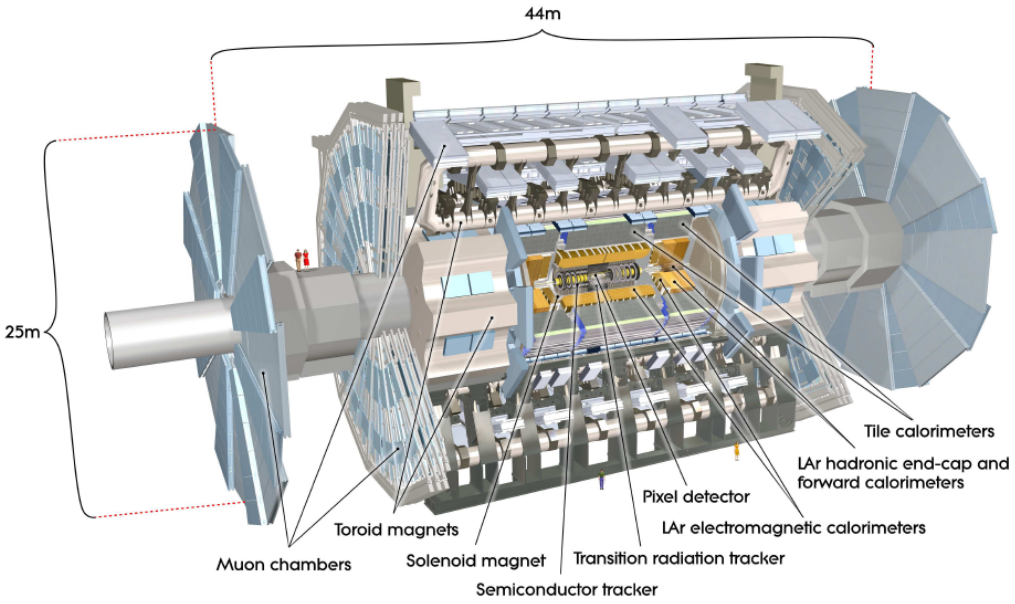


Figure 3.4: The ATLAS detector [72].

its mass. Distance in the η-ϕ plane can be measured by the approximately boost-invariant variable

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

As the total momentum is zero transverse to the beam axis, but not along it[1] , it is usual to use transverse quantities to describe particle trajectories and properties, such as transverse momentum p_{T} and transverse energy E_{T}, as well as transverse quantities for a whole event (such as missing transverse energy $E_{\mathrm{T}}^{\mathrm{miss}}$ and total transverse energy $\sum E_{\mathrm{T}}$)[2]. The T subscript signifies the transverse (parallel to the x-y plane) component of these quantities.

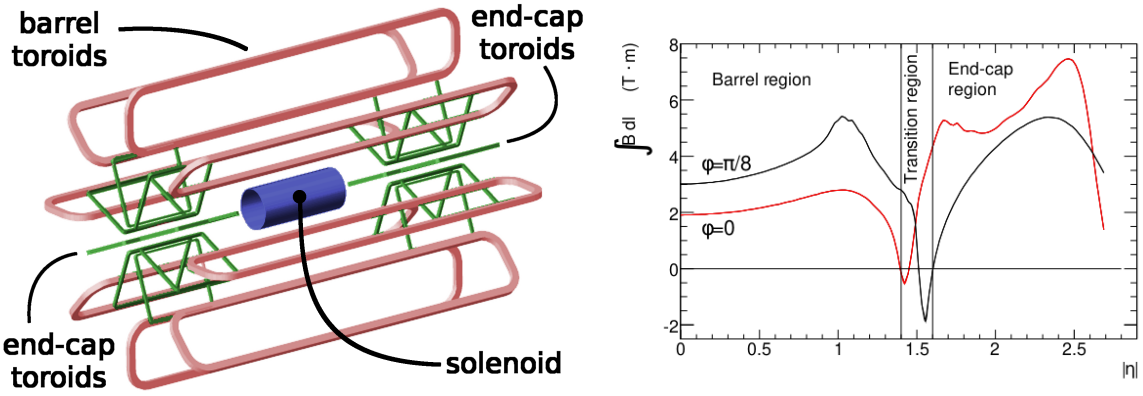
3.2.2 Magnet System

The ATLAS magnet system provides the bending power necessary for the measurement of the momentum of charged particles. The system consists of 4 superconducting magnets: the central solenoid (CS), which surrounds the inner detector, and three large air-core toroids generating the magnetic field for the muon spectrometer. The two end-cap toroids (ECT) are inserted inside the barrel toroid (BT) at each end, and line up with the CS (Figure 3.5a). The CS provides a central field of 2 T, with a peak magnetic field of 2.6 T at the superconductor itself. The peak magnetic fields on the superconductors in the BT and ECT are 3.9 and 4.1 T, respectively (the mean fields are approximately 0.5 and 1 T, respectively). The performance in terms of bending power is characterised by the field integral $\int Bdl$, where B is the azimuthal field component, and the integral is taken on a straight line trajectory between the inner and outer radius of the toroids. The BT provides 2 to 6 Tm and the ECT contributes with 4 to 8 Tm in the 0.0–1.3 and 1.6–2.7 pseudorapidity ranges, respectively.

[1] Protons are composite particles and events arise from interactions between constituent particles, or partons, within the colliding protons, which carry a fraction of the momentum of the colliding protons, according to the *Parton Distribution Functions* (discussed in section 5.2.2).

[2] Complete definitions are given in section 5.3.7.

The bending power is lower in the transition regions ($1.3 < |\eta| < 1.6$) where the two magnets overlap (Figure 3.5b).



(a) (b)

Figure 3.5: The ATLAS magnet system (a) and bending power of the toroid system as a function of $|\eta|$ (b) [72].

3.2.3 Inner Detector

The Inner Detector (ID) is located in the inner part of ATLAS and provides geometrical coverage of $|\eta| < 2.5$. It is designed for the measurement of tracks of charged particles, as well as for the measurement of primary and secondary vertices. The ID is contained inside a cylinder of 7 m length and 1.15 m radius, and immersed in the axial magnetic field of 2 T, produced by the central solenoid (fig. 3.6).

The ID consists of 3 sub-systems, as noted in the following, from inner to outer radii. The **pixel detector** is closest to the interaction point (IP), consists of 3 cylindrical layers in the barrel region ($|\eta| < 1.4$) and 3 discs in the end-cap regions, and is based on semiconducting silicon pixel sensor technology. A further cylindrical layer, called the insertable B-layer (IBL), was inserted closest to the beam pipe before the start of Run 2 [77, 78]. The **Semi-Conductor Tracker** (SCT) consists of ∼16000 modules arranged in 4 barrel layers, and 9 endcap discs on each side of the detector. Each module contains silicon microstrips of ∼12 cm length that produce an electrical pulse whenever a charged particle passes through. The **Transition Radiation Tracker** (TRT), has been designed to provide tracking of charged particles, as well as particle identification capabilities. The TRT consists of a barrel and 2 end-cap parts, and is based on the use of straw tubes, constructed from polyimide and carbon, and filled with a xenon gas mixture. The TRT spans up to $|\eta| = 2.0$.

3.2.4 Calorimeter System

The ATLAS calorimeters use 2 technologies: (1) high granularity liquid argon (LAr) sampling calorimeters and (2) steel/tile scintillator sampling calorimeters. The calorimeters are placed outside the CS and consist of a barrel part, 2 end-cap parts, and 2 forward calorimeters (FCal). The tile calorimeters consist of a barrel and 2 extended barrel parts. The whole calorimeter system covers the area $|\eta| < 4.9$.

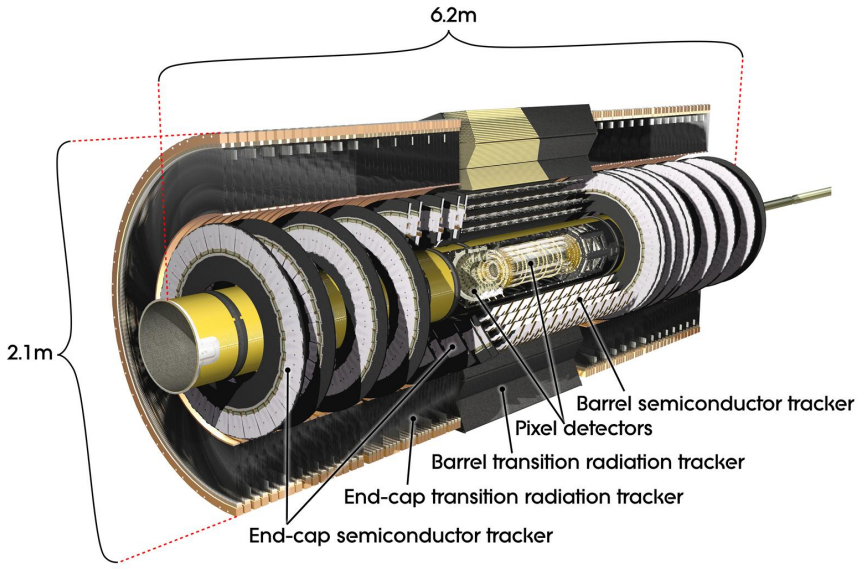


Figure 3.6: Representation of the ATLAS Inner Detector [79].

The precision LAr electromagnetic calorimeters use accordion geometry for the absorber (lead), with 4 read-out layers in the barrel region, 3 layers in intermediate η ($1.8 < |\eta| < 2.5$) and 2 layers for larger η ($|\eta| > 2.5$). The spatial precision is achieved by using high-granularity (0.003×0.1, measured in units of $\Delta\eta \times \Delta\phi$) energy measurements in terms of η in the first layer. The second barrel layer has a granularity of 0.025×0.025, while this is coarser in the end-caps (0.1×0.1 for $2.5 < |\eta| < 3.2$). A pre-sampler (PS) is located in front of the barrel calorimeter (for $|\eta| < 1.8$) and is used to correct for the energy lost in the material (ID, cryostats, coil) upstream of the calorimeter. Furthermore, a copper/LAr calorimeter is located in the forward regions ($3.1 < |\eta| < 4.9$) for complete geometrical coverage. The EM calorimeters correspond to between 23 and 39 radiation lengths (X_0) depending on the η region. Figure 3.7 gives a sketch of the granularity of the EM calorimeter.

The hadronic calorimeters are located outside the EM calorimeters. In the barrel region ($|\eta| < 1.7$), tile calorimeters, consisting of steel plates connected to scintillator tiles, have a depth of around 10 nuclear interaction lengths (λ), restricting the leakage of hadronic showers towards the muon system (punch through hadrons). The use of scintillator tiles as active material provides a good energy resolution for hadronic showers, though the granularity is coarse compared with the EM calorimeter (0.1×0.1 for the first two layers, 0.2×0.1 for the third layer). The end-cap hadronic calorimeters are copper/LAr type, with 4 sampling layers and covering the region $1.5 < |\eta| < 3.2$. The forward region ($3.1 < |\eta| < 4.9$) is also covered by the hadronic calorimeter, where a combination of copper and tungsten absorbers is used, along with LAr as the active material. As a whole, the hadronic calorimeter system is at least 10 interaction lengths deep, in the whole range of η. An overall sketch of the calorimeter system is given in fig. 3.8.

3.2.5 Muon Spectrometer

The muon detector is the outer part of ATLAS, extending to radii from 4.25 m to 11 m for the barrel region and distances from 7 m to 23 m from the interaction point for the end-caps. For

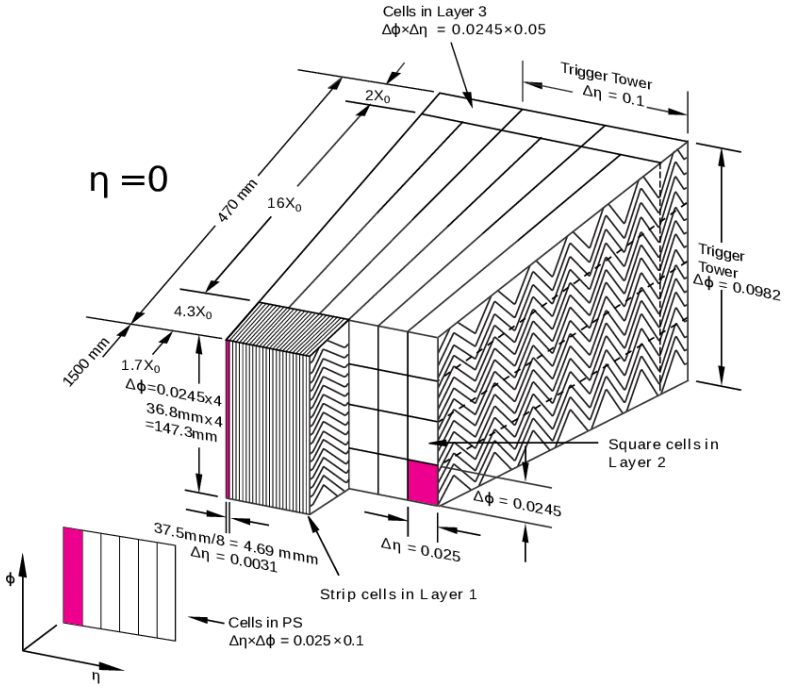


Figure 3.7: Sketch of the lateral and longitudinal segmentation of the ATLAS electromagnetic calorimeter around $\eta = 0$ [72].

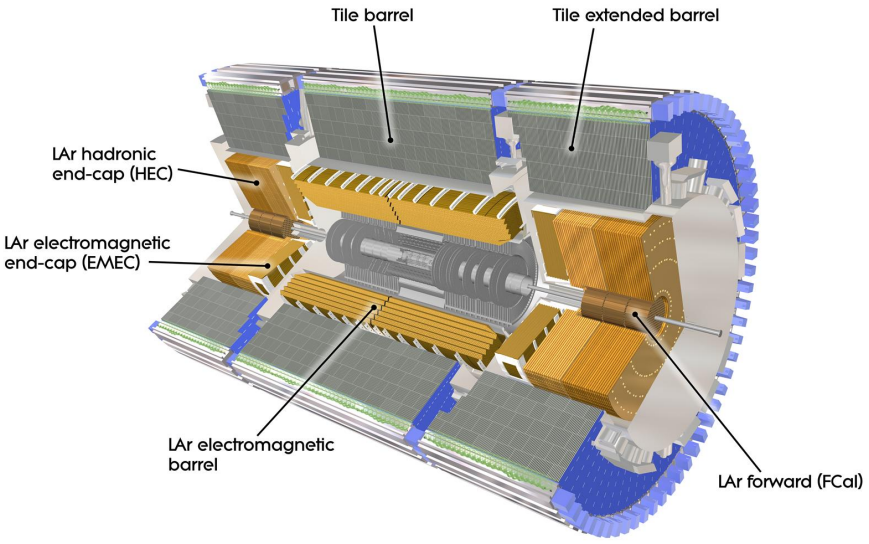


Figure 3.8: Representation of the ATLAS calorimeter system [79].

$|\eta| < 1$, the magnetic bending of muons is provided by the barrel toroid and for $1.4 < |\eta| < 2.7$ by the end-cap toroids. For the intermediate values, corresponding to a transition region, bending is provided by a combination of the end-cap and barrel toroids. In the barrel region tracks are measured by chambers arranged in three cylindrical stations, concentric with the beam axis. In the transition and end-cap regions, the chambers are arranged in 3 disks (or "wheels"), perpendicular to the beam axis. In all pseudorapidity ranges, the precise measurements of the coordinates of the muon trajectory perpendicular to the magnetic field are performed by the **Monitored Drift Tube** (MDT) chambers. However they do not provide measurement of the second coordinate, parallel to the magnetic field. For

large pseudorapidity and for regions closer to the interaction point, the **Cathode Strip Chambers** (CSC) are used, which are multiwire proportional chambers with cathodes segmented into strips. These have fine spatial and time resolution and can cope with the greater particle flux and high radiation background. They also provide information about the second coordinate.

The trigger function (discussed in section 3.3) in the barrel is provided by three stations of **Resistive Plate Chambers** (RPC). They are located on both sides of the middle MDT station, and directly inside the outer MDT station. In the end caps, the trigger is provided by three stations of **Thin Gap Chambers** (TGC) located near the middle MDT station. Both types of trigger chambers give fast signals, with time resolution of a few nanoseconds, which are used for the Level-1 Trigger (see section 3.3) as well as for the identification of the bunch-crossing. They also provide information about the second coordinate, i.e. in a direction approximately parallel to the magnetic field. An overall representation of the muon system is given in fig. 3.9.

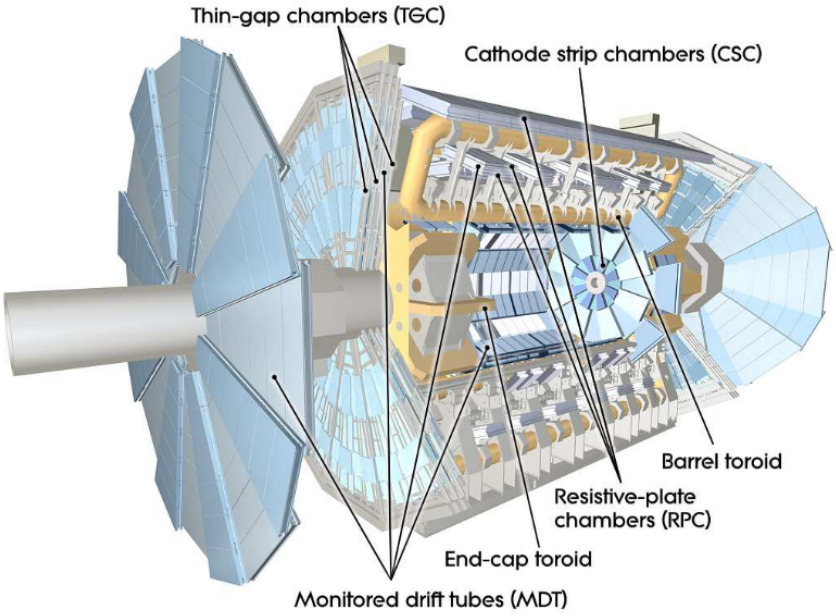


Figure 3.9: Representation of the ATLAS muon spectrometer [72].

3.3 The ATLAS Trigger System

In this section, a more detailed description of the ATLAS trigger system is provided, as it is relevant to the work performed in the context of its upgrade program that is presented in chapter 4. Some of the important upgrades performed on the Level-1 trigger subsystem in preparation for Run 3 are mentioned in section 3.3.2.

3.3.1 Design Principles

The design of the ATLAS trigger and data-acquisition (TDAQ) system [76, 80–84] is based on different levels of online event selection, each level refining the decisions made at the previous level by applying additional selection criteria. Starting from a bunch-crossing frequency of 40 MHz, the rate of selected

events has to be reduced to the order of 1 kHz for permanent storage, while retaining an optimal efficiency for rare physics processes.

The Level-1 (LVL1) trigger selects events based on reduced-granularity information from a subset of detectors. Muons are identified using only the trigger chambers (RPCs and TGCs), by the Level-1 Muon (L1Muon) trigger subsystem. The Level-1 Calorimeter (L1Calo) trigger subsystem is responsible for identifying high-p_{T} electrons and photons, jets, hadronically decaying τ-leptons (by applying dedicated algorithms which use energy isolation criteria) as well as large missing and total transverse energies, based on reduced-granularity information (from **trigger towers** of 0.1×0.1 in $\Delta\eta \times \Delta\phi$, see fig. 3.7) from all the calorimeters. A number of p_{T} thresholds are applied to provide separate trigger information. The missing and total scalar transverse energies are calculated by summing over all trigger towers. In addition, a trigger based on the scalar sum of jet transverse energies is also available. The maximum output event rate of LVL1 triggers, initially limited to ~ 70 kHz, was upgraded to 100 kHz for Run 2.

The LVL1 trigger is also responsible for uniquely identifying the bunch-crossing of interest. With a 25 ns bunch-crossing interval, this task is quite challenging. The physical size of the muon spectrometer implies times-of-flight of muons comparable to the bunch-crossing period, while the pulse shape of the calorimeter signals extends over many bunch crossings. The LVL1 *latency* (the time from the *pp* collision until the trigger decision is available to the next trigger level) has to be minimised and kept below 2.5 μs (2.0 μs has been achieved). "Pipeline" memories are used to store the trigger information during that time. The events selected by LVL1 are read out from the front-end electronics systems of the detectors into readout drivers (RODs) and then into readout buffers (ROBs).

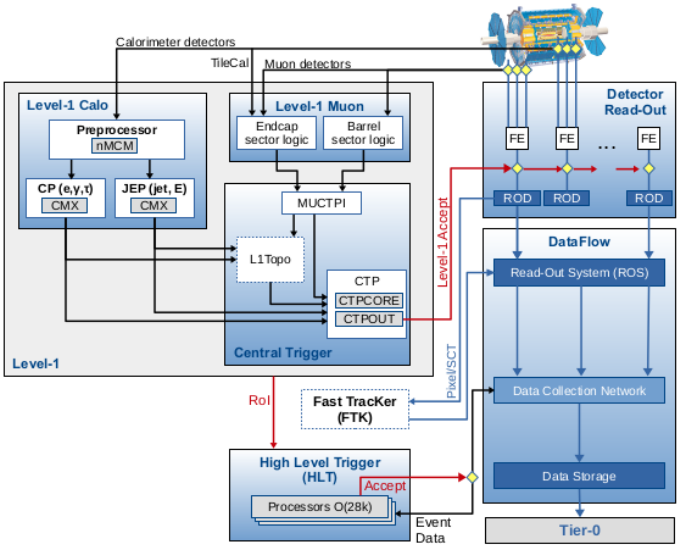
The topological trigger processor (L1Topo) system allows topological selections to be applied in the Level-1 trigger, combining kinematic information from multiple calorimeter and muon trigger objects, such as angular separation, invariant mass requirements, or global event quantities such as the sum of the transverse momenta of all Level-1 jet objects. The central trigger processor (CTP) forms the Level-1 trigger decision based on the information received from L1Topo, L1Calo and L1Muon systems, and distributes the Level-1 accept (L1A) signal and LHC timing signals to the detector readout subsystems, via the Timing, Trigger and Control (TTC) system.

The higher-level triggering (separated into a level-2 (LVL2) and an Event Filter (EF) during Run 1, while merged into a single High-Level Trigger (HLT) farm [85] for Run 2 onwards) is made possible by use of "region-of-interest" (ROI) information provided by the LVL1 trigger, which comprises information on the position (η and ϕ) and p_{T} of candidate objects for physics (high-p_{T} muons, electrons, photons, hadrons, τs and jets), and energy sums (missing-E_{T}, vector and scalar E_{T} values). The ROI mechanism allows for the reduction of the data processed, since typically only data corresponding to limited regions centred on the objects indicated by the LVL1 trigger are needed; in such cases the full precision data and granularity are accessed. The event rate of LVL2 was ~ 3 kHz, the latency being in the range of 1-10 ms (depending on the events processed).

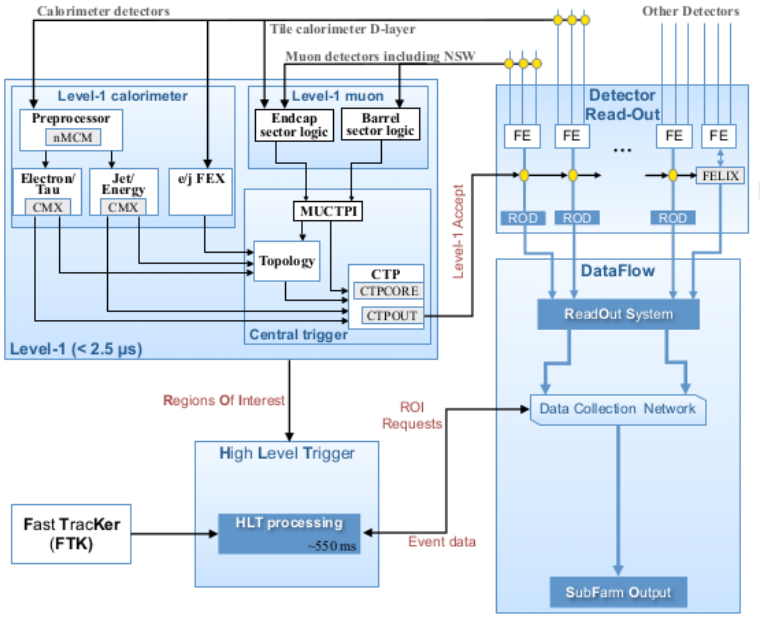
The use of information from the precision muon chambers and the ID allows improvement of the resolution of the p_{T} of muons, and information from the calorimeters around the muon candidates allows the application of isolation requirements. In this way, as well as increasing the muon p_{T} threshold if necessary, muon candidates which are not real muons can be rejected. For isolated electrons, the

rejection power comes from using the full-granularity calorimeter information and requiring a matching high-p_{T} charged track in the ID; the TRT provides additional rejection power. For photons, less rejection power can be achieved than for electrons, since the ID cannot always be used as a veto given the relatively high probability for photon conversion in the ID material.

The last part of the online selection relies on the use of offline algorithms and methods, the most up-to-date calibration and alignment information and the magnetic field map. The final selection of physics events is written to mass storage for subsequent full offline analysis. The final trigger output rate (after HLT) could be raised to ~ 1 kHz by upgrades of the system prior to Run 2, corresponding to a data rate of around 1 GB/s (assuming event sizes of around 1 MB). A sketch of the Trigger System for Run 2 is given in fig. 3.10a.



(a)



(b)

Figure 3.10: Schematic overview of the Trigger and DAQ system (a) for Run 2, (b) for Run 3. The FTK [86] was not realised. Taken from [83, 87].

3.3.2 Phase-1 upgrade of L1Calo trigger

During Long Shutdown 2, further upgrades [87] were made to the TDAQ system, to prepare for the ATLAS Run 3 data-taking period [88] (a sketch is given in fig. 3.10b). In this section we focus on changes to the L1Calo subsystem and the upstream calorimeter electronics [89].

Three new subsystems were added to L1Calo: the electromagnetic, jet and global feature extractors (eFEX, jFEX and gFEX). To achieve the increased discriminatory power necessary to handle the LHC luminosities planned beyond LS2, these process calorimeter data at a finer level of granularity than the pre-LS2 L1Calo system. Initially at least, they augment the existing L1Calo electronics, operating in parallel with the Cluster Processor (CP) and Jet Energy Processor (JEP) (Run 1 & Run 2 legacy systems).

Prior to LS2, the LAr and tile calorimeter electronics provided analogue data, which were digitised in the Pre-Processor Modules (PPMs). During Run 3 the LAr calorimeter data are digitised by a new LAr Trigger Digitiser Board (LTDB) and provided to L1Calo by the Digital Processing System (DPS). For each tower of 0.1×0.1, eFEX receives up to ten samples, derived from longitudinal segments and transverse sums of groups of calorimeter cells, each sum forming a SuperCell (fig. 3.11). Analogue tile calorimeter data are digitised in an upgraded pre-processor and transmitted via the new Tile Rear Extension (TREX) pre-processor extension module [90].

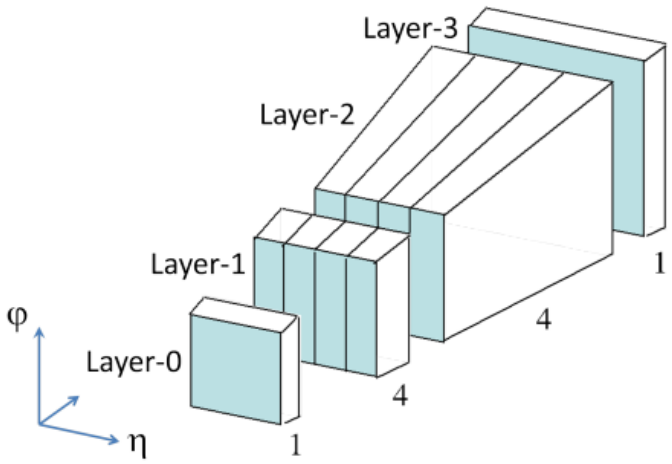


Figure 3.11: The trigger granularity from each trigger tower after the Phase-1 upgrade of the LAr Calorimeter electronics. Ten E_{T} values are provided from "1-4-4-1" (per layer) longitudinal/transverse samples, each forming a SuperCell. The hadronic calorimeter in the same trigger tower (not shown) provides one more SuperCell. Taken from [87].

The function of the eFEX module is to identify isolated energy deposits indicative of electrons, photons and τ leptons. The jFEX subsystem identifies energetic jet candidates and also performs the large-area τ, $E_{\mathrm{T}}^{\mathrm{miss}}$ and total E_{T} trigger algorithms. The gFEX subsystem is responsible for the identification of large radius jets, additional $E_{\mathrm{T}}^{\mathrm{miss}}$ computations and other jet-specific calculations. The results of the processing done by all the FEX modules give Trigger Objects (TOBs, comprising the object type, energy measured and η, ϕ coordinates), which are transmitted to L1Topo over optical fibres. In L1Topo, results from all the FEX subsystems are used as input to topological algorithms and the

results are transmitted to the CTP. The FEX modules comply with the Advanced Telecommunications Computing Architecture standard (ATCA) [91], and are based on field-programmable gate array (FPGA) technology. The eFEX comprises two shelves of 12 eFEX modules each and the jFEX comprises a single shelf of six jFEX modules; one gFEX module is sufficient for the implementation of the gFEX processing system.

On receipt of an L1A signal, the FEX modules provide ROI and readout data to the HLT and DAQ system, respectively. Each FEX module outputs these data to the shelf backplane. Two *Hub* modules in each shelf aggregate the data and implement the required ROD functionality on daughter boards. Additionally, the Hub modules provide nodes on the TTC, control and monitoring networks. The hardware of the Hub modules is common to the eFEX, jFEX and gFEX subsystems, but the readout firmware differs in order to handle the different data.

Figure 3.12 gives a schematic view of the L1Calo Trigger subsystem architecture (as a "block diagram") for Run 3.

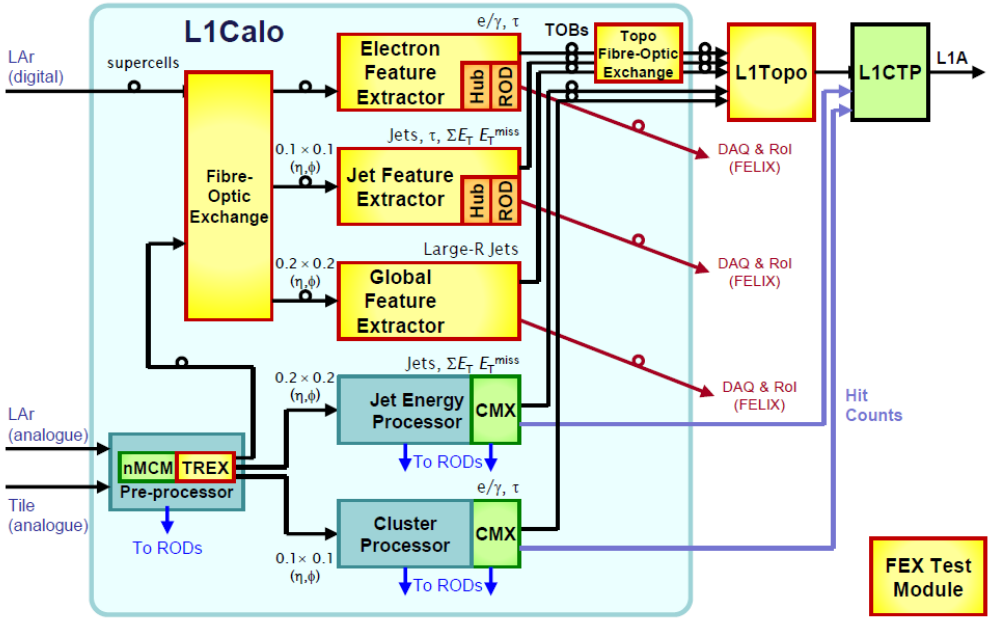


Figure 3.12: L1Calo system block diagram after completion of the Phase-1 upgrade; the eFEX, jFEX and gFEX are among the new processing hardware modules introduced during the Phase-1 upgrade. In green are the Run 2 additions, while in yellow/orange are the Run 3 additions (L1Topo has been redesigned for Run 3, hence the yellow colour). From [89].

Chapter 4

Tests of FEX Test Module and of eFEX firmware

4.1 Introduction

In the previous chapter it was mentioned that in the context of the ATLAS Phase-1 Upgrade of the Level-1 Calorimeter trigger system, vital processing of digital data coming from the ATLAS calorimeters will be performed by new, FPGA-based, modules called eFEX and jFEX. The Feature extractor (FEX) modules receive and transmit many multi-gigabit data streams over fibre optics and the validation of these data paths requires a complementary test module to transmit and receive this data.

In this chapter we describe the testing procedure for the Feature extractor Test Module (FTM), with emphasis on the verification of fundamental connections and monitoring of vital values via Python scripts, after explaining its role and layout. We also provide a description of the testing of the mapping implementation in the eFEX firmware. The mapping of incoming data to the position of energy deposits in the calorimeters is vital for the algorithm eFEX uses to identify and trigger physics objects.

These tests were performed as part of the commissioning of the FTM and eFEX modules, the latter now playing a vital role in the ATLAS Run 3 triggering system, and constituted the *qualification task* of the writer, which is an essential stage for a member of ATLAS to be registered in the ATLAS author list, by having contributed to some aspect of the detector technology or algorithms.

4.2 FTM testing

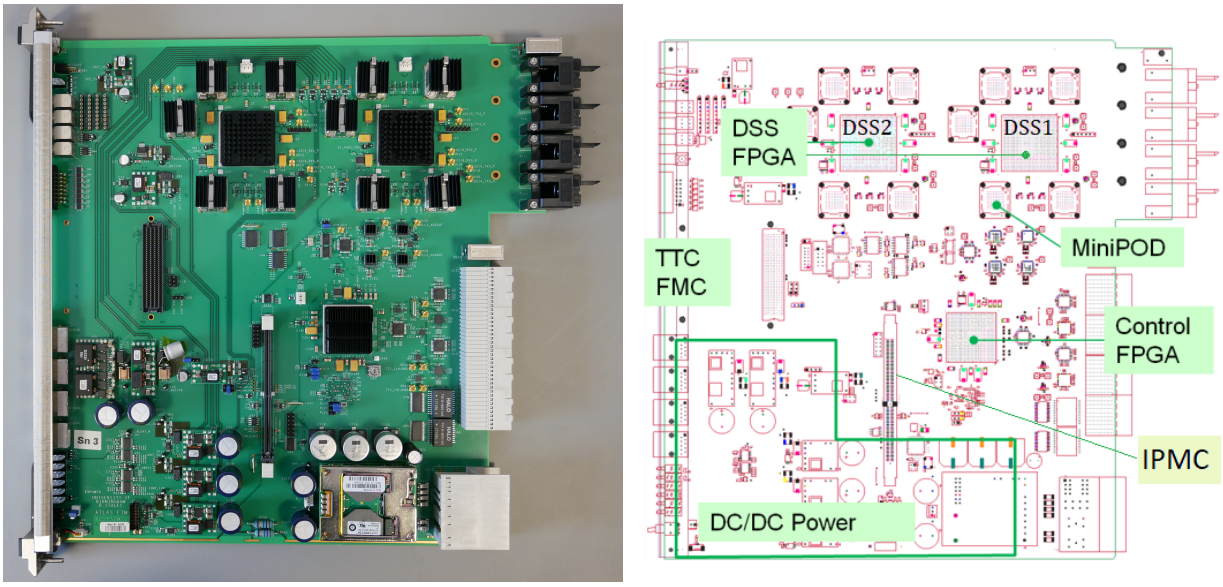
4.2.1 FTM role and design principles

The FEX Test Module's main purpose is to generate test data patterns (that imitate the real data that will be coming from the calorimeters) for transmission to the FEX modules and record the results from their outputs. This test module was built before the FEXs, hence it has also been used to confirm the high-speed signalling and technology choices intended for the FEX modules. Building this test module has allowed the FEX designers to gain early experience with the Xilinx Virtex-7 family [92] of FPGAs, and more specifically with the Multi-Gigabit Transceivers (MGTs) integrated inside them. The MGTs are integrated circuits used for ultra-fast (up to 13.1 Gbps in the case of FEXs) serial data-transmission to optical modules. Another aim for designing the FTM was to gain experience with the Advanced Telecommunications Computing Architecture (ATCA) standard. ATCA is of interest to the physics community due to its high availability, scalability, flexibility and massive input/output capability. In

brief, it is a standard which defines the module sizes as well as fundamental external supplies and connections.

The majority of the test module functions are provided by two large Xilinx Virtex-7 FPGA devices, used for data-transmission and reception to/from the FEXs, aptly named Data Source and Sink (DSS) FPGAs. All MGT transmitters within each DSS FPGA are routed to four MiniPOD (electrical to optical converter) sockets (with the code name Tx, "T" standing for "transmitter"). For the data sink, only one MiniPOD receiver (Rx) is connected to each DSS FPGA. These MiniPODs receive the eFEX outputs, as well as being used for the FTM self-testing.

For an overall layout of the FTM board see fig. 4.1. A smaller Xilinx Kintex-7 Control FPGA (C-FPGA) provides control and timing functions. The FTM can use the detector's TTC clock signal, or provide a local clock and real-time control signals. The Control FPGA also provides the interface to the external network (users), which is used for the functional control of the module and environmental monitoring.



(a) (b)

Figure 4.1: A photo of a FEX Test Module board (a) and a graphical representation of its layout (b).

The FTM accommodates a number of supplementary devices/components (Power Supply Unit, MiniPODs, ADCs etc.). Some of them will be described in the testing section (section 4.2.2). Each device is characterised by its register map: essentially memory blocks storing 32-bit values specific to the device's function [93, 94]. Various types of connections (or buses [95, 96], such as I2C [97–99] and SPI buses) are used for the communication (data transmission) between the Control FPGA and these devices.

In order to access the values stored in registers, an IPbus [100] connection is implemented between the user and the Control FPGA. IPbus connection of the C-FPGA to user PCs is made via an Ethernet cable. A script repository has been created by the Birmingham team, containing various Python classes defined for the IPBus communication to all devices (essentially to access their registers) on the board.

The FTM has a DIMM slot to accommodate an Intelligent Platform Management Controller Card/device (IPMC), an essential Fast Local Control utility (for the monitoring of power, temperatures

and interconnections). The IPMC connects to two I2C buses. In the absence of an IPMC, an I2C Bridge integrated circuit (IC) allows the IPMC sensor and management bus to be accessed by the control FPGA for test purposes.

A representation of the logic of interconnections between the various devices on the FTM printed circuit board (PCB) is given in fig. 4.2.

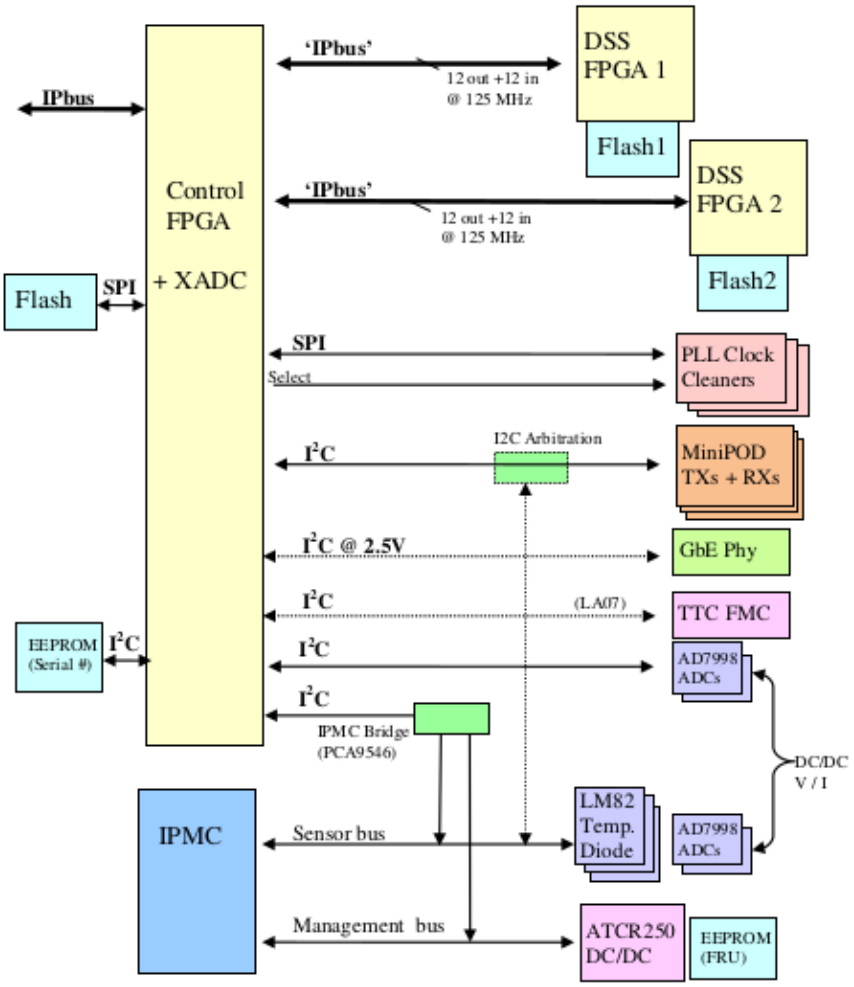


Figure 4.2: Interconnections of the various devices present on the FTM board [94].

4.2.2 Testing procedure

Upon reception of an FTM board from the manufacturer, a series of tests had to be performed for its commissioning. Generally, the first test of the FTM board begins with a JTAG [95] scan for a "direct" hardware check of the board's FPGAs, such as loading essential configuration files into the FPGAs or their FLASH memory, and booting the firmware design. Other initial tests may also be performed, such as the bit error ratio (BER) measurement of the optical transmission links.

After that, we proceed to two testing chains via software (Python scripts). The first testing chain is more fundamental and checks all I2C connections, reads power supply values and temperatures, thus verifying connections, and checks that the values are within limits. An end-user script has been produced by the writer, that contains all relevant steps.

The second testing chain (not described here) is relevant to testing the MGT source and sink RAMs by sending Test Data. It also checks the SPI buses to the FLASH memory of all FPGAs.

In the following section we will describe the first testing chain. Four production boards (with serial numbers 6, 7, 8, 9) have been tested.

4.2.2.1 I2C address space scan

The first step is to scan the I2C address space. There are four I2C interfaces (for MiniPODs, Bridge, ADCs and EEPROM; for the latter see section 4.2.2.7) that have to be scanned separately, as each one may contain a number of devices. The devices associated with each interface have 7-bit addresses, so for each interface we perform a scan of 128, i.e. 2^7, addresses, by implementing an I2C transaction up to an initial acknowledge signal that states that the device is reachable (thus there is a proper physical connection, and the protocol implementation by the code is verified). The addresses found are then compared to the expected addresses, as defined in xml files [93]. The scan was successful for all production boards tested, with the responding addresses matching the expected ones.

4.2.2.2 Power Supply Unit reading

The incoming −48V supply from the backplane connector feeds into a DC/DC converter which has the following power outputs:

- 3.3 V Management power for CERN IPMC + EEPROM (15 W of power is available from this supply rail, the “control” power supply)
- 12 V Payload power. Feeds DC/DC converters for FPGAs, Optics etc. (250 W of power is available from this supply rail, the “main power supply”). The different supply voltages needed by various components are derived from the 12V supply using further DC/DC converters.

We access the values stored in the registers of the Power Supply Unit via the bridge connection to the IPMC management bus. The code implements I2C transactions to access the values of these registers. The values read out in this way from testing FTM#9 are given in table 4.1.

Power Supply Unit			
Supply	Value	Low Limit	High Limit
48V Voltage (V)	46.64	37.6	72
48V Current (A)	1.77	0	5.9
12V Voltage (V)	11.84	10.82	13.22
12V Current (A)	5.09	0	18.72
3.3V Voltage (V)	3.32	3.13	3.47
Temperature (°C)	32	–	116

Table 4.1: Voltage, current and temperature measurements for the Power Supply Unit and for FTM#9.

The values measured are well within limits for all the boards.

4.2.2.3 ADC reading

The FTM Board accommodates two ADC modules that are connected to the C-FPGA via I2C, and two ADCs connected to the C-FPGA via the Bridge. The latter are connected to the IPMC card via the

I2C sensor bus, when this card is inserted into the board. These ADCs are responsible for measuring various supply voltages and currents intended for the FPGAs and the MiniPODS, as well as other ICs on board. Measurements from a test on FTM#9 are given in table 4.2.

	Current (A)			Voltage (V)			
	Measurement via			Measurement via			Nominal
Supply ID	I2C direct	I2C bridge	FSD value	I2C direct	I2C bridge	FSD value	value
VCC3V3	9.67	9.93	25.00	3.31	3.31	4.00	3.30
VCC2V5	4.03	4.25	25.00	2.50	2.50	4.00	2.50
VCC1V8	2.52	2.64	25.00	1.80	1.80	2.50	1.80
VADJ	0.80	0.80	5.00	2.50	2.52	4.00	2.50
MGTAVTT	3.80	3.76	12.50	1.25	1.25	2.50	1.20
MGTVCCAUX	0.48	0.49	2.50	1.79	1.80	2.50	1.80
MGTAVCC	21.78	21.87	50.00	1.02	1.02	2.50	1.00
VCCINT	15.53	15.59	50.00	1.00	1.00	2.50	1.00

Table 4.2: Voltage and current measurements by the ADC units, via the two separate I2C connections, for FTM#9. The VCC3V3, VCC2V5 supplies are used by the C-FPGA, MiniPODs and other ICs on the board, while VCC1V8, MGTAVTT, MGTVCCAUX, MGTAVCC, VCCINT are used by the FPGAs (all three of them). Full Scale Deflection values (FSD) are given, as the highest acceptable limits.

VADJ is an auxiliary voltage, not currently used. The current taken from each supply is converted to a voltage for measurement. Values measured from I2C directly and via the bridge should be compatible (within around 5%, to account for noise), and this has been verified. Full Scale Deflection (FSD) values are the maximum values that can be measured by the ADC sensors, and can be regarded as the highest acceptable limits, though in practice lower limits must be satisfied. In the case of VCCINT, MGTAVCC, MGTAVTT supply current measurements, a replacement of the current sensor amplifier had to be made because the device tested showed large errors at low currents. In fig. 4.3 we present the current measurements for various supply rails, and for all four FTM boards tested.

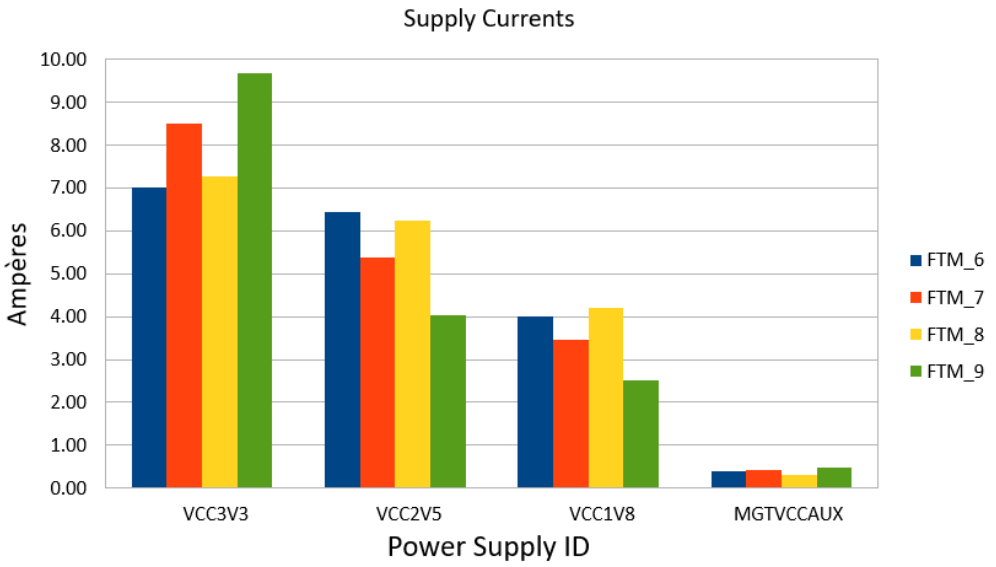


Figure 4.3: Current measurement (reading ADC registers via direct I2C connection) for various supplies intended for the FPGAs and other components on board.

The currents show some differences when testing different boards. This variation is acceptable, and can be attributed to manufacturing differences and properties of the transistors used. Current

measurements are subject to fluctuations (around 3-6%), but this does not change the picture presented in fig. 4.3. The voltages on the other hand are very stable (1% variation or less).

4.2.2.4 XADC reading

Each FPGA has one XADC (Xilinx ADC) module integrated onto it [101–103], which is a very efficient dual ADC device responsible for monitoring FPGA power supplies (voltages only) and temperature. It supports up to 17 flexible and user-configurable analogue inputs multiplexed to the ADC. The most recent measurement results (together with maximum and minimum readings) are stored in dedicated registers.

Reading the XADCs is important, because as well as providing additional monitoring, it can validate other measurements, e.g. by comparison to the ADCs or externally to FPGA temperature sensors. The test also verifies IPbus access to FPGA registers. The XADC readings for FTM#8 are given in table 4.3. Similar values are obtained from all boards tested.

	C-FPGA	DSS1	DSS2
Temperature (°C)	50.8	64.36	61.75
VCC_Aux (V)	1.81	1.8	1.8
VCC_Int (V)	0.98	0.96	0.97
VCC_Ram (V)	0.98	0.96	0.97

Table 4.3: Reading XADC registers for each FPGA directly via IPBus, from a test of FTM#8.

4.2.2.5 FPGA Temperature Sensor reading

Three LM82 Digital Diode Temperature Sensors [104], one for each FPGA, are placed on the FTM board. Each one provides two measurements, one on the FPGA and one in its vicinity (indicative of a "board" temperature). The measurements from a test of FTM#8 appear in table 4.4.

	Temperature (°C) on	
	FPGA	Board
Control FPGA	46	43
DSS1 FPGA	61	45
DSS2 FPGA	59	38

Table 4.4: Reading FPGA temperatures via I2C connection through bridge to IPMC Sensor Bus, for FTM#8.

A comparison with table 4.3 shows that the XADC temperature values are higher, as expected since the XADC modules are integrated inside the FPGA. Fluctuations of the temperature values are of the order of ±1°C. Temperature limits (for the XADC which provides the highest possible value) are set around 80°C; values higher than that should cause concern. Effective cooling of the FPGAs as well as ventilation of the crate is crucial.

4.2.2.6 MiniPODs monitoring

Each MiniPOD can be connected to a 12-way fibre ribbon, in cases of data transfer to a FEX board. We need to verify the I2C access to the register map of the MiniPODs, read the values stored and

verify that they are as expected or within accepted limits. Dedicated optical sensors measure the light intensity on each of the 12 channels (where single optical fibres will be connected) for each MiniPOD. Measurements of values relevant to the MiniPODs are presented in table 4.5.

	DSS1					DSS2				
	Tx0	Tx1	Tx2	Tx3	Rx	Tx0	Tx1	Tx2	Tx3	Rx
Temperature (°C)	59	52	65	56	49	54	60	51	56	43
Voltage3.3 (V)	3.28	3.25	3.23	3.26	3.32	3.27	3.26	3.28	3.26	3.29
Voltage2.5 (V)	2.45	2.44	2.43	2.44	2.47	2.44	2.44	2.44	2.44	2.46
I_{opt} (mW) Ch0	0.79	0.95	1.08	0.77	0.00	1.02	1.04	0.85	0.99	0.00
I_{opt} (mW) Ch1	0.83	0.99	1.05	0.90	0.00	0.87	1.07	0.82	1.17	0.00
I_{opt} (mW) Ch2	0.82	0.89	1.08	0.77	0.00	0.96	1.00	0.90	1.04	0.00
I_{opt} (mW) Ch3	0.81	0.96	0.97	0.88	0.00	0.91	0.97	0.81	1.11	0.00
I_{opt} (mW) Ch4	0.79	0.85	1.02	0.82	0.00	1.00	0.94	0.98	1.00	0.00
I_{opt} (mW) Ch5	0.83	0.97	0.90	0.90	0.00	0.90	0.97	0.80	1.09	0.00
I_{opt} (mW) Ch6	0.84	0.87	1.02	0.81	0.00	0.98	0.95	0.89	1.01	0.00
I_{opt} (mW) Ch7	0.75	0.99	0.97	0.91	0.00	0.92	1.03	0.80	1.09	0.00
I_{opt} (mW) Ch8	0.76	0.89	0.98	0.81	0.00	1.00	0.89	0.93	1.03	0.00
I_{opt} (mW) Ch9	0.74	0.95	0.92	0.90	0.00	1.03	0.91	0.80	1.03	0.00
I_{opt} (mW) Ch10	0.68	0.92	0.98	0.81	0.00	1.04	0.91	0.94	1.00	0.00
I_{opt} (mW) Ch11	0.68	0.91	0.99	1.00	0.00	1.00	1.00	0.87	1.18	0.00

Table 4.5: MiniPOD values read via direct I2C Bus connection from testing FTM#9.

The light output (I_{opt}) is subject to a small fluctuation of $\pm 1\%$. Since the testing was performed without optical fibres connected, nor in the process of data transfer, the values presented for the optical intensity could be regarded as default values. As in all previous steps, testing has been repeated for four production FTM boards. In fig. 4.4 and fig. 4.5 the comparison of each channel's light intensity for each Tx MiniPOD between different boards is shown.

Small variations around a value of 1 mW are observed, but can be justified by small manufacturing differences in electro-optical semiconductor devices used in the PCB production. On the other hand, the light intensities are consistent per board tested. The tests verified the connections to the MiniPODs. The values obtained by accessing the MiniPODs' registers are as expected by the designers.

4.2.2.7 EEPROM reading

As a last step, the Python script reads the FTM board's serial number, by performing an I2C transaction to read the register of the FTM's Electrically Erasable Programmable ROM (EEPROM). The serial number is used to identify uniquely each board and assign IP and MAC addresses to the FTM for remote access. A single number is output; in our case we correctly read the serial numbers 6, 7, 8, 9 of the boards tested.

4.2.3 Discussion

The FTM testing was successful. It verified the connections, and the values were measured to be within accepted limits. Some differences in the MiniPODs' light output intensities for the production boards tested, can be expected and related to manufacturing differences.

With a scan like this, as well as when trying to retrieve values from registers as described in the previous section, we were able to identify a hardware problem. This was a missing I2C connection on

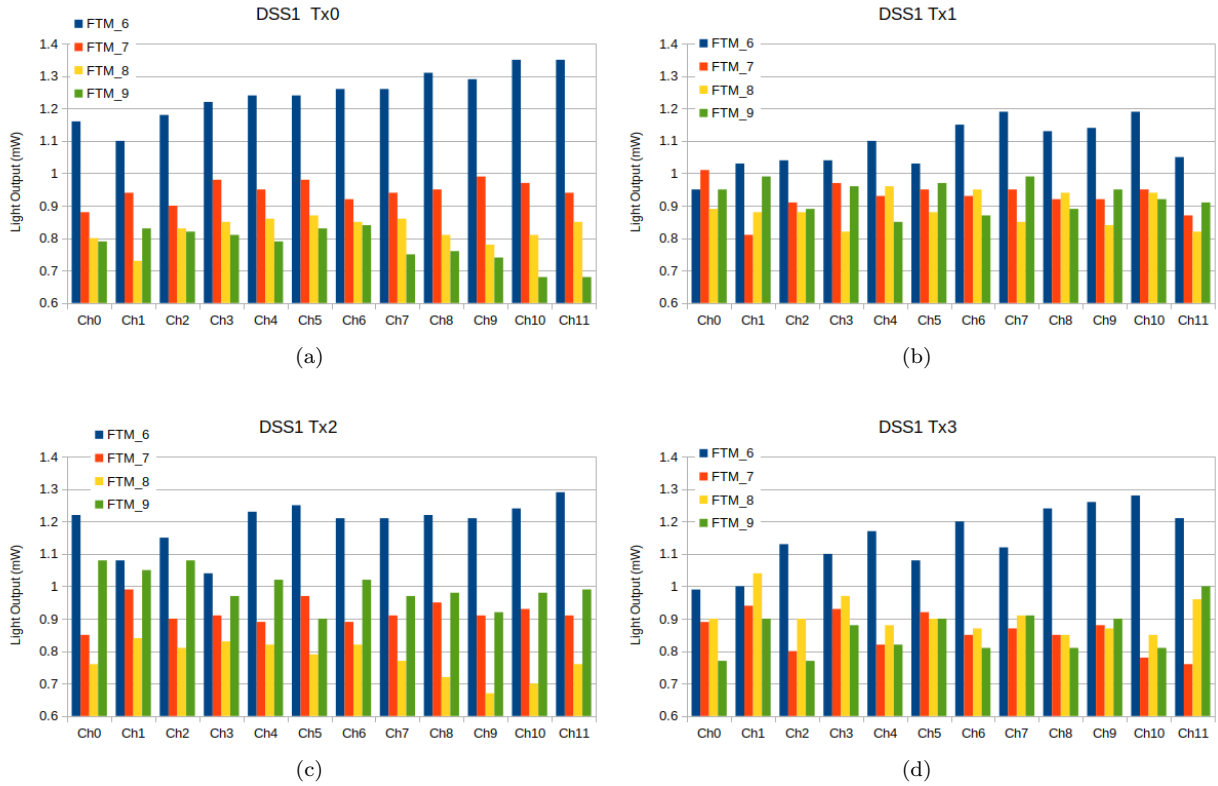


Figure 4.4: Light output intensity in mW for each optical fibre, or channel, in the four Transmitter MiniPODs (named Tx0, Tx1...) of the DSS1 FPGA, for the four different FTM boards tested.

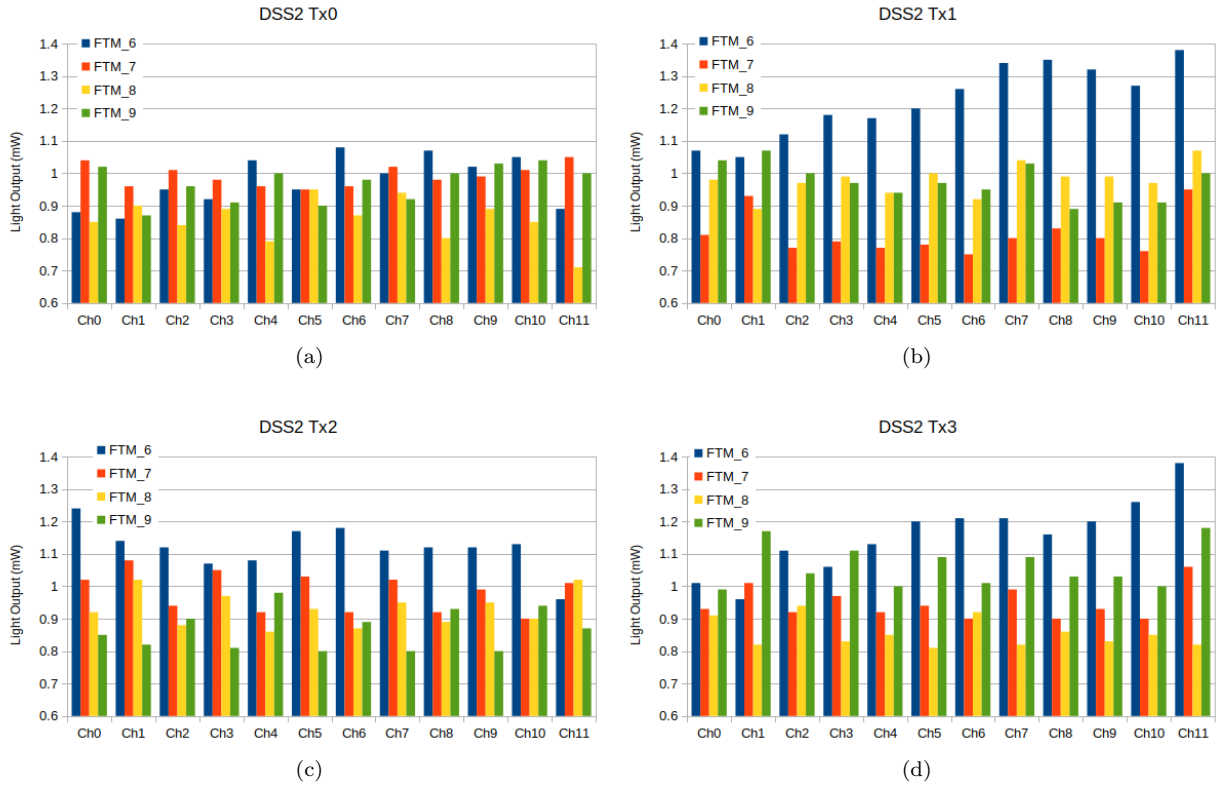


Figure 4.5: Light output intensity in mW for each optical fibre, or channel, in the four Transmitter MiniPODs (named Tx0, Tx1...) of the DSS2 FPGA, for the four different FTM boards tested.

the bridge component, which was subsequently fixed. This has highlighted the importance of the testing via software/Python scripts.

4.3 Tests of the eFEX mapping

4.3.1 eFEX role and design principles

The electron Feature extractor (eFEX) [105, 106] has architectural similarities to the FTM. The data processing is provided by four Xilinx Virtex-7 *processor* FPGAs (pFPGAs), while a *control* FPGA (cFPGA) of the same family (but of a previous type) is used for communication with the user, and as an interface to most devices on the board (fig. 4.6).

The energies measured in the SuperCells (depicted, for the detector's barrel region, in fig. 3.11) of the calorimeters are routed as digitised signals prepared by the LATOME [107, 108] modules, which are part of the DPS and act as a front-end to the LAr calorimeter, through optical fibres to the eFEX modules for processing. Each optical fibre transmits data with a speed of 11.2 Gbps. The optical fibres are connected to MiniPODs, which convert optical signals to electrical signals and feed the MGTs of the pFPGAs. The eFEX *algorithm* is responsible for distinguishing electrons and photons from the dominant jet background, by use of dedicated shower-shape and isolation variables defined from the energy of the SuperCells.

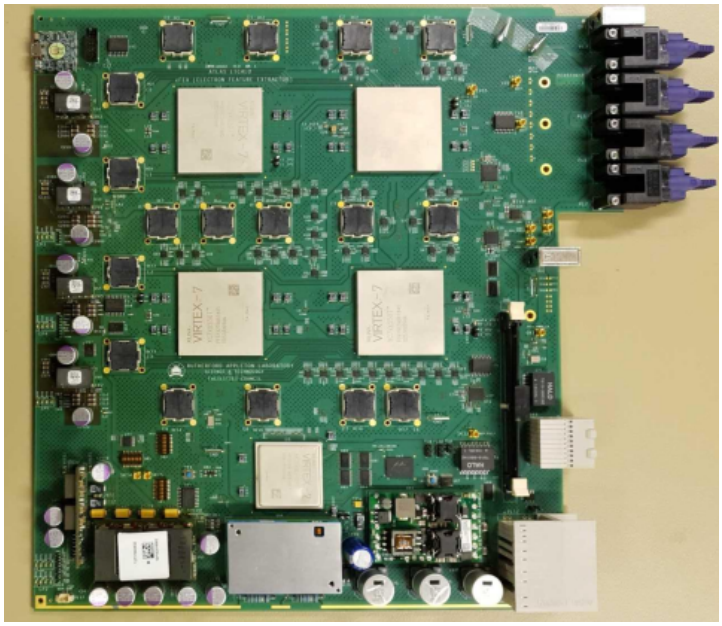


Figure 4.6: A photo of an eFEX prototype board. The processor FPGAs can be seen as large silver squares and the cFPGA is the smaller silver square below them. The pFPGAs are surrounded by small black squares which are the MiniPODs, where the optical fibres are connected to route the data to the MGTs that are integrated into the FPGAs.

Twenty-four (24) eFEX modules are used in total, and they account for the calorimeter pseudorapidity coverage of $|\eta| \leq 2.5$. Each eFEX module is responsible for producing candidate objects from a *core* area of 1.7×0.8 (in terms of $\Delta\eta \times \Delta\phi$), using an augmented region that includes also neighbouring cells of the calorimeter, which are necessary for the trigger algorithm (fig. 4.7).

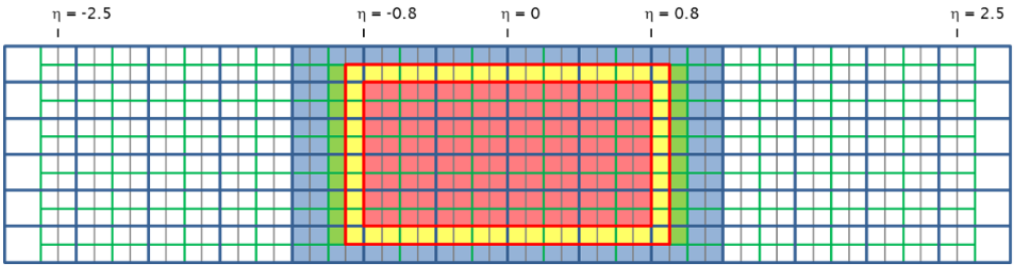


Figure 4.7: eFEX module processing window in $\eta-\phi$ (with ϕ in the vertical axis). In red is the core area for which trigger candidates are formed. In yellow is the neighbouring area examined by the algorithms. In green and blue are extra areas (from LAr+Tile and Tile Calorimeter respectively) carried within fibres but not used by the algorithms [87].

4.3.2 Tests of the mapping

The digitised signals that an eFEX module receives are formatted into 10-bit data "words": 10 words from the electromagnetic and 1 word from the hadronic calorimeter per trigger tower. Each pFPGA of an eFEX module contains 64 MGTs responsible for receiving the energies formatted into these data words. The data words are further arranged into seven 32-bit numbers (or equivalently into seven 8-digit hexadecimal numbers) according to the scheme shown in table 4.6. Each MGT carries up to 20 data words, corresponding to the information from 2 trigger towers, for each bunch crossing (BC).

W #	b 31	b 30	b 29	b 28	b 27	b 26	b 25	b 24	b 23	b 22	b 21	b 20	b 19	b 18	b 17	b 16	b 15	b 14	b 13	b 12	b 11	b 10	b 9	b 8	b 7	b 6	b 5	b 4	b 3	b 2	b 1	b 0
0	BCID(4..3)		DATA1(9..0)										DATA0(9..0)										BCID(6..5)		K28.5 / QUALITY(7...0)							
1	D19(9:8)		DATA4(9..0)										DATA3(9..0)										DATA2(9..0)									
2	D19(7:6)		DATA7(9..0)										DATA6(9..0)										DATA5(9..0)									
3	D19(5:4)		DATA10(9..0)										DATA9(9..0)										DATA8(9..0)									
4	D19(3:2)		DATA13(9..0)										DATA12(9..0)										DATA11(9..0)									
5	D19(1:0)		DATA16(9..0)										DATA15(9..0)										DATA14(9..0)									
6	CRC(8..0)									BCID(2..0)			DATA18(9..0)										DATA17(9..0)									

Table 4.6: Format of the incoming data words handled by an MGT of an eFEX module. Twenty words are formatted into seven 32-bit words. A cyclic redundancy check (CRC) and a bunch crossing "quality" are included.

The way the eFEX firmware interprets the input energies, in terms of position in $\eta-\phi$ in the calorimeter, is called *mapping*. These energy deposits and their position (Trigger Tower + SuperCell) are then the input to the eFEX algorithm.

The mapping test starts with the generation (via Python scripts) of a series of input files, in order to fully populate the available MGTs of an eFEX FPGA with energy values. A representation of a pFPGA, holding 64 MGTs each accepting 20 data words, is given in table 4.7.

We may choose to assign different values of energy to populate the table 4.7 with a unique value (so that we can recognise them in the output, after the firmware simulation has been run). We also have the freedom to generate multiple BCs, and fill only one MGT at a time with energy values. The latter path has been chosen: 64 BCs are generated at a time, but only one MGT is filled for each BC.

After defining the input files, by arranging the incoming energies (i.e. the data-words) according to table 4.6, the second step of the mapping test consists of running the eFEX firmware simulation in the Xilinx Vivado [109] framework. The tests were performed by simulating the eFEX FPGAs with the

MGT \ word	0	1	2	···	19
0					
1					
2					
⋮					
63					

Table 4.7: Table representing the arrangement of the data words in the available MGTs per pFPGA of an eFEX module. An input pattern is generated by assigning user defined values to the separate cells shown. Since this can be done per bunch crossing, we are free to generate multiple BCs and fill one MGT each time.

ModelSim/QuestaSim package [110, 111]. The simulation produces an output file, which contains blocks of 60 lines, each line formatted as in table 4.8, corresponding to 60 trigger towers processed by a pFPGA per bunch crossing.

0	1111	2222	3	H

Table 4.8: Representation of the format of a line of the output file of the firmware simulation. Each number represents a layer in the EM Calorimeter, "H" being the hadronic calorimeter; for layers 1 and 2 we get 4 SuperCells. Each number represents a 3-digit hexadecimal number, the energy word found. 60 such lines correspond to a single bunch crossing, and account for the 60 trigger towers that a pFPGA can process.

The output is decoded by a separate Python script, given the input pattern that we provided. After decoding, we obtain a file which contains the mapping of the energies to the location of the detector, as illustrated in table 4.9.

In summary, the test consists of generating user-defined input patterns of the incoming data words, formatted as in table 4.6, by use of dedicated Python scripts. After running a simulation of the eFEX firmware, we obtain a single output file formatted according to table 4.8. The latter is decoded by a separate Python script in order to finally have the mapping in a structure that is readable, as in table 4.9.

The mapping testing was repeated for all 4 pFPGA modules and for each eFEX module, by changing the eFEX code number in the firmware. In most cases a unique mapping of the energies to a location in the detector was validated, while in the cases where duplicate mapping was found, a further resolution was followed up in the firmware, by the firmware designer. As a result, the validation of the mapping implementation in the eFEX firmware was achieved.

4.4 Conclusion

In this chapter the contribution of the writer to the Phase-1 Upgrade program of the ATLAS detector has been presented. Firstly, a Python script which contains various steps for checking the connections on an FTM board and the validity of the readings from the various devices on the board, has been developed. The tests performed by the use of the script, have been part of the successful commissioning of the FTM boards.

MGT	WORD	η	ϕ	Layer	SuperCell
12	0	2	4	0	
12	1	2	4	1	4
12	2	2	4	1	3
12	3	2	4	1	2
12	4	2	4	1	1
12	5	2	4	2	4
12	6	2	4	2	3
12	7	2	4	2	2
12	8	2	4	2	1
12	9	2	4	3	
12	10	1	4	0	
12	11	1	4	1	4
12	12	1	4	1	3
12	13	1	4	1	2
12	14	1	4	1	1
12	15	1	4	2	4
12	16	1	4	2	3
12	17	1	4	2	2
12	18	1	4	2	1
12	19	1	4	3	

Table 4.9: Table showing the mapping of each word for a single MGT. One MGT (#12 of pFPGA#1) of an eFEX module covering a central region in η (eFEX code #00) is shown for illustration purposes. Values of $\eta - \phi$ are given in units of 0.1, in the region covered by the eFEX module under study. The total number of MGTs is 64 per pFPGA. Only part of these MGTs carry data words: some MGTs do not receive data words ("inactive"), while some MGTs receive fewer than 20 data words (full calorimeter coverage can be achieved by partial usage of the available MGTs). Such cases, where no mapping is available, are flagged appropriately by the output decoder.

Secondly, two separate Python scripts have been developed in order to test the interpretation of the digitised incoming energies as localised energy deposits in the calorimeters (mapping) as implemented in the eFEX firmware. These tests have been part of the successful eFEX commissioning.

The FTM boards have been repeatedly used to feed simulated data to the FEX modules, and played an instrumental role in the commissioning of the eFEX and jFEX modules, which now play a vital role in the ATLAS trigger system and the Run 3 data-taking period. The FEX modules will continue being used during the High Luminosity LHC (HL-LHC) operation, coping with even greater luminosities and elevated pileup conditions. The challenges to the ATLAS trigger system by the increased *pp* collision rates during Run 4, will be met by supplementing the FEX modules with additional digital systems (such as the forward FEX, fFEX), which together will form the new *L0Calo* trigger system [112].

Chapter 5

Search for charged-lepton flavour violation in top-quark processes

5.1 Introduction

The goal of this analysis is to search for the exotic, charged-lepton-flavour violating (CLFV) interaction vertex, $te\mu q$, $q = \{u, c\}$, both in top quark decay, $t \to e^{\pm}\mu^{\mp}q$, and in top quark production, $gq \to e^{\pm}\mu^{\mp}t$. Leptonically decaying τ-leptons are also considered in the analysis; they are allowed for "inclusive" limit setting, while for the Wilson coefficient limit setting, they are excluded, because the EFT operators are "flavour-dependent" as described in section 2.4, and we target only the e, μ leptons.

For the charged-lepton-flavour violating top decay, the signal (referred to as "decay signal" hereafter) considered consists of the production of a top-quark pair ($t\bar{t}$ events), where one of the top quarks decays in the CLFV mode, $t \to \ell^{\pm}\ell'^{\mp}q$, where $q = \{u, c\}$ and $\{\ell, \ell'\} = \{e, \mu\}$ or $\{\mu, e\}$, and where the other top quark of the pair decays semi-leptonically, i.e. $t \to W^{+}(\to \ell^{+}\nu_{\ell})b$, according to the SM[1]. An example Feynman diagram was given in section 2.4, fig. 2.13a.

For the charged-lepton-flavour violating top production process, the signal (referred to as "production signal" hereafter) is the production of a single-top quark with a CLFV vertex, $gq \to \ell^{\pm}\ell'^{\mp}t$, where the top-quark produced decays semi-leptonically according to the SM, as depicted in figs. 2.13b and 2.13c.

For all diagrams, the final state of the interaction is characterised by three charged leptons, two being of opposite-sign (OS) and different-flavour, and one b-quark initiated jet (called a b-jet), together with one additional light (u-, c-) jet for the CLFV decay process only, along with missing transverse momentum, $E_{\mathrm{T}}^{\mathrm{miss}}$. The formation of hadronic jets is described in section 5.2.2, while the $E_{\mathrm{T}}^{\mathrm{miss}}$ calculation is described as part of section 5.3.

The main backgrounds are from processes with a similar final state, i.e. processes which produce three leptons in the final state along with hadronic jets. Processes with three prompt (i.e. arising from the primary vertex; described in section 5.3), well-isolated leptons, which contribute significantly to the background, are diboson (WZ and ZZ) production. Secondary background processes are the production of a top-quark pair along with a vector boson or with a Higgs boson ($t\bar{t}W$, $t\bar{t}Z$, $t\bar{t}H$), and triboson (VVV, $V = \{W, Z\}$) production, as well as single-top production in association with a Z boson (tZq). Processes containing two prompt and one non-prompt lepton are $t\bar{t}$, $t\bar{t}\gamma$, $Z+$jets, $Z\gamma$ and tW production. These backgrounds contain either one lepton that originates from a heavy-flavoured hadron

[1] Charge-conjugate decays of the top quark, $\bar{t} \to \ell^{\pm}\ell'^{\mp}\bar{q}$, $\bar{t} \to W^{-}\bar{b}$, are assumed to be included throughout.

decay or, in the case of $Z\gamma$ and $t\bar{t}\gamma$, contain one electron from a photon conversion, customarily referred to as a "fake" electron. A small fraction of events contain an electron (again called "fake"), which is due to light jets with depositions in the calorimeter that are misidentified as an electron shower. We will collectively be referring to the non-prompt/fake leptons, as well as the respective backgrounds, as non-prompt.

The analysis utilises the full Run 2 data set collected by the ATLAS detector during 2015-2018, which corresponds to an integrated luminosity of 139 fb^{-1}. The structure of this chapter is as follows. In section 5.2, the basic ideas of Monte Carlo simulation and the simulated samples used, are described. The definitions of the physics objects used in the analysis are given in section 5.3. The analysis strategy is described in section 5.4. In that section, the event selection into analysis regions is defined. The events are separated into 12 regions based on the flavour and tightness (with respect to (w.r.t.) isolation and identification; described in section 5.3) of the electrons and muons. Six of these regions are high-purity Control regions (CRs) which target specific background sources, mainly sources of fake and non-prompt leptons. Three of these regions are Validation regions (VRs), and there are three Signal regions (SRs).

The SRs are designed for maximal signal acceptance while suppressing the background contributions. The SRs are produced from a single "inclusive" region, firstly by cutting on the jet multiplicity (> 1 and $= 1$) in order to obtain two SRs. A "diagonal" cut, based on the leptons' p_{T}, is placed on the $N_{\text{jets}} > 1$ SR, in order to discriminate CLFV in top quark production ("SR-production") and CLFV in top quark decay ("SR-decay"), resulting in two distinct SRs. A kinematic reconstruction, described in section 5.5, is performed in the SR-decay region, defined in order to reconstruct the SM and CLFV top quarks of the decay signal. Kinematic variables are then used as inputs to multivariate discriminators (boosted decision trees - BDTs), which are trained and deployed in the SR-decay region, to discriminate decay signal events from background events, as described in section 5.6.

The CRs are designed to target certain backgrounds, with minimal signal contamination, and their purpose is to control the normalisation of these backgrounds as part of the profile likelihood fit presented in section 5.9. A theoretical introduction to the fit is provided in section 5.8.1. In case no signal is observed, an exclusion limit is set on the cross section of the signal processes, computed by a simultaneous fit in the SRs and the CRs. The VRs are included as spectator regions in the fit, and are used to test the results of the fit in independent regions. Systematic uncertainties, described in section 5.7, are considered in the fit through nuisance parameters, and normalisation factors are utilised for the non-prompt backgrounds, which are determined by the fit in the CRs, as described in section 5.9.

5.2 Data and simulation samples

5.2.1 Data

The data analysed amount to an integrated luminosity of $139\,\text{fb}^{-1}$ collected by the ATLAS experiment during the years 2015-2018, with a centre-of-mass energy of $\sqrt{s} = 13\,\text{TeV}$, and a bunch spacing of 25 ns. All the events in the data sample belong to the ATLAS *Good Runs Lists* (GRLs), corresponding to detector conditions when all subsystems were performing correctly.

5.2.2 Monte Carlo simulation

Physics processes arising from the pp collisions happening in the ATLAS detector are simulated via Monte Carlo (MC) event generators. The use of MC generators is essential as these quantum mechanical processes are highly stochastic. MC simulation is also required to model the complicated variation of efficiency of the ATLAS detector and the analysis, as functions of angle, momentum etc. MC simulation is, in short, used to develop the analysis strategy and to evaluate the efficiency, the level of backgrounds and sensitivity of the analysis.

To model $pp \to X$, the MC generators have to take into consideration various steps [113]:

- Firstly, the quantum mechanical Matrix Element (ME) of the hard scatter of the partons (quarks and gluons) inside the incoming protons is computed to a given order [usually leading-order (LO), or next-to-leading-order (NLO), or next-to-next-to-leading-order (NNLO)] in perturbation theory of QCD [10, 15].

- To obtain the cross-section of the physics process ($pp \to X$), the parton distribution functions (PDFs) of the incoming protons have to be taken into account. The PDFs give the probability of a parton having a certain momentum fraction x of the total proton momentum. The PDFs are factorised at a given resolution scale (Q^2 or *factorisation scale*[2]) and then estimated at different energy scales, at LO or NLO using the Dokshitzer-Gribov-Lipatov–Altarelli-Parisi (DGLAP) [10, 114, 115] evolution equations.

- The strong coupling constant (α_s, usually quoted at a *renormalisation energy scale* $Q = M_Z$) is also taken into account by use of the Renormalisation Group Equations [10], which evaluate α_s at the different energy scales that appear in the generated process.

- Missing higher orders in the perturbative expansion of the ME, as well as in the DGLAP expansion of the PDFs, are usually taken into account by assigning uncertainties, evaluated by variation of the renormalisation and factorisation scales used.

- Initial State Radiation (ISR), either of gluons associated with QCD and/or of EM bremsstrahlung photons off the incoming partons, as well as Final State Radiation (FSR) off the outgoing particles, also has to be modelled.

[2]The factorization scale can be viewed as a way to separate high-energy/short-distance/perturbative from low-energy/long-distance/non-perturbative QCD interactions.

- The high-energy partons that arise from the initial hard collision emit further partons as they propagate in space-time; gluons are emitted from quarks and quark-antiquark pairs may be emitted by gluon splitting, so long as the partonic energy is high enough for them to remain unconfined. This results in a *parton shower* (PS). Parton shower calculations are made in perturbative QCD, and usually corrected by using (next-to-)next-to-leading-logarithmic [(N)NLL] resummation [116, 117] to fix divergences when low-p_{T}, "soft and collinear", gluons are radiated.

- Partons within the incoming protons, which do not account for the main physics process under study, may produce further *multiple interactions*, resulting in the so-called *underlying event*.

- As partons propagate away from each other, and as their energy decreases, confinement dictates that the quarks combine into colourless hadrons, mainly light mesons, a process called *hadronisation* (or fragmentation). The result is the formation of a hadronic jet. The PDF and hadronisation calculations are non-perturbative.

- A renormalisation and factorisation scale is also assigned when parametrising the parton shower and fragmentation processes.

- In the case of unstable particles, such as heavy (b-, c-) hadrons, a secondary decay into further particles may happen. In such cases a secondary vertex might be produced within a jet of light hadrons.

- The interaction of particles with the detector material is taken into account by the simulation, usually within the GEANT4 [118] framework. Alternatively, the faster ATLFAST [119] simulation (AFII) can be used, which uses reduced and parametrised information for the calorimeters.

- Reconstruction of the particles which traverse the detector is performed using the same techniques as used for the real data [120].

Figure 5.1 illustrates the evolution from the hard scatter to the hadronisation and possible secondary hadronic decays.

The MC generators most commonly used by the ATLAS Collaboration are the following [122].

- Multi-purpose generators, which include parton showering and fragmentation, as well as the initial ME simulation: Pythia [123, 124], Herwig [125–127], Sherpa [121, 128].

- Powheg Box [129–133], MadGraph/MadEvent [134–137] MadGraph 5 [138], MC@NLO [139, 140], MadGraph5_aMC@NLO [141] are parton-level ME generators, which need to be interfaced to a generator from the previous group for the PS and fragmentation.

- There also exist add-on packages to the generators for specific purposes, such as TAUOLA [142], dedicated to τ-decays; Photos [143], used for QED corrections in weak boson decays; EvtGen [144], used for the decay of heavy (b-, c-) hadrons.

In the following we provide information on the simulation of the samples used in this analysis (summarised in table 5.1).

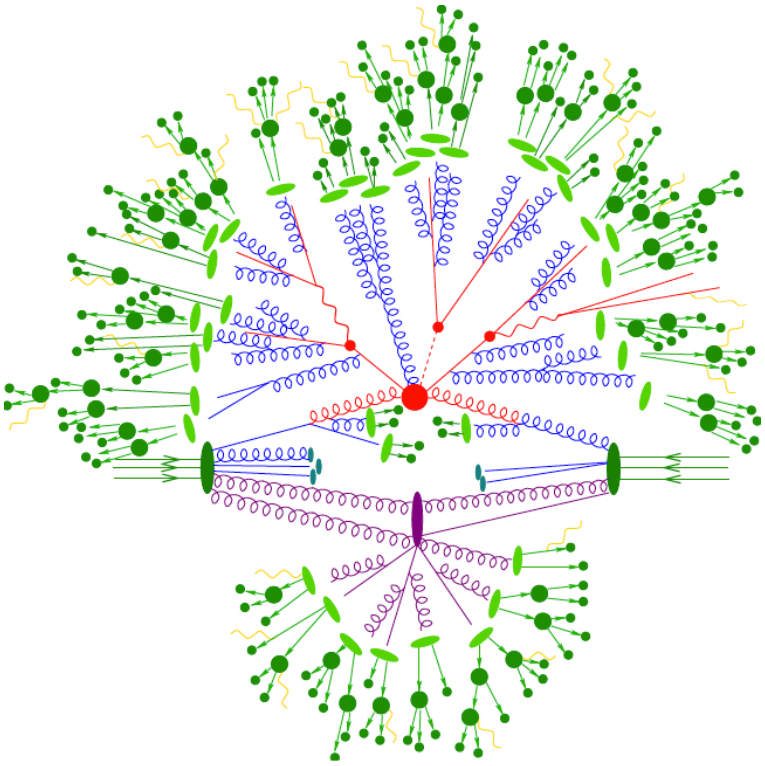


Figure 5.1: Image showing the formation of a parton shower. Partons from the incoming protons produce the hard scatter (in red). The partons produced further radiate gluons (parton shower/QCD bremsstrahlung, in blue); at lower energies/higher distances, the partons recombine (hadronisation, in green). Hadrons may decay further (dark green), possibly emitting QED bremsstrahlung (yellow). The underlying event is initiated by the hard process shown in purple. Image by SHERPA [121].

5.2.3 Signal sample

The signal processes are simulated by use of a *UFO model* [145] containing the EFT operators listed in table 2.2. It has been created with FEYNRULES 2.0 [146], using the DIM6TOP model [147] as a starting point. Events are generated at leading order in QCD with MADGRAPH5_AMC@NLO2.9.5 for the hard process in combination with PYTHIA 8.306 for parton showering and hadronisation. The renormalisation and factorisation scales (μ_R, μ_F) are dynamic and correspond to the centre of mass energy of the incoming partons for the decay diagrams and half the sum of the transverse masses of all final state particles and partons for the production diagrams. The NNPDF31_NLO_AS_0118 parton distribution function (provided by the NNPDF group [148]) is chosen; PYTHIA8 is configured according to the A14 set of tuned parameters [149].

To evaluate the parton shower uncertainty on the signal process, a set of signal samples is produced using MADGRAPH 2.9.5 for the hard scattering process interfaced to HERWIG 7.1.6 for parton showering and hadronisation. These events are also produced at leading order in QCD with the NNPDF31_NLO_AS_0118 PDF.

5.2.4 Background samples

5.2.4.1 $t\bar{t}(+\gamma)$

The production of $t\bar{t}$ ($\rightarrow \ell\ell$ +jets) events is modelled using the Powheg Box v2 generator, which provides matrix elements at NLO in the strong coupling constant α_s, and the NNPDF3.0nlo [148] parton distribution function. The h_{damp} parameter[3], which regulates ("damps") the first high-p_{T} gluon radiation beyond the Born approximation (1st order in perturbation theory of quantum scattering, see e.g. [150]) against which the $t\bar{t}$ system recoils [151], was set to $1.5\,m_t$, where m_t is the top quark mass [152]. The QCD renormalisation and factorisation scales are set to $\sqrt{m_t^2 + p_{\text{T},t}^2}$, where $p_{\text{T},t}$ is the top quark's transverse momentum. The events are interfaced with Pythia 8.230 for the parton shower and hadronisation, using the A14 tune [153] and the NNPDF2.3lo [148] set of PDFs.

The size of the $t\bar{t}$ sample is normalised to the cross-section prediction at NNLO in QCD, including the resummation of NNLL [116] soft-gluon terms calculated using Top++ 2.0 [154–160].

The impact of using a different parton shower and hadronisation model is evaluated by comparing the nominal $t\bar{t}$ sample with an event sample also produced with the Powheg Box v2 generator, but interfaced with Herwig 7.1.3.

The $t\bar{t}\gamma$ sample is simulated at LO by MadGraph5_aMC@NLO 2.3.3 with the NNPDF2.3lo PDF set, interfaced with Pythia 8.212 using the A14 tune. The photon could be radiated from an initial-state charged parton, an intermediate top quark, or any of the charged final-state particles.

An *overlap removal* procedure, based on the presence of a high-p_{T} photon as identified by "truth" information (information on the origin and type of the final-state particles given by the generator), is applied to avoid double-counting of events in $t\bar{t}$ and $t\bar{t}\gamma$ samples. The decays of bottom and charm hadrons are simulated using the EvtGen 1.6.0 program.

5.2.4.2 Single-top samples

Single-top associated tW production is modelled using the Powheg Box v2 generator with the NNPDF3.0nlo PDF set. A *diagram removal* scheme (DR) [161] is employed to handle the interference with $t\bar{t}$ production [152], by removing events in which the W boson and an associated b quark form a top quark. The events are interfaced with Pythia 8.230 using the A14 tune and the NNPDF2.3lo PDF set. The inclusive cross-section is corrected to the theoretical prediction calculated at NLO in QCD with NNLL soft-gluon corrections [162, 163]. Systematic uncertainties on the tW background component are considered by comparing the diagram removal and the *diagram subtraction* schemes, the former removing resonant $t\bar{t}$ effects at the amplitude level, while the latter is applied at the cross-section level [164, 165].

The production of tWZ events is modelled using the MadGraph5_aMC@NLO 2.3.3 generator at NLO with the NNPDF3.0nlo PDF set. The events are interfaced with Pythia 8.212 using the A14 tune and the NNPDF2.3lo PDF set. The renormalisation and factorisation scales are set to the top-quark mass. The diagram removal scheme described in [161] is employed to handle the interference

[3] h_{damp} is related to the choice of resummation scale, separating hard emissions described by the matrix element calculation from soft emissions which are included instead by the parton shower generator.

between tWZ and $t\bar{t}Z$, and is applied to the tWZ sample.

The tZq sample is simulated using the MadGraph5_aMC@NLO 2.3.3 generator at NLO with the NNPDF3.0nlo PDF set. The events are interfaced with Pythia 8.230 using the A14 tune and the NNPDF2.3lo PDF set. Following the discussion in [130], the functional form of the renormalisation and factorisation scales is set to be $4\sqrt{m_b^2 + p_{\mathrm{T},b}^2}$, where the b-quark chosen is the one produced by a gluon-splitting in the event. The tZq total cross-section is calculated at NLO using MadGraph5_aMC@NLO 2.3.3 with the NNPDF3.0nlo PDF set. The decays of bottom and charm hadrons are simulated using the EvtGen program.

5.2.4.3 $t\bar{t}V$

The production of $t\bar{t}W$ events is modelled using Sherpa 2.2.10 [121]. The sample is generated using NLO accuracy for matrix elements with up to one additional jet and LO accuracy for up to two additional jets, and using the NNPDF3.0NNLO PDF set.

The associated production of a top quark-antiquark pair with a leptonically decaying Z boson ($t\bar{t}Z$) is modelled using the MadGraph5_aMC@NLO 2.8.1 generator, with the NNPDF3.0nlo PDF set. Top quark decays are modelled at LO using MadSpin [166, 167] to preserve all spin correlations. The events are interfaced with Pythia 8.244 for the parton shower and hadronisation, using the A14 tune and the NNPDF3.0nlo PDF set. The decays of bottom and charm hadrons are simulated using EvtGen 1.7.0.

The renormalisation and factorisation scales are set to $H_{\mathrm{T}}/2$ for both samples, where

$$H_{\mathrm{T}} = \sum_i m_{\mathrm{T},i} = \sum_i \sqrt{m_i^2 + p_{\mathrm{T},i}^2}\,,$$

where the sum is over all outgoing partons in the matrix-element calculation.

5.2.4.4 $t\bar{t}H$

The production of $t\bar{t}H$ events is modelled using the Powheg Box v2 generator with the NNPDF3.0nlo PDF set. The functional form of the renormalisation and factorisation scales is set to $\sqrt[3]{m_{\mathrm{T}}(t) \cdot m_{\mathrm{T}}(\bar{t}) \cdot m_{\mathrm{T}}(H)}$. The events are interfaced to Pythia 8.230 [124] using the A14 tune and the NNPDF2.3lo PDF set. The decays of bottom and charm hadrons are performed by EvtGen 1.6.0. The cross-section is calculated at NLO in QCD, with NLO corrections in the electroweak (EW) coupling, using MadGraph5_aMC@NLO, as reported in [168].

5.2.4.5 VV, VVV

Samples of diboson final states ($VV : WZ \to 3\ell\nu$, $ZZ \to 4\ell$) are simulated with Sherpa 2.2.2, including off-shell effects and Higgs boson contributions, where appropriate. Fully leptonic final states and semileptonic final states, where one boson decays leptonically and the other hadronically, are generated using matrix elements at NLO accuracy in QCD for up to one additional parton emission, and at LO accuracy for up to three additional parton emissions. Samples for the loop-induced processes $gg \to VV$ are generated using LO-accurate matrix elements for up to one additional parton emission for both

the cases of fully leptonic and semileptonic final states. The matrix element calculations are matched and merged with the SHERPA parton shower based on Catani–Seymour dipole factorisation [169, 170] (a way of treating infrared divergences in the parton showering calculations) using the MEPS@NLO prescription [171–174]. The virtual QCD corrections are provided by the OPENLOOPS library [175–177]. The NNPDF3.0NNLO set of PDFs is used, along with the dedicated set of tuned parton-shower parameters developed by the SHERPA authors.

The production of triboson (VVV; $V = W, Z$) events is simulated with the SHERPA 2.2.2 generator using factorised gauge-boson decays. Matrix elements, accurate to NLO for the inclusive process and to LO for up to two additional parton emissions, are matched and merged with the SHERPA parton shower as done for VV.

5.2.4.6 $Z+$jets$(+\gamma)$

The production of $Z(\to \ell\ell)$+jets is simulated with the SHERPA 2.2.1 [121] generator. NLO-accurate matrix elements for up to two partons, and LO-accurate matrix elements for up to four partons are calculated with the Comix [169] and OPENLOOPS libraries. The default SHERPA parton shower [170] based on Catani–Seymour dipole factorisation and the cluster hadronisation model [178] is used. They employ the dedicated set of tuned parameters developed by the SHERPA authors and the NNPDF3.0NNLO PDF set [148].

The NLO matrix elements for a given jet multiplicity are matched to the parton shower using a colour-exact variant of the MC@NLO algorithm [171]. Different jet multiplicities are then merged into an inclusive sample using an improved CKKW matching procedure (a scheme for merging the ME with the PS by use of *Sudakov logarithms*, essential for the treatment of infinities at the low-energy gluon emission limit in the PS calculations) [173, 174], which is extended to NLO accuracy using the MEPS@NLO prescription [172]. The size of the Z+jets samples is normalised to the NNLO prediction [179].

The production of $Z(\to \ell\ell) + \gamma$ final states is simulated with the SHERPA 2.2.4 [121] generator. Matrix elements at LO accuracy in QCD for up to three additional parton emissions are matched and merged with the SHERPA parton shower using the MEPS@LO prescription [171–174]. Samples are generated using the NNPDF3.0NNLO PDF set, along with the dedicated set of tuned parton-shower parameters developed by the SHERPA authors. The ME+PS matching [171] is employed for different jet multiplicities, which are then merged into an inclusive sample using CKKW matching, as for $Z+$jets.

An overlap removal procedure is applied to remove events in the Z+jets samples that have a photon in the matrix element. The discriminant used to remove the partial overlap between Z+jets and $Z\gamma$ is the presence of a photon that has been emitted by a charged lepton, but is separated from it by $\Delta R > 0.1$ and has a p_{T} larger than 7 GeV.

5.2.4.7 Other SM processes

The following samples make minor contributions to the total background and are discussed briefly.

Higgs associated production, WH, ZH, is produced inclusively at LO by PYTHIA8+EVTGEN, while

the cross-sections are corrected by k-factors[4] to the NLO prediction.

The $2q2\ell$ samples from WZ and ZZ are modelled with Sherpa 2.2.1, while $2\ell2\nu$ samples from ZZ are modelled with Sherpa 2.2.2 in a similar way to the dibosons of section 5.2.4.5.

The $t\bar{t}t$ and $t\bar{t}WW$ samples are modelled with MadGraph5_aMC@NLO at NLO in QCD, and interfaced to Pythia8 for PS and hadronisation.

The $t\bar{t}t\bar{t}$ events are modelled using the MadGraph5_aMC@NLO 2.3.3 generator with the NNPDF3.1nlo PDF set. The events are interfaced with Pythia 8.230 for the parton shower and hadronisation, using the A14 tune and the NNPDF2.3lo PDF set. The decays of bottom and charm hadrons are simulated using the EvtGen 1.6.0 program.

Physics process	MC generator (+ PS, hadronisation)	Accuracy (ME, PDF)
Signal	MadGraph5_aMC@NLO + Pythia8	LO, NLO
WZ ($\to 3\ell\nu$)	Sherpa 2.2.2	(N)LO, NNLO
ZZ ($\to 4\ell$)	Sherpa 2.2.2	(N)LO, NNLO
$Z(\to \ell\ell)$+jets	Sherpa 2.2.1	(N)LO, NNLO
$Z(\to \ell\ell) + \gamma$	Sherpa 2.2.4	(N)LO, NNLO
$t\bar{t}$ ($\to \ell\ell + jets$)	Powheg Box + Pythia8	NLO, NLO
$t\bar{t}\gamma$	MadGraph5_aMC@NLO + Pythia8	LO, LO
$t\bar{t}H$, $H \to WW, ZZ, \tau\tau$	Powheg Box + Pythia8	NLO, NLO
$t\bar{t}Z$ ($Z \to \ell\ell$)	MadGraph5_aMC@NLO + Pythia8	NLO, NLO
$t\bar{t}W$	Sherpa 2.2.10	NLO, NNLO
tZq	MadGraph5_aMC@NLO + Pythia8	NLO, NLO
tW	Powheg Box + Pythia8	NLO, NLO
tWZ	MadGraph5_aMC@NLO + Pythia8	NLO, NLO
VVV	Sherpa 2.2.2	(N)LO, NLO
others:		
WZ ($\to 2\ell2q$)	Sherpa 2.2.1	(N)LO, NNLO
ZZ ($\to 2\ell2q$)	Sherpa 2.2.1	(N)LO, NNLO
$2\ell2\nu$	Sherpa 2.2.2	(N)LO, NNLO
$3t$	MadGraph5_aMC@NLO + Pythia8	NLO, NLO
$4t$	MadGraph5_aMC@NLO + Pythia8	NLO, NLO
$t\bar{t}WW$	MadGraph5_aMC@NLO + Pythia8	NLO, NLO
WH, ZH	Pythia8	LO, LO

Table 5.1: Information on the MC generators used for the simulation of physics processes considered in the analysis. In the third column the accuracy of the simulation in the matrix element (ME) and PDF calculations is quoted; in case different approximations in the perturbative expansion of the ME are used for a sample, depending on the jet multiplicity, a parenthesis is used. e.g. (N)LO for WZ. Only the nominal samples are presented. For "alternative" samples, used for the evaluation of systematic uncertainties, see section 5.2.4.

[4]The k-factors are scale factors applied to the cross-section to normalise it to a higher-order precision in the ME approximation.

5.3 Physics object definitions

The physics *objects* considered in this analysis are electrons, muons, jets, b-jets and missing transverse momentum. Each object is characterized by the reconstruction and identification criteria specified below.

5.3.1 Tracks and primary vertex

Tracks are reconstructed in the inner detector, by combining signals from adjacent channels in the Pixel and SCT subdetectors into clusters. These are interpreted as deposits left by individual traversing charged particles [72]. Pairs of one-dimensional SCT clusters on either side of a sensor module, or single pixel clusters, are then further converted into three-dimensional *space-points*, with position uncertainties determined by the detector geometry and sensor pitch. Sets of three space-points are used to form a large number of track seeds. These seeds are extended along "search roads" through a *combinatorial Kalman filter* [180], which searches for adjacent clusters both outwards and inwards in radius, while attempting to smooth the trajectory [181]. The potential track candidates then undergo further refinement by implementation of algorithms to resolve *ambiguities* in the cluster-to-track association. Finally an extension outwards, to the TRT, is attempted by several algorithms. Again a Kalman filter is used to build candidate extensions. The TRT hits are added and whole tracks are again refitted with a global χ^2 fitter [182].

Candidate *primary vertices* are formed from reconstructed tracks that are spatially compatible with the luminous interaction region, which is determined by the average collision point and the sizes of the two proton beams at the interaction point [183]. The primary vertex (PV) associated with the hard-scattering interaction, is chosen to be the vertex with the highest $\sum_{tracks} p_{\mathrm{T}}^2$ [184], where the sum extends over all associated tracks with $p_{\mathrm{T}} > 0.5\,\mathrm{GeV}$ [185]. At least two tracks are required to form the primary vertex [186].

Each reconstructed track is characterised by five *perigee* parameters with respect to a reference point (hereafter identified as the primary vertex). The transverse and longitudinal impact parameters, d_0 and z_0, are defined as the transverse and longitudinal distances of the single point (the *perigee* point) of closest approach to the primary vertex; d_0 is measured[5] in the x-y (i.e. transverse) plane, while z_0 is measured along the z (i.e. the beam-axis) direction. The other three parameters are the azimuthal angle ϕ and the polar angle θ of the track momentum at the primary vertex, and the ratio $q/|\vec{p}|$ of the charge of the reconstructed track divided by the magnitude of its momentum [182] (fig. 5.2).

Heavy flavoured hadrons, such as mesons containing a c or b quark, typically have a large lifetime[6] compared to particles which decay before entering the detector (such as the heavy vector bosons); leptons originating from their semi-leptonic decays tend to acquire a large impact parameter (of the order of a few mm) [187]. We can therefore exploit different requirements on d_0 to preferentially select processes which contain such non-prompt (i.e. not arising from the PV) leptons.

[5] d_0 is signed according to the track's angular momentum w.r.t. the z-axis.
[6] Indicatively, the mean lifetime of a $B^{\pm}$ meson is 1.6×10^{-12} s.

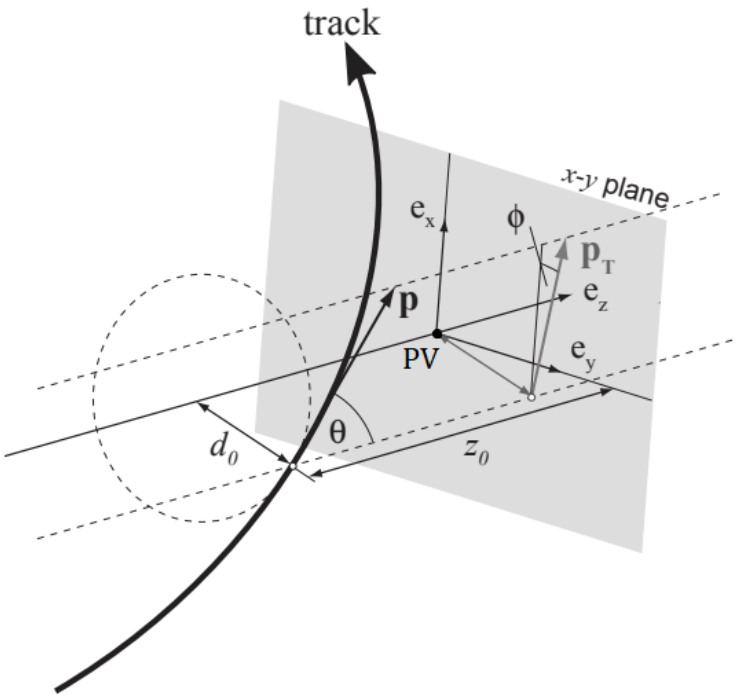


Figure 5.2: The *perigee* parameters describe the track at its point of closest approach to the primary vertex (PV). The parameter d_0 is the impact parameter measured in the transverse (or x-y) plane, while z_0 is measured along the beam (or z) axis [182].

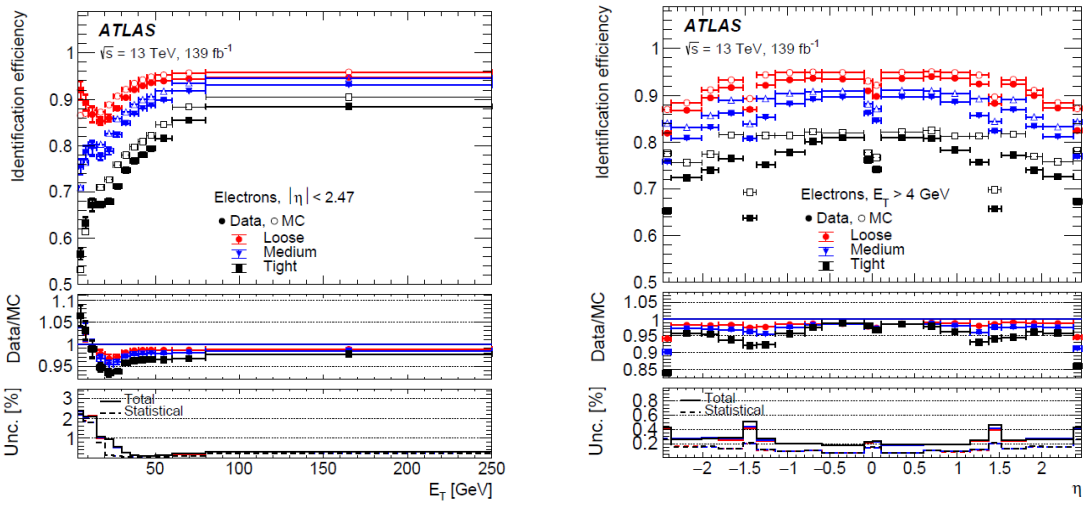
5.3.2 Electrons

Electron candidates are reconstructed from localised energy deposits in *clusters* of the electromagnetic calorimeter cells, that are associated with charged particle tracks reconstructed in the inner detector [188–190]. Only candidates with $|\eta_{\text{clu}}| < 2.47$ and $p_{\text{T}} > 10\,\text{GeV}$ are considered in the analysis. Candidates in the transition region between different electromagnetic calorimeter components, $1.37 < |\eta_{\text{clu}}| < 1.52$, are rejected.

The reconstruction efficiency depends on the $|\eta|$ and E_{T} of the electron candidate. The method used to obtain the efficiencies is described in Refs. [189, 191] as a *tag-and-probe* method. The efficiencies are evaluated in data and in simulated samples using electrons from $Z \to ee$ and $J/\psi \to ee$ decays, the latter targeting low-p_{T} electrons. The events are selected on the basis of the electron–positron invariant mass. The *tag* electron is required to satisfy a strict selection, while the selection on the *probe* electron is loosened. The efficiency of a given requirement is determined by applying it to the probe electron. For electrons with E_{T} above 15 GeV, the reconstruction efficiency is around 99% [191].

For the *identification* (*ID*) of electrons, a multivariate likelihood discriminant, combining shower shape and track information, is used to distinguish real electrons from hadronic showers [192]. Since different analyses have different requirements for the electron selection efficiency and background rejection, several "working points" (WPs) are defined by using different values of the likelihood. The likelihood threshold values are varied according to the p_{T} and $|\eta|$ of the electron candidate, so that the selection efficiency varies smoothly with the electron p_{T}. The *ID* WPs used in this analysis are the `TightLH` and the `LooseAndBLayerLH` WPs; the latter requires a hit in the IBL (to suppress candidates originating from photon conversions), in addition to the `LoooseLH` WP [192]. Figure 5.3 shows the p_{T}

and η dependence of the efficiency of the commonly used WPs, obtained by the tag-and-probe method using a $Z \to ee$ sample. For `TightLH` electrons, the efficiency degrades at momenta/energies lower than 50 GeV, while for `LoooseLH` electrons, the efficiency is close to 90% even for low electron energies, which means that most electrons are retained by the looser criterion.



(a) (b)

Figure 5.3: (a) Efficiency of different identification working points used in Run 2, as a function of the electron transverse energy E_{T} (a), and of the electron pseudorapidity η (b). In the top panel, the inner and outer error bars represent the statistical and the total uncertainties. The middle panel shows the ratios of the efficiencies measured in data to those in MC simulation. The bottom panel shows the statistical and the total uncertainties in the data/MC ratio. The efficiency is lowered in the transition region between the electromagnetic calorimeter barrel and endcap, $1.37 < |\eta| < 1.52$, to keep background levels manageable. From [193]. Additional plots in [188, 190, 194].

Isolation (*ISO*) variables are used to reduce the background from non-prompt electrons produced in hadronic decays; these electrons are typically found inside hadronic jets. The calorimetric isolation variable `topoetcone30` is defined as the sum of transverse energies of calorimeter clusters within a cone of $\Delta R = 0.3$ around the electron candidate, excluding the candidate itself. The track isolation variable `ptvarcone30` is based on the sum of transverse momenta of tracks, again within a cone of $\Delta R = 0.3$ around the electron candidate[7]. These variables, divided by the lepton p_{T}, along with the ΔR between the lepton and the track jet axis, the ratio of the lepton p_{T} and the track jet p_{T}, the number of tracks in the track jet, and finally two b-tagging (see section 5.3.5) related variables[8] are used as input to a multivariate algorithm, which outputs variables called "Prompt Lepton Veto" (PLV) WPs [195–198].

The efficiency of each isolation WP varies with the lepton p_{T}. The efficiencies of `PLVTight` and `PLVLoose` WPs for prompt electrons (70% and 90% at a p_{T} of 20 GeV, respectively) are compatible with the simpler *fixed-cut* (FC) `FCTight` and `FCLoose` WPs, which are based on simple cuts on the track and calorimeter isolation variables, as defined in [191]. For non-prompt electrons, originating from light or heavy-flavoured hadrons, the PLV WPs lead to more efficient rejection than the FC WPs.

The "nominal" longitudinal *impact parameter* criterion, $|z_0 \sin(\theta)| < 0.5\,\mathrm{mm}$, is applied to all

[7] The cone radius is allowed to vary, $\Delta R = \min(0.3, 10\,\mathrm{GeV}/p_{\mathrm{T}}^{e}[\mathrm{GeV}])$, so that for very high energy leptons it is reduced (hence the "var" in the variable name) [191].

[8] For a low p_{T} (< 12 GeV) lepton, the last two, b-tagging related, variables are dropped.

electrons, to ensure consistency with them originating from the primary vertex. The "nominal" (or "tight") requirement on the transverse impact parameter significance, defined as d_0/σ_{d_0}, where σ_{d_0} is the resolution of the track's d_0 [199], is that its absolute value should be less than five[9].

Several classes of electrons are used in the analysis as summarised in table 5.2. Tight electrons use the normal definitions, whereas different types of Loose electrons are used in the context of control regions for the data-driven estimation of the normalisation of backgrounds from non-prompt leptons. The *Isolation*, *ID* and $|d_0|/\sigma_{d_0}$ criteria of the electrons are varied, to loosen or tighten the definition of the electrons used in the analysis.

	Tight	*Loose ISO*	*Inv-d_0*	*InvIso*
Identification working point	`TightLH`	`TightLH`	`LooseAndBLayerLH`	`TightLH`
Isolation working point	`PLVTight`	`None`	`None`	`!PLVTight`
$\lvert d_0\rvert/\sigma_{d_0}$	< 5	< 5	> 5	< 5

Table 5.2: Tight and Loose electron definitions. "!" refers to electrons that fail the criterion. ID and isolation working points are described in [200].

5.3.3 Muons

Muon candidates are reconstructed by combining inner detector tracks with full tracks or track segments reconstructed in the muon spectrometer [192, 201, 202]. The corresponding muon types are called *combined* and *segment-tagged* muons. In the region $|\eta| < 0.1$[10], muon candidates are also reconstructed from inner detector tracks matched to isolated energy deposits in the calorimeters, consistent with the passage of a minimum-ionizing particle (*calorimeter-tagged* muons). The muon candidates used in this analysis are required to satisfy $p_{\mathrm{T}} > 10\,\mathrm{GeV}$ and $|\eta| < 2.5$.

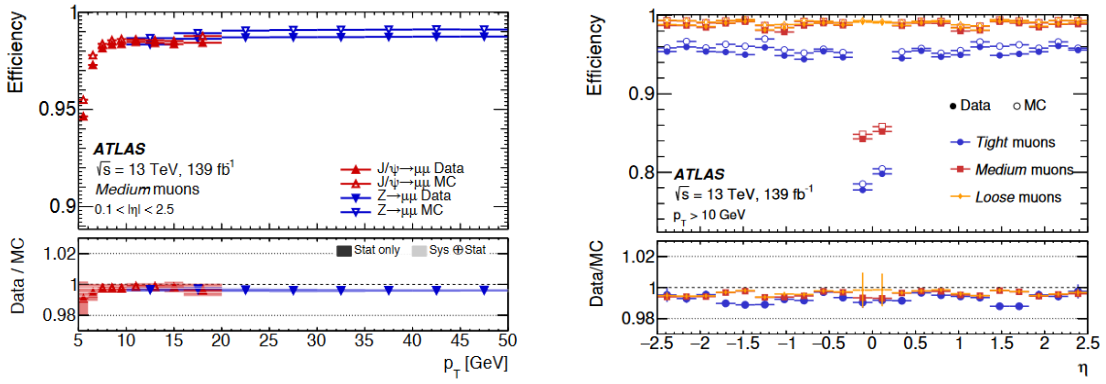
Additional requirements on the reconstructed muon candidates (such as on the number of hits in the different ID subdetectors and the different MS stations, on the track fit properties, and on variables that test the compatibility of the individual measurements in the two detector systems[11]) result in different *Identification* WPs. Figure 5.4 shows the p_{T} and η dependence of identification working points used in Run 2. They have been evaluated using a similar technique to the electron *tag-and-probe* method, using muons from $Z \to \mu\mu$ and $J/\psi \to \mu\mu$ decays. The *tag* muon is required to satisfy the `Medium` *ID* working point and the single-muon trigger requirements described in section 5.3.9, while the *probe* muon is subject to a very loose selection. The WP used in this analysis is `Medium` for all muons, with an efficiency of around 98% [192].

The *isolation* working points used are `PLVTight` and `PLVLoose`, defined by use of a multivariate likelihood discriminant combining track- and calorimeter-based information, in a similar fashion to that described for the electrons. There exist also simple, fixed-cut (FC), isolation variables defined in [201]. The efficiency of `FCTight` is comparable to the `PLVTight` for prompt muons; for non-prompt muons though, the `PLVTight` criterion leads to more efficient rejection (lower efficiency), as in the case of electrons.

[9] An alternative notation, used in the following, for the transverse impact parameter significance is d_0^{sig} .

[10] In the $|\eta| < 0.1$ region, there is a gap in the MS, essential for the passage of services to the calorimeters, the magnet solenoid and the inner detector.

[11] These criteria reduce the background from in-flight decays of light-flavour hadrons, which often result in kinked tracks.



(a) (b)

Figure 5.4: (a) Efficiency of the `Medium` identification working point used in Run 2 as a function of the muon p_{T}, for muons with $0.1 < |\eta| < 2.5$. (b) Efficiency of the `Loose, Medium, Tight` identification WPs measured in $Z \to \mu\mu$ events as a function of η, for muons with $p_{\mathrm{T}} > 10$ GeV. The efficiency is lowered in the $|\eta| < 0.1$ region, which lacks MS coverage. When not negligible, the statistical uncertainty in the efficiency measurement is indicated by the error bars. The panel at the bottom shows the ratio of the measured to predicted efficiencies, with statistical and systematic uncertainties. From [202].

The longitudinal *impact parameter* criterion is similar to that used for electrons, $|z_0 \sin(\theta)| < 0.5$ mm. The nominal ("tight") transverse impact parameter cut for muons is $|d_0|/\sigma_{d_0} < 3$.

Tight and *Loose* selections for muons used in this analysis are reported in table 5.3. Different requirements are set on $|d_0|/\sigma_{d_0}$ and on the *Isolation* working point.

	Tight	*Loose ISO*	*Inv-d_0*	*InvIso*
Identification working point	`Medium`	`Medium`	`Medium`	`Medium`
Isolation working point	`PLVTight`	`PLVLoose`	`PLVLoose`	`!PLVTight`
$\lvert d_0\rvert/\sigma_{d_0}$	< 3	< 3	> 3	< 3

Table 5.3: Tight and Loose muon definitions. "!" refers to muons failing the criterion. ID and isolation working points are described in [201].

5.3.4 Jets

Jet candidates are reconstructed from clusters of topologically connected calorimeter cells [203], using the anti-k_t [204] jet algorithm with radius parameter $R = 0.4$ implemented in the FastJet [205] software package. Jet properties are calibrated using the *Particle Flow* (PFlow) algorithm [206], which exploits both calorimeter and ID information. After calibration, jet candidates are required to have $p_{\mathrm{T}} > 25$ GeV and $|\eta| < 2.5$. To suppress jets originating from pile-up collisions, a cut on the Jet Vertex Tagger (JVT) [207], as defined in the *JVTTight* working point, is applied for jets with p_{T} below 120 GeV.

5.3.5 *b*-tagging

Jets containing *b*-hadrons, as illustrated in fig. 5.5, are identified (*b*-tagged) using the `DL1r` algorithm [208–210]. This algorithm augments the `DL1` algorithm [211, 212] by use of the `RNNIP` output [213] (a multivariate output using the impact parameters of the jet tracks), and combines it with inputs from the impact parameters with respect to displaced vertices, as well as the topological

properties of secondary and tertiary vertices within a jet. These are passed to a neural network which outputs three values, representing the probabilities of the jet being light-flavour, c-jet or b-jet. These are then combined into a single discriminant. The analysis uses the 77% efficiency working point.

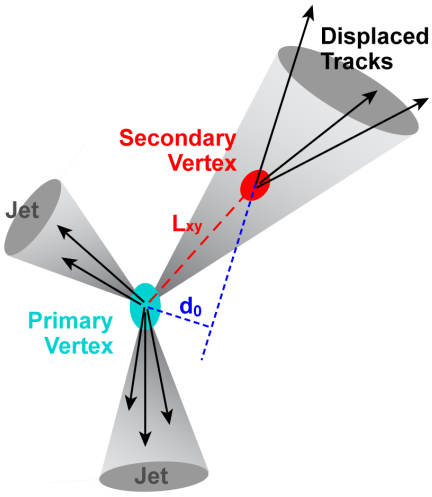


Figure 5.5: Three jets emerging from the primary vertex [214]. Two are jets originating from a light-flavour quark. One jet contains a b-hadron. The formation of a secondary vertex (SV) is shown, when the b-hadron decays semi-leptonically, typically after flying for a few mm. The transverse impact parameter d_0 of its tracks are relatively large, as is the separation L_{xy} of the secondary vertex itself (both measured in the transverse, or x-y, plane).

5.3.6 Overlap removal

Once reconstructed, some physics objects can show features that are typical of more than one category: for example an electron shower could be identified as an electron but also as a jet; or a muon originating from a semi-leptonic b-hadron decay could both be assigned to the jet and be considered as a stand-alone muon. To avoid such double counting, an overlap removing (OR) algorithm [215] is deployed. The rules applied by this OR tool are the following [216]:

- Remove any calorimeter-tagged muon sharing a track with an electron within a cone of $\Delta R < 0.2$.
- Remove any electron sharing a track with a remaining muon.
- Reject any jet within a cone of $\Delta R < 0.2$ of an electron.
- Reject any electron within a cone of $\Delta R < 0.4$ of a remaining jet.
- Reject any jet within a cone of $\Delta R < 0.2$ or *ghost-matched* [217] to a muon, if the number of tracks associated to the jet is less than three.
- Reject any muon within a cone of $\Delta R < 0.4$ of a remaining jet.

The OR algorithm is run with leptons satisfying the *Loose* lepton definitions (`LooseAndBLayerLH` electron *ID* and *ISO* `None`/`PLVLoose` for e/μ), so that there is consistency when selecting leptons with tightened or loosened criteria to form different analysis regions (regions defined in section 5.4).

5.3.7 Missing transverse momentum

The missing transverse momentum vector $\mathbf{p}_{\mathrm{T}}^{\mathrm{miss}}$, whose magnitude is indicated by $E_{\mathrm{T}}^{\mathrm{miss}}$ (or plainly MET), is defined in [218, 219]. It essentially corresponds to minus the vectorial sum of all transverse momenta, $-\sum \mathbf{p}_{\mathrm{T}}$, as the initial energy of the particles travelling transverse to the beam axis is zero, so the conservation of energy dictates that $\sum \mathbf{p}_{\mathrm{T}} = 0$. The presence of a large $E_{\mathrm{T}}^{\mathrm{miss}}$ is usually related to the presence of neutrinos which pass the detector material undetected. The $\mathbf{p}_{\mathrm{T}}^{\mathrm{miss}}$ vector is defined by

$$\mathbf{p}_{\mathrm{T}}^{\mathrm{miss}} = -\sum_{\substack{\text{selected}\\\text{electrons}}} \mathbf{p}_{\mathrm{T}}^{e} - \sum_{\substack{\text{selected}\\\text{muons}}} \mathbf{p}_{\mathrm{T}}^{\mu} - \sum_{\substack{\text{accepted}\\\text{jets}}} \mathbf{p}_{\mathrm{T}}^{\mathrm{jet}} - \sum_{\substack{\text{unused}\\\text{tracks}}} \mathbf{p}_{\mathrm{T}}^{\mathrm{track}} . \tag{5.1}$$

This is the sum of terms obtained respectively from the vectorial sums of the transverse momenta of preselected electrons, muons and jets, as defined in sections 5.3.2 to 5.3.4, with the lepton-jet overlap removal applied as in section 5.3.6. The last term, called the "soft term", is built from tracks that are not associated to any reconstructed physics object, but still associated with the primary vertex. It can contain contributions from the hard scattering process, as well as from the underlying event and from pileup interactions.

Events also have to satisfy jet cleaning and trigger matching requirements described in sections 5.3.8 and 5.3.9.

5.3.8 Jet cleaning

The event has to pass the LooseBad selection [220] of the JetCleaningTool. This tool, run after the overlap removal, discards events containing at least one fake jet (originating from non-collision background) or fake signals in the calorimeter (e.g. noise bursts).

5.3.9 Triggers

Events are selected using trigger sets that include the lowest unprescaled[12] single-lepton triggers [84] that were active for each data-taking period.

The electron triggers select a calorimeter cluster matched to a track. Electron candidates must then satisfy a likelihood-based identification criterion. In 2015, electrons had to satisfy medium identification and have $E_{\mathrm{T}} > 24\,\mathrm{GeV}$. In 2016-2018, electrons had to satisfy tight identification together with isolation criteria and have $E_{\mathrm{T}} > 26\,\mathrm{GeV}$. Two other triggers were also available to avoid efficiency losses due to identification and isolation at high p_{T}. These selected medium electrons with $E_{\mathrm{T}} > 60\,\mathrm{GeV}$ and loose electrons with $E_{\mathrm{T}} > 120\,\mathrm{GeV}$ ($140\,\mathrm{GeV}$ in 2016-2018).

Muons are triggered by matching tracks reconstructed in the muon spectrometer and in the inner detector. In 2015, muons had to satisfy a loose isolation requirement and have $p_{\mathrm{T}} > 20\,\mathrm{GeV}$. In 2016-2018, the isolation criteria were tightened and the threshold increased to $p_{\mathrm{T}} > 26\,\mathrm{GeV}$. During

[12] *Prescale* means that only a fraction of the total events that pass the trigger are accepted. They can be regarded as "support triggers", related to the data-taking conditions or used for monitoring purposes, testing of low-E_{T} threshold (and loose or no isolation) triggers, and are mainly used for performance and background studies. Primary triggers which are used for physics measurements typically have no prescale applied.

each of these years, another muon trigger without any isolation requirement was available to avoid efficiency losses due to isolation requirements at high p_{T}. This selected muons with $p_{\mathrm{T}} > 50\,\mathrm{GeV}$. The following triggers were used for each year:

2015:

`HLT_e24_lhmedium_L1EM20VH`,

`HLT_e60_lhmedium`,

`HLT_e120_lhloose`,

`HLT_mu20_iloose_L1MU15`,

`HLT_mu50`.

2016-2018:

`HLT_e26_lhtight_nod0_ivarloose`,

`HLT_e60_lhmedium_nod0`,

`HLT_e140_lhloose_nod0`,

`HLT_mu26_ivarmedium`,

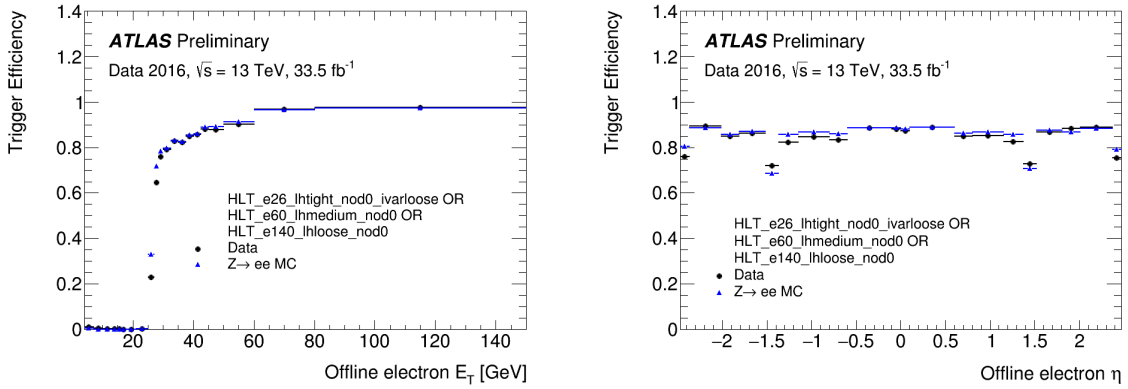
`HLT_mu50`

The naming convention is as follows: HLT indicates a High Level Trigger, e (mu) indicates an electron (muon) trigger and the subsequent number indicates the transverse energy threshold. Next, the identification level is identified (tight, loose or medium), where "lh" is used to indicate a likelihood based trigger. "nod0" indicates that no transverse impact parameter cuts are required. "i" indicates the isolation requirement applied on the lepton, e.g. ivarloose indicates a variable-sized cone isolation requirement. So, e.g. the HLT_e26_lhtight_nod0_ivarloose trigger requires an electron candidate with $E_{\mathrm{T}} > 26$ GeV satisfying the likelihood-based tight identification without applying transverse impact parameter requirements but applying variable-size cone isolation.

The L1 signifies explicitly the corresponding Level-1 trigger. EM (MU) indicates an electromagnetic (muon) trigger and the subsequent number indicates the transverse energy threshold. V indicates a pseudorapidity dependent transverse energy threshold, H indicates hadronic isolation and I indicates EM isolation. For muon triggers where L1 is not specified, the L1MU20 trigger has been used. For electrons where L1 is not specified, the Level-1 trigger requires an isolated electromagnetic cluster with $E_{\mathrm{T}} > 22$ GeV.

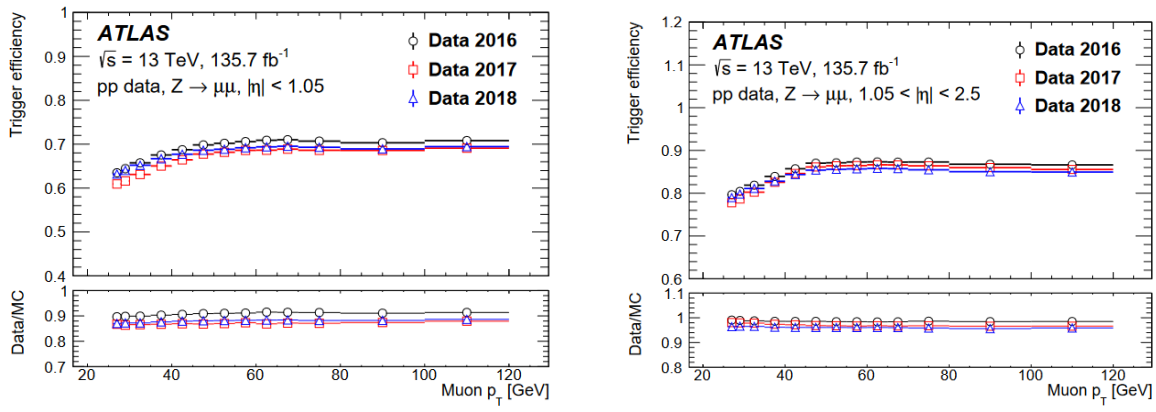
Dilepton triggers have also been considered, but their inclusion was found not to be beneficial for the analysis.

Figure 5.6 and fig. 5.7 show the p_{T} dependence (*turn-on curves*) of the trigger efficiency for electrons [221, 222] and muons [223]. The efficiencies have been evaluated with a *tag-and-probe* method using offline reconstructed leptons from $Z \to \ell\ell$ and $J/\psi \to \ell\ell$, $\ell = \{e, \mu\}$, samples. The trigger efficiency is defined as the ratio of the number of trigger-matched probe leptons relative to the total number of probe leptons.



(a) (b)

Figure 5.6: Efficiency of the logical OR between HLT_e26_lhtight_nod0_ivarloose, HLT_e60_lhmedium_nod0 and HLT_e140_lhloose_nod0 triggers as a function of the offline electron candidate's transverse energy E_{T} (a) and η (b). The efficiencies were measured with a tag-and-probe method using $Z \to ee$ decays in data and Monte Carlo. The error bars show the statistical uncertainties. From [222]. A collection of plots can be found in [224].



(a) (b)

Figure 5.7: The efficiency of passing either the HLT_mu26_ivarmedium or the HLT_mu50 trigger in the barrel (a) and endcap (b) regions as a function of the probe muon p_{T}, derived by the $Z \to \mu\mu$ tag-and-probe method, for *Medium* quality muons. Computed using data taken in 2016–2018. The error bars show the statistical uncertainties only. From [223].

Trigger matching

The presence in the event of at least one reconstructed lepton matched to a fired trigger is required. Such a lepton must have a transverse momentum larger than the nominal trigger threshold by at least 1 GeV for electrons [225] and at least 5% for muons [226].

5.3.10 Efficiency corrections

The full efficiency for a given lepton can be written as the product of the efficiencies of the separate algorithms (reconstruction, *ID*, *ISO*, trigger, as well as track-to-vertex association (TTVA) for the muons), measured by tag-and-probe methods. The differences between the efficiency measured in data for a given algorithm X, $\epsilon^{Data}(X)$ and the corresponding efficiency in simulation, $\epsilon^{MC}(X)$, shown e.g. in figs. 5.3 and 5.4, are taken into account by the ratio of these two numbers, called the efficiency scale

factor (SF) [189, 202]:

$$\mathrm{SF} = \frac{\epsilon^{Data}(X)}{\epsilon^{MC}(X)} \,. \tag{5.2}$$

It is possible to distinguish between the efficiencies of prompt leptons and leptons originating from hadronic decays, by variations of the tag-and-probe methods, e.g. by using $t\bar{t}$ or $W(\rightarrow \ell\nu)$+jets samples to probe non-prompt and non-isolated leptons. The SFs are a function of the lepton kinematics, such as the p_{T} and η. The SF quantifies the deviation of the simulation from the real detector behaviour, and is therefore used in physics analysis to correct the simulation by reweighting the simulated events.

5.4 Event selection

The events used in the analysis are separated into distinct regions, according to different selection criteria. All regions are subject to the same preselection cuts shown in table 5.4. At least (most) three loose (tight) leptons and at least one jet are required; at most two b-jets are allowed. The sum of the lepton charges must be ± 1. Low mass resonances are vetoed by excluding any opposite-sign same-flavour (OSSF) lepton pair[13] with an invariant mass below 15 GeV.

	Preselection
Description	**Value**
Lepton flavours	e or μ
Baseline lepton definition	$p_{\mathrm{T}} > 10$ GeV, $\lvert\eta\rvert < 2.5$
Baseline jet definition	$p_{\mathrm{T}} > 25$ GeV, $\lvert\eta\rvert < 2.5$
Number of leptons	$N_{\mathrm{lep}} \geq 3$
Number of tight leptons	$N_{\mathrm{lep}}^{\mathrm{tight}} \leq 3$
Number of jets	$N_{\mathrm{jets}} \geq 1$
Number of b-jets (DL1r @ 77%)	$N_b \leq 2$
Leading lepton p_{T}	$p_{\mathrm{T}}^{\mathrm{lep.\ 1}} > 27$ GeV
Sum of lepton charges	$\sum q_i = \pm 1$
OSSF lepton pair mass	$m_{\ell\ell}^{\mathrm{OSSF}} \geq 15$ GeV
Electron isolation	None
Electron ID	LooseAndBLayerLH
Muon isolation	PLVLoose
Muon ID	Medium
z_0 impact parameter	$\lvert z_0 \sin\theta_{\mathrm{lep}}\rvert < 0.5$ mm

Table 5.4: The preselection criteria which are common to all the analysis regions. Individual regions may tighten these selections, but not loosen them. Leptons are numbered 1, 2, 3, according to their p_{T} ordering.

Twelve analysis regions are then defined: three signal regions (SRs), six control regions (CRs) to target specific backgrounds ($t\bar{t}$ and $Z+$jets, separately for non-prompt e and μ; $Z\gamma$ and WZ), and three validation regions (VRs) to check the post-fit modelling. The CRs will be used for a data-driven estimation (using a simultaneous maximum likelihood fit) of the $t\bar{t}$, $Z+$jets($+\gamma$) and WZ backgrounds, since the MC modelling is not expected to be sufficiently robust[14].

The application of the `PLVTight` criterion to all leptons in the SRs significantly suppresses the non-prompt backgrounds, notably from $Z+$jets and $t\bar{t}$. The $Z+$jets background is further reduced in the SRs by applying a cut on the p_{T} (> 15 GeV) of the third leading (in p_{T}) lepton. Even though the remaining $Z+$jets background is small in the SRs, we still estimate it by use of dedicated CRs, so that we are confident of its normalisation in the VRs, where it has a considerable presence.

The primary motivations and key defining aspects of each region are as follows.

- **Signal regions**: All signal regions require 3 tight leptons with flavours matching the CLFV requirement (i.e. an OSSF lepton pair is needed). The third leading lepton p_T threshold is set to 15 GeV. By not allowing the OSSF dilepton invariant mass to be within 10 GeV of the Z-mass

[13] A lepton pair will also be referred to as *dilepton*.

[14] Non-prompt leptons present an additional challenge for the MC generators, in order to correctly model the secondary decays and the interaction of the decay products with the detector. For the WZ background, a mismodelling in comparison to data has been observed as N_{jets} increases (which has been observed also in [227], fig. 6).

(Z-mass veto), we are suppressing backgrounds containing a Z boson, which decays into a lepton pair, mainly the diboson and Z+jets contributions. The signal regions are split by the N_{jets} requirement. When requiring $N_{\text{jets}} \geq 2$, the resulting region is split into two parts, by use of a "diagonal cut" [15], which essentially separates low lepton p_{T} events from high p_{T} ones. In the first case, the SR targets the CLFV decay signal and suppresses the production signal (**SR-decay**), while in the latter case, the CLFV production-signal is targeted (**SR-production**). To enhance the production signal, an additional SR (**SR-1jet**) is defined by changing the requirement on the number of jets to N_{jets} =1; no diagonal cut is applied on SR-1jet. The production signal has a much higher yield overall in the SRs compared to the decay signal. The production signal is heavily suppressed in the SR-decay region by use of the diagonal cut, while the decay signal-to-background ratio is similar before and after the cut is applied. Separation of the decay signal versus the background in the SR-decay region is achieved by use of multivariate methods (described in section 5.6).

- **CRWZ**: The WZ diboson control region is used to estimate the normalisation of the production of the WZ background, which is the largest prompt lepton background process. The main difference from the SRs is to invert the Z-mass veto, to select events with a Z boson. An $E_{\text{T}}^{\text{miss}}$ cut is also applied to suppress the ZZ background.

- **CR**$Z\gamma$: The $Z\gamma$ control region is used to estimate the normalisation of the production of electrons from photon conversions. $Z\gamma$ events are used for this purpose for two reasons: firstly the primary photon-conversion contribution in the SRs comes from $Z\gamma$ events, and secondly the Z boson provides a suitable environment in which to identify two tight leptons, and hence focus the normalisation on the fake electron. A three-lepton invariant mass cut is used in order to target $Z\gamma$ and suppress $Z+$jets events. The electron ambiguity type, a variable related to the reconstruction quality of an electron [228], is used to select events with photon-like electron signatures. The lepton isolation is loosened to increase the number of events.

- **CR**$Zj\mu(e)$: The muon(electron)-based Z+jets control region is used to estimate the normalisation of the production of non-prompt muons (electrons) from heavy-flavour hadronic decays. The third leading lepton is required to be a muon (electron) with a p_{T} requirement that has been inverted with respect to the SRs, in order to preferentially select soft-p_{T} lepton events, which are more likely to be non-prompt. The isolation requirement on the third leading lepton is loosened, and the nominal d_0 significance requirement inverted. A Z-mass window is applied to the OSSF lepton pair invariant mass, to ensure a Z boson is present in the event.

- **CR**$t\bar{t}\mu(e)$: The muon(electron)-based $t\bar{t}$ control region is also used to estimate the normalisation of the production of non-prompt muons (electrons) from heavy-flavour hadronic decays. Since the kinematics of $t\bar{t}$ and $Z+$jets events are different, we choose to separate the non-prompt

[15] The "diagonal cut" is defined by $3p_{\text{T}}^{\ell_1} + \sum m_{\ell\ell(\prime)}^{\text{OS}} < 400$ GeV for the SR-decay and > 400 GeV for the SR-production region, where with ℓ_1 we denote the highest p_{T} lepton. The $m_{\ell\ell(\prime)}^{\text{OS}}$ is the dilepton invariant mass for same flavour or different flavour leptons, hence two terms enter the sum (two OS lepton pairs exist per event). The relevant study was performed in a 3 muon region by the $\mu\tau$ channel analysis [27]. The variable used for the diagonal cut is plotted in fig. 5.10.

background normalisation according to these source samples. The p_{T} requirement on the third leading lepton has been loosened with respect to the SRs, in order to enhance soft-lepton events, while the isolation and d_0 cuts are the same as in CRZjμ/e. The $t\bar{t}$ purity is improved by requiring a leading $e\mu$ (i.e. different-flavour) lepton pair, and also by vetoing the Z-mass window for same-flavour dileptons. Application of the "diagonal cut" used in the SR-decay region, suppresses contamination from the production signal.

- **VR**$low\ell_3 p_{\mathbf{T}}$: The low p_{T} third-lepton validation region is used to check the post-fit modelling of, mainly, the non-prompt lepton background distributions in a region that is close to the signal region. It is similar to the SR-decay region definition, but with an inverted cut on the third leading lepton p_{T}.

- **VR**$\mu(e)$: The three-muon (-electron) validation region is used to check the post-fit modelling of the backgrounds, in a region close to the signal region. It is identical to the SR-decay region definition but with a three-muon (-electron) flavour requirement and a loosened p_{T} cut on the third leading lepton.

The SRs, VRs and the CRWZ can collectively be described as "Tight" regions, because they require three *tight* leptons. It is possible that these events contain an additional (4^{th}) lepton that is loose (but not tight) as a very small fraction of the total yield. The control regions CR$Z\gamma$, CRZ$j\mu/e$ and CR$t\bar{t}\mu/e$, can be collectively called "Non-prompt" regions, as they target backgrounds that contain one fake or non-prompt lepton. These regions require two *tight* leptons (the leading ones, with the exception of CR$Z\gamma$, where the isolation is loosened for all three leptons), plus one *loose* lepton.

A diagram summarising the basic selection criteria of the analysis regions is given in fig. 5.8. The full details of the selection criteria are provided in tables 5.5 to 5.7.

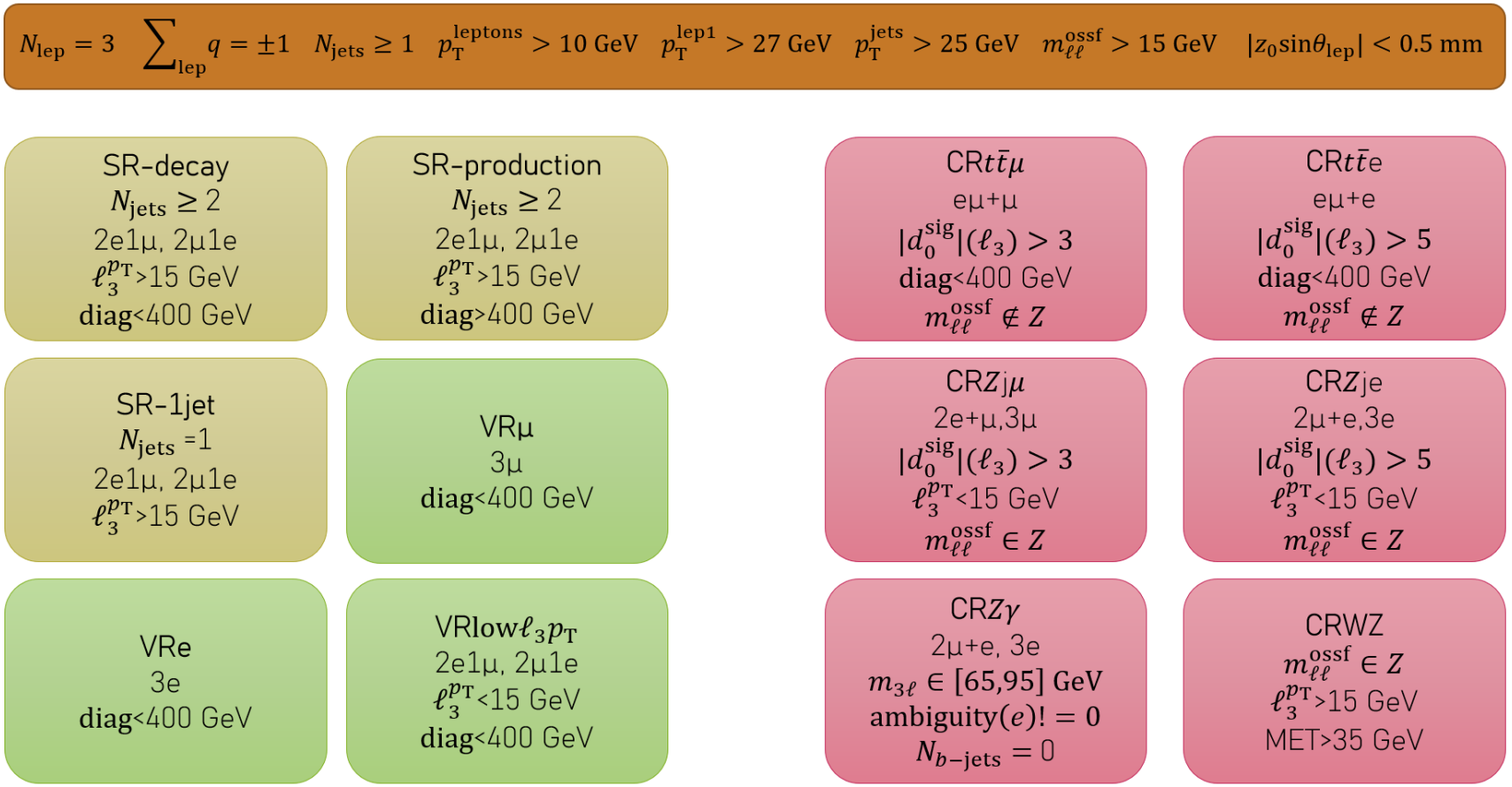
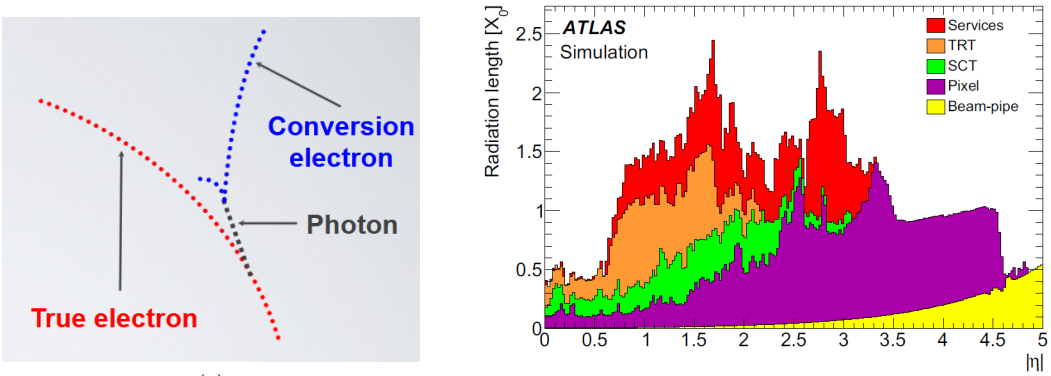


Figure 5.8: A diagram summarising the basic selection criteria of the analysis regions. "diag" refers to the application of the diagonal cut, defined in footnote 15. Leptons (ℓ) are labelled with a 1, 2, 3 subscript according to their p_T ordering. The regions on the left (signal regions and validation regions) and CRWZ contain *Tight* leptons. The Control Regions $CRt\bar{t}(\mu/e)$ and $CRZj(\mu/e)$ have an *Inv-d_0* lowest p_T lepton (i.e. third leading lepton). Finally $CRZ\gamma$ has *Loose ISO* leptons. For the definition of *Tight*, *Inv-d_0*, *Loose* leptons, refer to tables 5.2 and 5.3.

Electron charge mis-reconstruction

The electron charge sign can be incorrectly reconstructed due to an instrumental or a physical effect. The former is a result of a wrong reconstruction of the curvature of the associated track in the magnetic field of the Inner Detector, especially for high p_{T} electrons [229]. The latter may occur when an electron emits a hard bremsstrahlung photon, which converts asymmetrically in the detector material. If the positron created in the photon conversion carries most of the energy of the original electron, the electron can be reconstructed, but with opposite charge (fig. 5.9a). The effect is most prominent at higher $|\eta|$ values, as the detector material in front of the EM calorimeter increases (fig. 5.9b). This electron charge misidentification has been taken into account by applying an "ECIDS" (an acronym stemming from "Electron Charge ID Selector") cut on all tight electrons. The ECIDS tool [230] is a multivariate technique, using BDTs with electron kinematic and reconstruction variables as input (such as the p_{T} and η, the energy measured in the calorimeter divided by the track's momentum, and the number of SCT hits), which indicates whether an electron is likely to have been subject to a "charge-flip". In the "Non-prompt" Control Regions, the ECIDS cut has not been applied to the third lepton due to a significant decrease in the yield, as well as the fact that the backgrounds targeted by these CRs will be subject to a normalisation factor obtained by the fit to the data.



(a) (b)

Figure 5.9: (a) Schematic diagram of an electron emitting a bremsstrahlung photon, which converts into e^+e^- pair [231]. (b) Amount of material in front of the cryostat, housing the solenoid and the EM calorimeters, in units of radiation length, X_0, traversed by a particle as a function of $|\eta|$. The contributions of the different detector elements, including the services, are shown separately by filled coloured areas [188, 232].

Event weights

The MC simulated event yields are evaluated by summing the event weights, as given by the MC simulation packages; the corresponding statistical error is evaluated by summing the squares of the weights and then taking the square root of the result. The event weights are a product of individual weights which account for the corresponding luminosity of the data-taking period, the cross section of the process (corrected by the relevant k-factor; defined in footnote 4) and the available MC statistics. Weight corrections related to the MC generator scales and PDF used, the pileup and jet-to-vertex (JVT) association, the b-tagging efficiency, as well as the lepton efficiency corrections (scale factors), have also been included. The application of the ECIDS cut has also been accounted for by the relevant weight

correction.

For a "good" (i.e. with low statistical uncertainty) MC statistical sample, it is best to have a large number of events, with small weights (with values < 1). A small number of MC events with large event weights (> 5) have been observed in some of the regions. These events belong to the Z+jets sample, and one event belongs to the ZZ sample. These weights result in "non-physical" spiky structures in the distributions of the variables in the respective regions. These weights have been rounded to ± 1, following the recommendations of the ATLAS Physics Modelling Group [233].

The MC and data event yields in each region, broken down by the physics processes, are given in tables 5.8 to 5.11. The signal processes are normalised to a cross-section corresponding to all Wilson coefficients (defined in section 2.4) equal to the same value of $|C| = 0.3$. The comparison with the cross-section values shown in tables 2.5 and 2.6 can be done by multiplying the table's values with 0.09, taking into account that $\sigma_{\text{cLFV}} \propto |C|^2$. This normalisation is motivated by the limits obtained by the previous ATLAS study of 80 fb^{-1} [2], as well as in order to moderate the increase of the signal yield in the analysis regions.

The $t\bar{t}$, $Z+\text{jets}$ and $Z\gamma$ processes are split by the non-prompt lepton content. The fraction of $t\bar{t}$ and $Z+\text{jets}$ that does not contain a heavy-flavour (HF) hadronic decay (a photon conversion for $Z\gamma$), is denoted by *other* in these tables, and may refer to events containing leptons from τ decays, light flavour hadronic decays, and γ-conversions (in $Z+\text{jets}$) or HF hadron decays (in $Z\gamma$).

The H_{T}, defined as the scalar sum of p_{T} of leptons and jets, and $E_{\text{T}}^{\text{miss}}$ distributions for each region are shown in figs. 5.11 to 5.14. Most regions have good agreement between data and MC simulation; a mismodelling is observed in CR$Z\gamma$, VRe, VRlow$\ell_3 p_{\text{T}}$ and in CRWZ. The production signal in SR-production region, having highly boosted leptons in the final state, exhibits a good separation from the backgrounds, as it is concentrated in high values of H_{T}. It can also be seen that the Control Regions are highly pure in the backgrounds that they target ($t\bar{t}$, $Z+\text{jets}$, $Z\gamma$, WZ). Additional kinematic distributions are plotted in Appendix A.

Name of region	SR-decay	SR-prod.	SR-1jet
Description	**Value**		
Number of leptons	$N_{\text{lep}}^{\text{tight}} = 3$		
Lepton flavours	$2\mu 1e$, $2e1\mu$		
Third leading lepton p_{T}	$p_{\text{T}}^{\text{lep. 3}} > 15$ GeV		
Number of b-jets	$N_b \leq 1$		
Z-mass cut on the OSSF dilepton	$\lvert m_{\ell\ell}^{\text{OSSF}} - m_Z \rvert > 10$ GeV		
Trilepton inv. mass $m_{3\ell}$	–		
Electron isolation	PLVTight		
Electron ID	TightLH		
Electron ambiguity type	0 (electron-like)		
Muon isolation	PLVTight		
Muon ID	Medium		
Electron transverse impact parameter significance	$\lvert d_{0,e}^{\text{sig}} \rvert < 5$		
Muon transverse impact parameter significance	$\lvert d_{0,\mu}^{\text{sig}} \rvert < 3$		
Number of jets	$N_{\text{jets}} \geq 2$	$N_{\text{jets}} \geq 2$	$N_{\text{jets}} = 1$
Diagonal cut $3p_{\text{T}}^{\ell_1} + \sum m_{\ell\ell(\prime)}^{\text{OS}}$	< 400 GeV	> 400 GeV	–

Table 5.5: Selection criteria for the Signal Regions. The tight leptons are numbered 1, 2, 3, according to their p_{T} ordering.

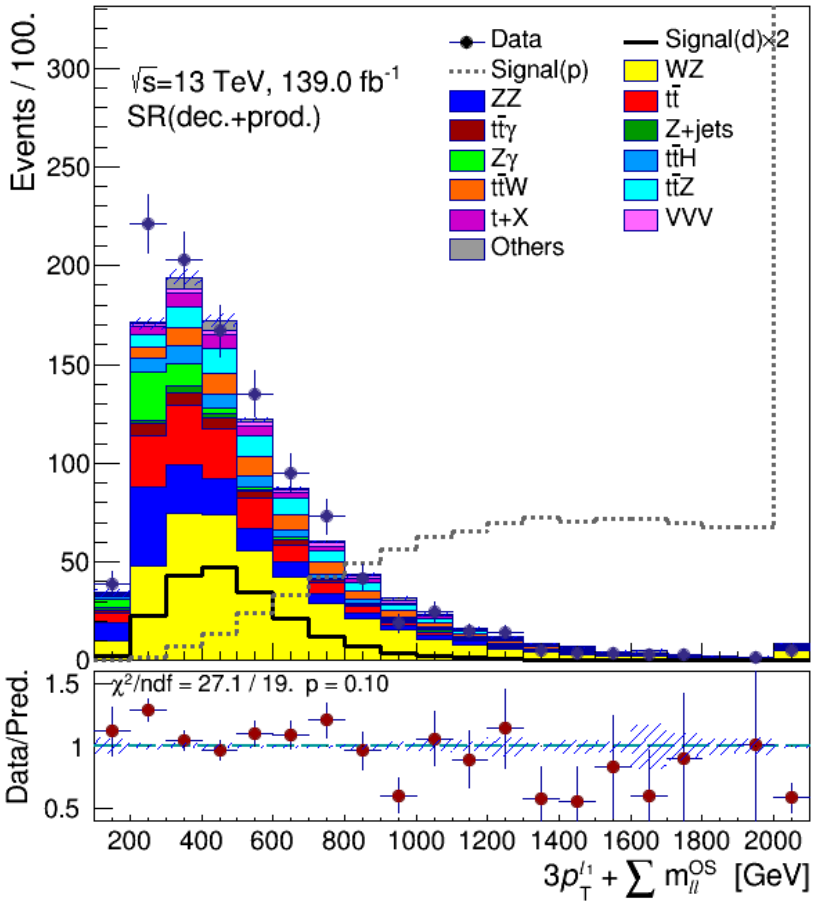


Figure 5.10: Distribution of the variable used for the diagonal cut, employed to isolate the production-signal, in the merged SR-decay and SR-production regions. For values below 400 GeV the production-signal has a minor contribution, while for values above 400 GeV it has the dominant contribution. The Signal/Background ratio for the decay-signal is similar for the regions $<$ 400 GeV (SR-decay) and $>$ 400 GeV (SR-production). The overflow is added in the last bin (the production-signal reaches $\sim$ 1250 events in the last bin).

Name of region	CRWZ	CR$Z\gamma$	CR$Zj\mu$	CRZje	CR$t\bar{t}\mu$	CR$t\bar{t}e$
Description	**Value**					
Number of leptons	$N_{\text{lep}}^{\text{tight}} = 3$	$N_{\text{lep}} = 3$				
Lepton flavours	–	$\ell_1 = \ell_2$			$\ell_1 \neq \ell_2$	
		$\ell_3 = e$	$\ell_3 = \mu$	$\ell_3 = e$	$\ell_3 = \mu$	$\ell_3 = e$
$E_{\text{T}}^{\text{miss}}$	> 35 GeV	–	–		–	
Third leading lepton p_{T}	> 15 GeV	–	< 15 GeV		–	–
Number of b-jets	≤ 1	$= 0$	≤ 1	≤ 1	≤ 2	≤ 2
$\lvert m_{\ell\ell}^{\text{OSSF}} - m_Z \rvert$	< 10 GeV	–	< 10 GeV		> 10 GeV	
Trilepton inv. mass $m_{3\ell}$	–	$\in [65, 95]$ GeV	–	–	–	–
Electron isolation	PLVTight	None	$\ell_{1,2}$ PLVTight, ℓ_3 None			
Electron ID	TightLH	TightLH	$\ell_{1,2}$ TightLH, ℓ_3 LooseAndBLayerLH			
Electron ambiguity type	0	$\neq 0, \geq 1e$	0	0	0	0
Muon isolation	PLVTight	PLVLoose	$\ell_{1,2}$ PLVTight, ℓ_3 PLVLoose			
Muon ID	Medium	Medium	Medium			
Electron transverse impact parameter significance	$\lvert d_{0,e}^{\text{sig}} \rvert < 5$		$\lvert d_{0,e}^{\text{sig}} \rvert < 5$ for $\ell_{1,2}$, ≥ 5 for ℓ_3			
Muon transverse impact parameter significance	$\lvert d_{0,\mu}^{\text{sig}} \rvert < 3$		$\lvert d_{0,\mu}^{\text{sig}} \rvert < 3$ for $\ell_{1,2}$, ≥ 3 for ℓ_3			
Diagonal cut $3p_T^{\ell_1} + \sum m_{\ell\ell^{(\prime)}}^{\text{OS}}$	–	–	–	–	< 400 GeV	

Table 5.6: Selection criteria for the Control Regions. In CR$Z\gamma$, the ambiguity type must not be 0 for at least one electron. In case there is a second OSSF dilepton present in the event (in regions containing events with equal lepton flavours, i.e. 3μ or $3e$), then the dilepton entering the Z-mass requirement is the one with invariant mass closest to m_Z. Leptons are numbered 1, 2, 3, according to their p_{T} ordering.

Name of region	VRlow$\ell_3 p_{\text{T}}$	VRμ	VRe
Description	**Value**		
Number of leptons	$N_{\text{lep}}^{\text{tight}} = 3$		
Lepton flavours	$2\mu 1e, 2e1\mu$	3μ	$3e$
Third leading lepton p_{T}	≤ 15 GeV	–	–
Number of b-jets	$N_b \leq 1$		
$m_{\ell\ell}^{\text{OSSF}}$ closest to Z-mass	$\lvert m_{\ell\ell}^{\text{OSSF}} - m_Z \rvert > 10$ GeV		
Trilepton inv. mass $m_{3\ell}$	–		
Electron isolation	PLVTight		
Electron ID	TightLH		
Electron ambiguity type	0 (electron-like)		
Muon isolation	PLVTight		
Muon ID	Medium		
Electron transverse impact parameter significance	$\lvert d_{0,e}^{\text{sig}} \rvert < 5$		
Muon transverse impact parameter significance	$\lvert d_{0,\mu}^{\text{sig}} \rvert < 3$		
Diagonal cut $3p_T^{\ell_1} + \sum m_{\ell\ell^{(\prime)}}^{\text{OS}}$	< 400 GeV		

Table 5.7: Selection criteria for the Validation Regions.

Process	ℓ^{NP}	**SR-dec.**	**SR-prod.**	**SR-1jet**
WZ		132.3 ± 1.2	286.1 ± 1.6	472.1 ± 4.4
ZZ		73.9 ± 2.0	53.6 ± 2.1	231.8 ± 4.1
$t\bar{t}$	HFμ	36.7 ± 1.2	39.9 ± 1.2	63.1 ± 1.5
	HFe	22.3 ± 0.9	22.1 ± 0.9	38.4 ± 1.2
	other	1.9 ± 0.3	2.2 ± 0.3	3.5 ± 0.4
	total	60.9 ± 1.5	64.2 ± 1.6	105.1 ± 2.0
$t\bar{t}\gamma$		14.0 ± 0.7	17.9 ± 0.8	13.2 ± 0.7
$Z+$jets	HFμ	1.1 ± 0.6	1.2 ± 1.4	15.1 ± 5.1
	HFe	5.6 ± 2.6	4.2 ± 1.1	8.1 ± 4.6
	other	0.2 ± 0.3	0.5 ± 0.4	6.5 ± 3.6
	total	6.9 ± 2.7	5.9 ± 1.9	29.7 ± 7.8
$Z\gamma$	$\gamma \to e^+e^-$	39.0 ± 3.2	6.7 ± 1.1	97.6 ± 5.3
	other	0.6 ± 0.4	1.9 ± 1.1	5.1 ± 2.4
	total	39.6 ± 3.2	8.6 ± 1.6	102.7 ± 5.8
$t\bar{t}H$		17.1 ± 0.1	23.4 ± 0.1	3.7 ± 0.0
$t\bar{t}W$		15.9 ± 0.3	51.4 ± 0.5	22.3 ± 0.3
$t\bar{t}Z$		18.0 ± 0.3	51.5 ± 0.5	8.6 ± 0.2
$t+X$		11.2 ± 0.7	22.8 ± 1.0	23.4 ± 1.3
VVV		3.3 ± 0.1	14.9 ± 0.2	18.7 ± 0.3
Other		7.0 ± 1.8	11.8 ± 1.9	16.1 ± 2.8
Total BKG		400.4 ± 5.4	614.6 ± 4.6	1048.4 ± 12.0
Signal (decay)		34.1 ± 0.4	66.8 ± 0.6	19.4 ± 0.3
Signal (prod.)		9.0 ± 0.5	2164.9 ± 8.2	1030.8 ± 5.7
Total SIG		43.1 ± 0.7	2231.7 ± 8.2	1050.2 ± 5.7
SIG/BKG		(10.8 ± 0.2)%	(365 ± 3)%	(100 ± 1)%
Total MC		443.2 ± 5.5	2844.1 ± 9.4	2097.6 ± 13.3
Data		463 ± 22	611 ± 25	1238 ± 35
Data/total BKG		1.16 ± 0.06	1.00 ± 0.04	1.18 ± 0.04

Table 5.8: Pre-fit event yields in the Signal Regions. The signal is normalised by setting the Wilson coefficients equal to 0.3. The ℓ^{NP} column shows the type of fake or non-prompt lepton present in the process. HF$e/\mu$ processes include events with one e/μ from a heavy-flavour hadron decay. Uncertainties shown are statistical only.

Process	ℓ^{NP}	**CRWZ**	**CR**$Z\gamma$
WZ		7032.4 ± 14.2	1.5 ± 0.2
ZZ		532.1 ± 3.3	5.3 ± 0.8
$t\bar{t}$	HFμ	31.5 ± 1.1	0.1 ± 0.1
	HFe	15.0 ± 0.7	3.9 ± 0.4
	other	1.4 ± 0.2	1.1 ± 0.2
	total	47.9 ± 1.3	5.2 ± 0.4
$t\bar{t}\gamma$		6.5 ± 0.5	0.5 ± 0.1
$Z+$jets	HFμ	83.1 ± 9.4	0.0 ± 0.0
	HFe	32.1 ± 7.6	2.6 ± 2.0
	other	29.8 ± 6.0	7.5 ± 5.3
	total	145.0 ± 13.5	10.1 ± 5.7
$Z\gamma$	$\gamma \to e^+e^-$	17.8 ± 1.9	555.6 ± 16.0
	other	3.7 ± 1.4	3.0 ± 0.8
	total	21.5 ± 2.4	558.6 ± 16.0
$t\bar{t}H$		10.0 ± 0.1	0.0 ± 0.0
$t\bar{t}W$		14.3 ± 0.3	0.0 ± 0.0
$t\bar{t}Z$		344.5 ± 1.2	0.0 ± 0.0
$t+X$		344.6 ± 1.9	0.7 ± 0.3
VVV		31.6 ± 0.2	0.0 ± 0.0
Other		66.1 ± 5.2	3.9 ± 0.5
Total BKG		8596.9 ± 20.9	585.9 ± 17.0
Signal (decay)		8.4 ± 0.2	0.0 ± 0.0
Signal (prod.)		67.6 ± 1.4	0.0 ± 0.0
Total SIG		76.0 ± 1.5	0.0 ± 0.0
SIG/BKG		(0.91 ± 0.01)%	(0.007 ± 0.004)%
Total MC		8672.4 ± 20.9	585.9 ± 17.0
Data		7906.0 ± 88.9	784.0 ± 28.0
Data/MC		0.91 ± 0.01	1.34 ± 0.06

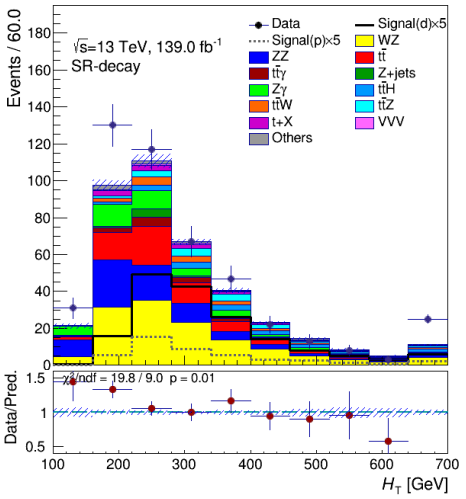
Table 5.9: Pre- fit event yields in the Control Regions targeting WZ and $Z\gamma$ processes. The ℓ^{NP} column shows the type of fake or non-prompt lepton present in the process. HF$e/\mu$ processes include events with one e/μ from a heavy-flavour hadron decay. Uncertainties shown are statistical only. The signal is normalised by setting the Wilson coefficients to 0.3.

Process	ℓ^{NP}	**CR**$Z\mathbf{j}\mu$	**CR**$Z\mathbf{j}e$	**CR**$t\bar{t}\mu$	**CR**$t\bar{t}e$
WZ		30.5 ± 1.0	11.4 ± 0.5	7.2 ± 0.5	3.8 ± 0.3
ZZ		15.6 ± 0.6	6.4 ± 0.3	1.8 ± 0.1	0.8 ± 0.1
$t\bar{t}$	HFμ	50.1 ± 1.4	0.0 ± 0.0	163.6 ± 2.5	0.6 ± 0.1
	HFe	0.0 ± 0.0	187.1 ± 2.7	0.5 ± 0.1	1674.6 ± 7.9
	other	1.7 ± 0.3	9.0 ± 0.6	6.4 ± 0.5	62.5 ± 1.5
	total	51.9 ± 1.4	196.2 ± 2.7	170.5 ± 2.5	1737.7 ± 8.1
$t\bar{t}\gamma$		0.6 ± 0.1	2.2 ± 0.3	2.8 ± 0.3	19.2 ± 0.9
$Z+$jets	HFμ	707.7 ± 37.0	0.0 ± 0.0	1.8 ± 1.2	0.0 ± 0.0
	HFe	0.0 ± 0.0	1402.8 ± 45.1	1.4 ± 1.4	2.1 ± 1.2
	other	59.8 ± 11.2	96.8 ± 12.6	0.0 ± 0.0	0.0 ± 0.0
	total	767.5 ± 38.7	1499.6 ± 46.8	3.1 ± 1.8	2.1 ± 1.2
$Z\gamma$	$\gamma \to e^+e^-$	0.0 ± 0.0	2.4 ± 1.1	2.8 ± 1.7	0.0 ± 0.0
	other	15.2 ± 4.4	21.8 ± 4.5	0.0 ± 0.0	0.0 ± 0.0
	total	15.2 ± 4.4	24.3 ± 4.7	2.8 ± 1.7	0.0 ± 0.0
$t\bar{t}H$		0.1 ± 0.0	0.2 ± 0.0	1.2 ± 0.0	2.3 ± 0.0
$t\bar{t}W$		0.1 ± 0.0	0.2 ± 0.0	1.0 ± 0.1	2.5 ± 0.1
$t\bar{t}Z$		1.7 ± 0.1	2.9 ± 0.1	0.9 ± 0.1	2.3 ± 0.2
$t+X$		6.5 ± 0.8	9.6 ± 1.1	14.6 ± 1.4	69.8 ± 3.0
VVV		0.1 ± 0.0	0.0 ± 0.0	0.2 ± 0.0	0.2 ± 0.0
Other		10.6 ± 0.9	26.9 ± 1.4	1.1 ± 0.5	1.4 ± 0.2
Total BKG		900.3 ± 39.0	1779.9 ± 47.2	207.3 ± 3.8	1841.9 ± 8.8
Signal (decay)		0.1 ± 0.0	0.1 ± 0.0	1.2 ± 0.1	1.1 ± 0.1
Signal (prod.)		0.8 ± 0.2	0.5 ± 0.1	1.3 ± 0.2	1.0 ± 0.2
Total SIG		0.9 ± 0.2	0.6 ± 0.1	2.5 ± 0.2	2.0 ± 0.2
SIG/BKG		(0.10 ± 0.02)%	(0.034 ± 0.008)%	(1.2 ± 0.1)%	(0.11 ± 0.01)%
Total MC		901.2 ± 39.0	1780.5 ± 47.2	209.8 ± 3.9	1843.9 ± 8.8
Data		948.0 ± 30.8	1815.0 ± 42.6	202.0 ± 14.2	1855.0 ± 43.1
Data/MC		1.05 ± 0.06	1.02 ± 0.04	0.96 ± 0.07	1.01 ± 0.02

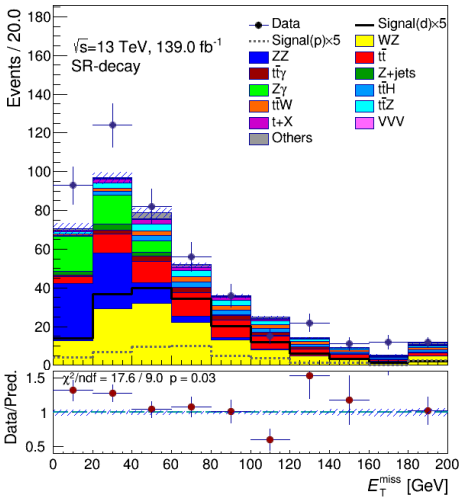
Table 5.10: Pre-fit event yields in the Control Regions targeting Z+jets and $t\bar{t}$ processes. The ℓ^{NP} column marks the type of fake or non-prompt lepton present in the process. HF$e/\mu$ processes include events with one e/μ from a heavy-flavour hadron decay. Uncertainties shown are statistical only. The signal is normalised by setting the Wilson coefficients to 0.3.

Process	ℓ^{NP}	**VRlow**$\ell_3 p_{\mathbf{T}}$	**VR**μ	**VR**e
WZ		126.6 ± 2.0	298.8 ± 3.5	94.6 ± 1.6
ZZ		184.2 ± 3.6	346.7 ± 5.6	122.6 ± 3.2
$t\bar{t}$	HFμ	119.7 ± 2.1	73.9 ± 1.7	0.0 ± 0.0
	HFe	67.0 ± 1.5	0.0 ± 0.0	20.0 ± 0.8
	other	13.6 ± 0.7	3.0 ± 0.3	1.2 ± 0.2
	total	200.3 ± 2.7	76.9 ± 1.7	21.3 ± 0.9
$t\bar{t}\gamma$		7.4 ± 0.5	2.3 ± 0.3	2.8 ± 0.3
$Z+$jets	HFμ	24.7 ± 8.5	67.2 ± 13.1	0.0 ± 0.0
	HFe	44.5 ± 9.5	0.0 ± 0.0	21.1 ± 7.1
	other	3.9 ± 5.5	2.8 ± 5.8	1.8 ± 3.5
	total	73.1 ± 13.9	70.1 ± 14.4	22.9 ± 8.0
$Z\gamma$	$\gamma \to e^+e^-$	100.6 ± 9.6	0.0 ± 0.0	85.3 ± 7.2
	other	7.0 ± 2.2	11.2 ± 3.2	5.2 ± 1.8
	total	107.6 ± 9.9	11.2 ± 3.2	90.6 ± 7.4
$t\bar{t}H$		5.8 ± 0.1	5.6 ± 0.1	2.2 ± 0.0
$t\bar{t}W$		5.4 ± 0.2	5.7 ± 0.2	2.2 ± 0.1
$t\bar{t}Z$		6.6 ± 0.2	12.2 ± 0.2	4.6 ± 0.1
$t+X$		22.3 ± 1.6	14.6 ± 0.9	4.8 ± 0.5
VVV		1.8 ± 0.1	2.1 ± 0.1	0.8 ± 0.1
Other		8.9 ± 1.8	5.4 ± 1.4	3.3 ± 1.4
Total BKG		750.3 ± 17.9	851.8 ± 16.3	372.6 ± 11.6
Signal (decay)		8.9 ± 0.2	2.0 ± 0.1	0.9 ± 0.1
Signal (prod.)		3.4 ± 0.3	1.8 ± 0.2	0.8 ± 0.2
Total SIG		12.2 ± 0.4	3.8 ± 0.3	1.7 ± 0.2
SIG/BKG		(1.63 ± 0.06)%	(0.45 ± 0.03)%	(0.46 ± 0.05)%
Total MC		762.4 ± 17.9	855.5 ± 16.3	374.3 ± 11.6
Data		917.0 ± 30.3	851.0 ± 29.2	515.0 ± 22.7
Data/MC		1.20 ± 0.05	0.99 ± 0.04	1.38 ± 0.07

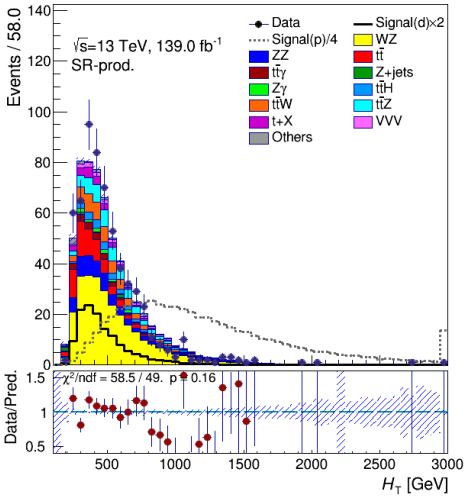
Table 5.11: Pre-fit event yields in the Validation Regions. The ℓ^{NP} column shows the type of fake or non-prompt lepton present in the process. HF$e/\mu$ processes include events with one e/μ from a heavy-flavour hadron decay. Uncertainties are statistical only. The signal is normalised by setting the Wilson coefficients to 0.3.



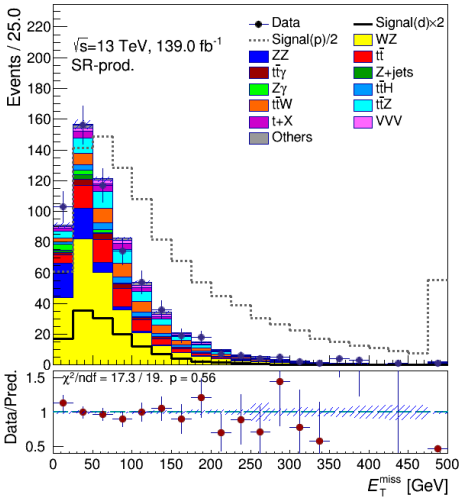
(a) Scalar sum of p_{T} of jets and leptons.



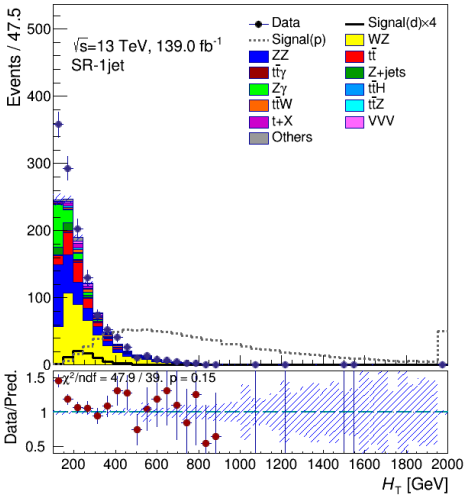
(b) Missing transverse momentum.



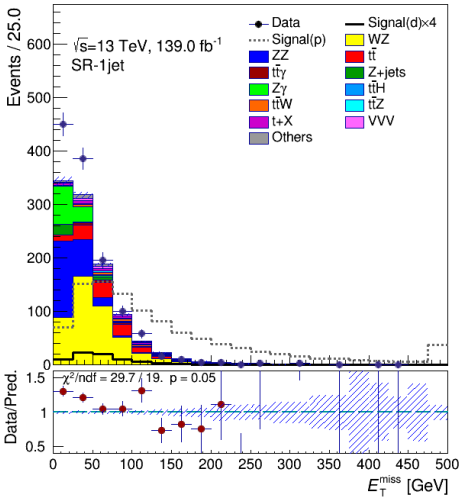
(c) Scalar sum of p_{T} of jets and leptons.



(d) Missing transverse momentum.

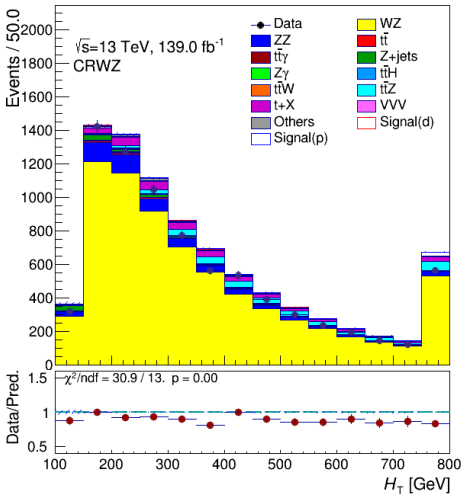


(e) Scalar sum of p_{T} of jets and leptons.

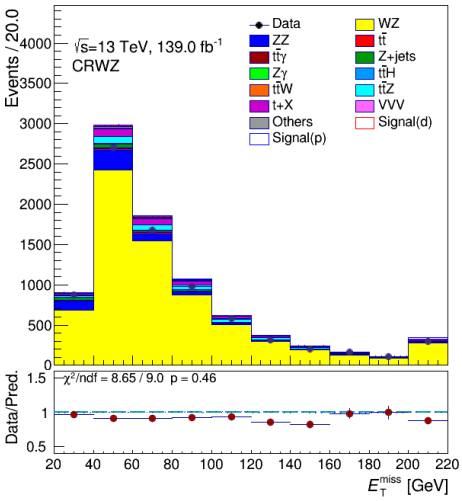


(f) Missing transverse momentum.

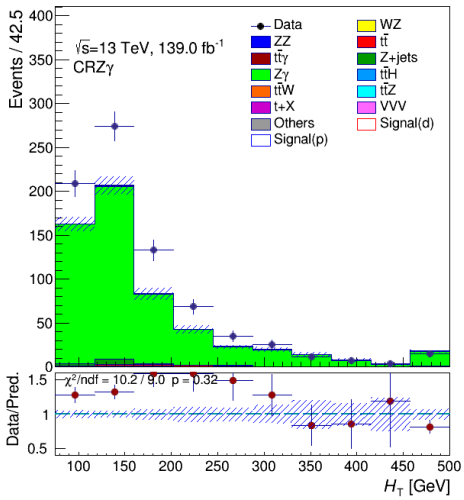
Figure 5.11: Kinematic distributions of the MC predictions vs data in the Signal Regions. Statistical uncertainties are shown for both MC (hatched area) and data (error bars). Overflow is added in the last bin. Signal has been scaled to fit in the vertical range. The ratio of data over MC prediction is given in the lower panel, along with its statistical uncertainty as error bars. The hatched area in the lower panel represents the total MC uncertainty over the total MC prediction.



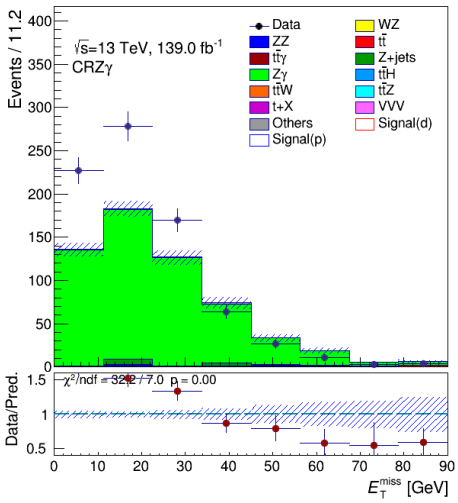
(a) Scalar sum of p_{T} of jets and leptons.



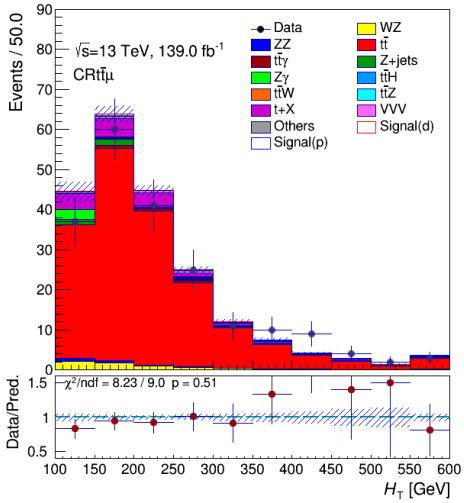
(b) Missing transverse momentum.



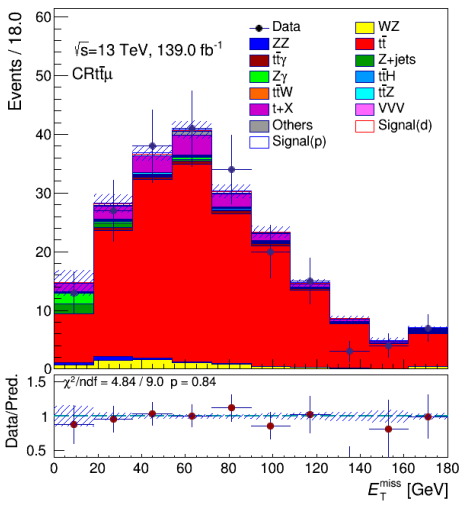
(c) Scalar sum of p_{T} of jets and leptons.



(d) Missing transverse momentum.

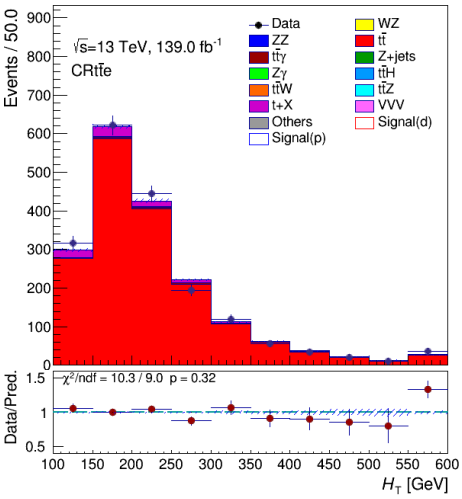


(e) Scalar sum of p_{T} of jets and leptons.

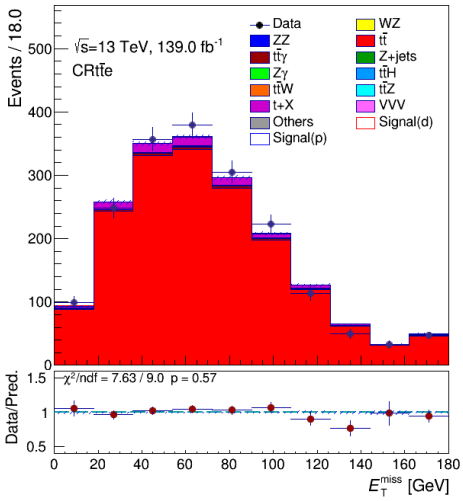


(f) Missing transverse momentum.

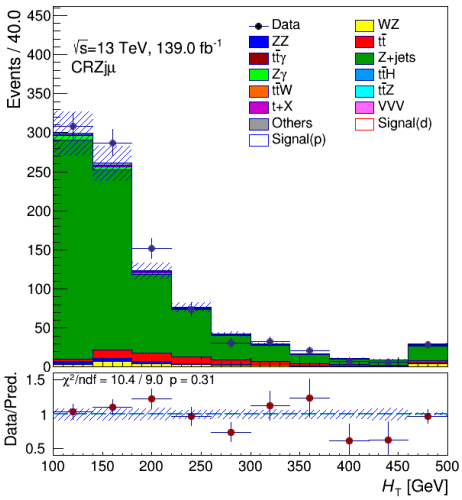
Figure 5.12: Kinematic distributions of the MC predictions vs data in the Control Regions CRWZ, CR$Z\gamma$ and CR$t\bar{t}\mu$. Statistical uncertainties are shown for both MC (hatched area) and data (error bars). Overflow is added in the last bin. The ratio of data over MC prediction is given in the lower panel, along with its statistical uncertainty as error bars. The hatched area in the lower panel represents the total MC uncertainty over the total MC prediction.



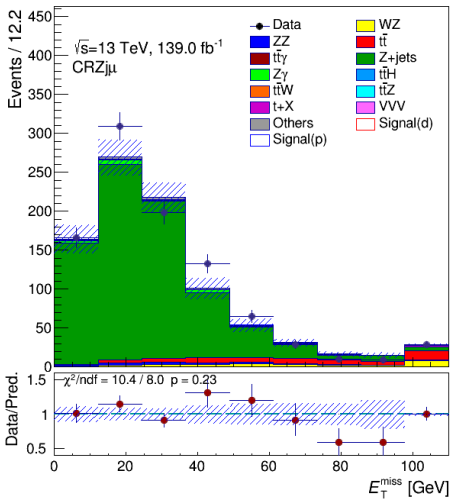
(a) Scalar sum of p_{T} of jets and leptons.



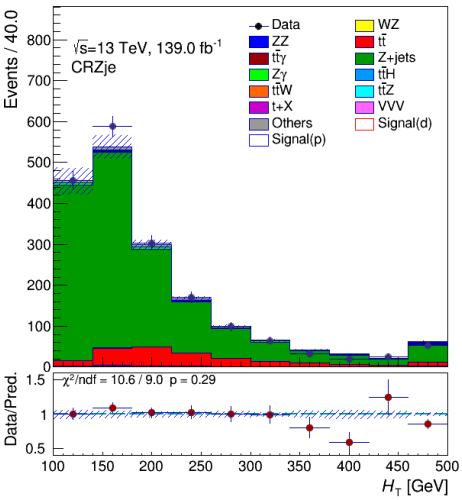
(b) Missing transverse momentum.



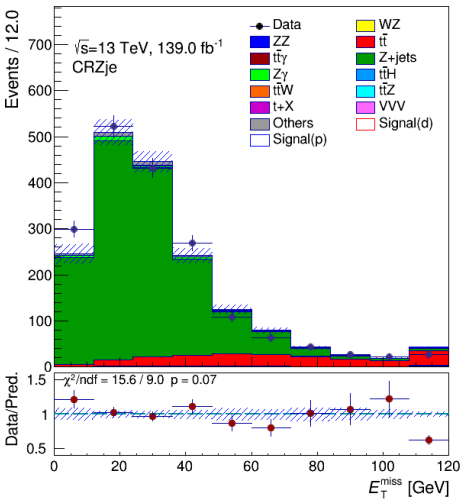
(c) Scalar sum of p_{T} of jets and leptons.



(d) Missing transverse momentum.

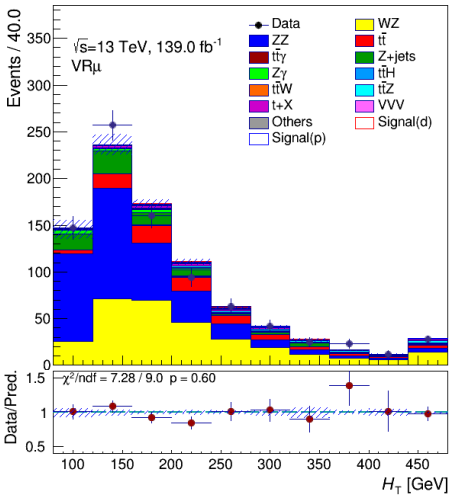


(e) Scalar sum of p_{T} of jets and leptons.

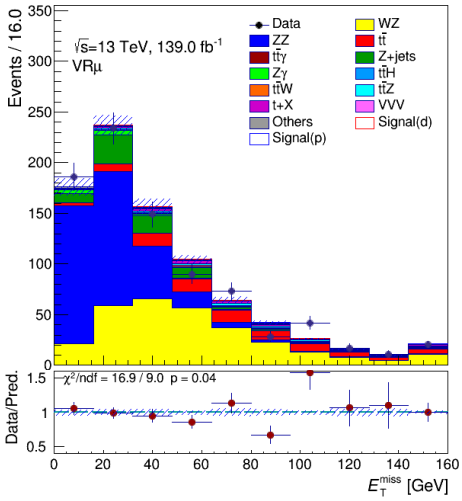


(f) Missing transverse momentum.

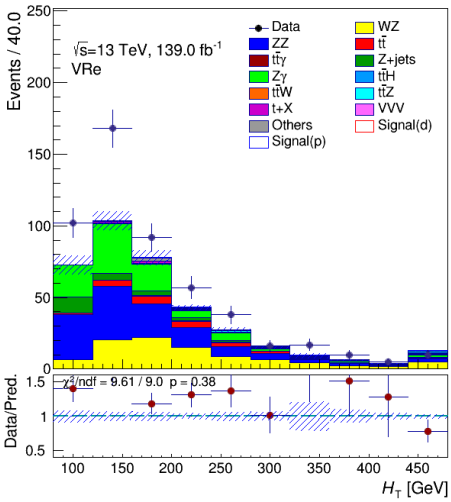
Figure 5.13: Kinematic distributions of the MC predictions vs data in the Control Regions CR$t\bar{t}e$, CRZjμ and CRZje. Statistical uncertainties are shown for both MC (hatched area) and data (error bars). Overflow is added in the last bin. The ratio of data over MC prediction is given in the lower panel, along with its statistical uncertainty as error bars. The hatched area in the lower panel represents the total MC uncertainty over the total MC prediction.



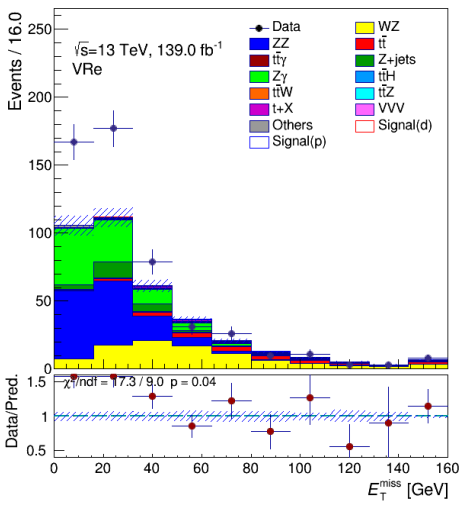
(a) Scalar sum of p_{T} of jets and leptons.



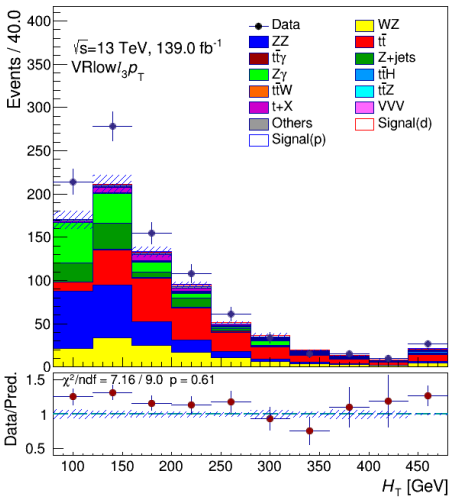
(b) Missing transverse momentum.



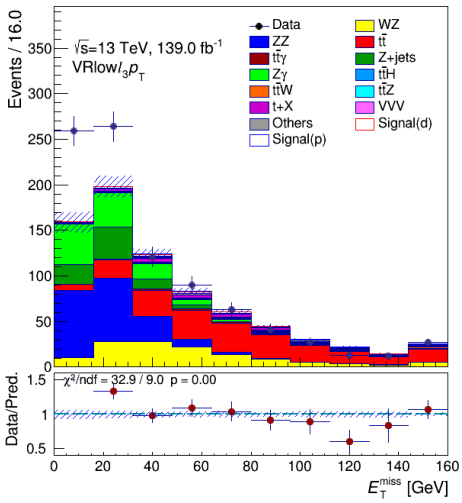
(c) Scalar sum of p_{T} of jets and leptons.



(d) Missing transverse momentum.



(e) Scalar sum of p_{T} of jets and leptons.



(f) Missing transverse momentum.

Figure 5.14: Kinematic distributions of the MC predictions vs data in the Validation Regions. Statistical uncertainties are shown for both MC (hatched area) and data (error bars). Overflow is added in the last bin. The ratio of data over MC prediction is given in the lower panel, along with its statistical uncertainty as error bars. The hatched area in the lower panel represents the total MC uncertainty over the total MC prediction.

5.5 Top quark reconstruction

The reconstruction of top quarks is performed in the SR-decay region only, as it is relevant only for the decay signal, where we can discriminate between two types ("SM" and "CLFV") of top-quark decays (see also section 5.6).

The first step is the reconstruction of the exotically decaying top quark, which we name "CLFV top", in contrast to "SM top", the other top quark of the pair. Each jet, excluding the one carrying the highest `DL1r` b-tagging score, is associated with every opposite-sign different-flavour lepton pair present in the event, to give a top-quark candidate. Among these candidates the one having an invariant mass closest to the top-quark mass (172.5 GeV) is the "CLFV top" and its decay products are excluded from further steps.

The reconstruction of the "SM top" starts with the reconstruction of the W boson decay products: while the charged lepton can be unambiguously identified as the one remaining, the neutrino momentum has to be calculated. Following Ref. [234], we assign to the x and y components of the neutrino four-momentum the respective components of the missing momentum of the event, $\mathbf{p}_{\mathrm{T}}^{\mathrm{miss}}$, and we assume the W to be on-shell. Thus we require the invariant mass of the charged lepton and the neutrino to equal the W mass and we neglect the charged lepton and neutrino masses:

$$M_W^2 = (E_\ell + \sqrt{(E_{\mathrm{T}}^{\mathrm{miss}})^2 + p_{z,\nu}^2})^2 - (\mathbf{p}_{\mathrm{T},\ell} + \mathbf{p}_{\mathrm{T}}^{\mathrm{miss}})^2 - (p_{z,\ell} + p_{z,\nu})^2 \,. \tag{5.3}$$

Solving eq. (5.3) for $p_{z,\nu}$ we obtain eq. (5.4):

$$p_{z,\nu}^{\pm} = \frac{\mu \cdot p_{z,\ell}}{p_{\mathrm{T},\ell}^2} \pm \sqrt{\frac{\mu^2 \cdot p_{z,\ell}^2}{p_{\mathrm{T},\ell}^4} - \frac{E_\ell^2 \cdot (E_{\mathrm{T}}^{\mathrm{miss}})^2 - \mu^2}{p_{\mathrm{T},\ell}^2}} \,, \tag{5.4}$$

where μ is given by:

$$\mu = \frac{M_W^2}{2} + \mathbf{p}_{\mathrm{T},\ell} \cdot \mathbf{p}_{\mathrm{T}}^{\mathrm{miss}} \,. \tag{5.5}$$

If the radicand of eq. (5.4) is positive the lower value of $|p_{z,\nu}|$ is taken as the solution [234]. However, due to finite $E_{\mathrm{T}}^{\mathrm{miss}}$ resolution or the presence of other neutrinos in the event, the radicand can be negative. This corresponds analytically to the situation in which the W transverse invariant mass $m_{\mathrm{T},W}$, calculated from the transverse momenta of the associated charged lepton and the neutrino, is larger than the W mass itself. In this case, while still respecting eq. (5.3), the radicand is set to zero by asking that $m_{\mathrm{T},W} = M_W$, which leads to a quadratic relationship between $p_{y,\nu}$ and $p_{x,\nu}$ with two solutions $p_{y_{1,2},\nu}(p_{x,\nu})$. In order to get a unique solution for the neutrino momentum we minimise the difference between the neutrino transverse momentum components and the measured $E_{\mathrm{T}}^{\mathrm{miss}}$, as defined by eq. (5.6).

$$\delta_{1\,2}(p_{x,\nu}) = \sqrt{(p_{x\,\nu} - E_{\mathrm{T}\,x}^{\mathrm{miss}})^2 + (p_{y_{1,2},\nu}(p_{x,\nu}) - E_{\mathrm{T}\,y}^{\mathrm{miss}})^2} \tag{5.6}$$

The smaller δ value is kept. The $p_{x,\nu}$ minimising it and the corresponding $p_{y,\nu}$ become the transverse momentum components, while its longitudinal momentum is still given by eq. (5.3). The charged lepton

and the neutrino are then combined with the jet having the highest `DL1r` score to give the "SM top".

In fig. 5.15, the distributions of some of the reconstructed variables, later used for the multivariate analysis, are given for both signal and background processes. More variable distributions can be found in Appendix A.

5.6 Multivariate analysis

5.6.1 The boosted decision tree

Introduction

The multivariate analysis algorithms comprise methods for supervised machine learning, where the problem of predicting the *output* of a process based on a set of *input variables* is tackled [235]. If the output is a quantitative variable, i.e. it takes continuous values, then we refer to the prediction task as a *regression* problem; if the output is qualitative, i.e. it is a discrete variable, then we are dealing with a *classification* problem. The input variables may also be called *independent variables*, or *features*, the latter term preferred in pattern recognition problems. The *output variables* are often called the *responses*, or classically the *dependent variables*.

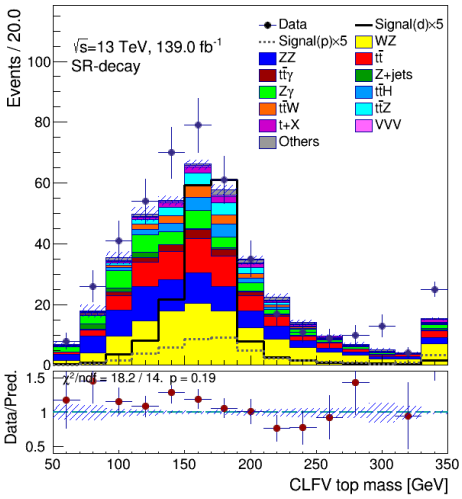
In terms of notation, usually the input variable set is denoted by X, which in general is a vector of p input variables. Each input variable is in turn a vector of N measurements, or observations. $\vec{X}$ will refer to the $N \times p$ set of inputs (N observations for each of the p input variables). A quantitative output is usually denoted by Y, and a qualitative one by G. The respective values obtained by the prediction model are typically denoted by $\hat{f}(X)$ or $\hat{G}(X)$ respectively. Observed values are usually written in lowercase, therefore the ith observed value of X is written as x_i and the respective observed value of Y as y_i.

In order to get a measure of the performance of the model, it is typical to split the available sample of (x_i, y_i) observed, or measured, values into a *training sample* and a *test sample*, the latter not taking part in the derivation of the model but used independently to evaluate the predicted outputs after the prediction model has been obtained. Methods which partition the available data into different training and test samples, are broadly referred to as *cross-validation* or *cross-training* methods.

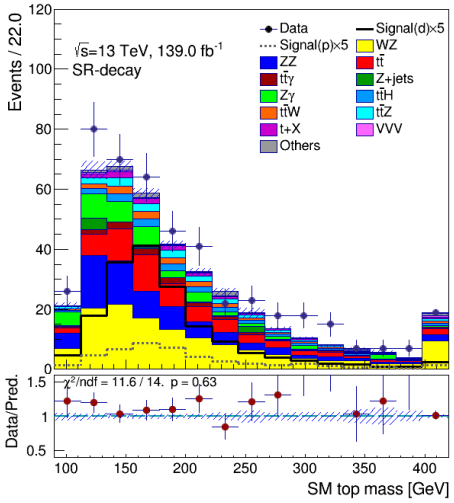
A measure of the deviation of the prediction compared to the real output values (or of the *misclassification* of the outputs) is the *loss function*. Typical choices for the loss function are:

$$L(Y, \hat{f}(X)) \begin{cases} (Y - \hat{f}(X))^2 & \text{squared error} \\ |Y - \hat{f}(X)| & \text{absolute error .} \end{cases} \tag{5.7}$$

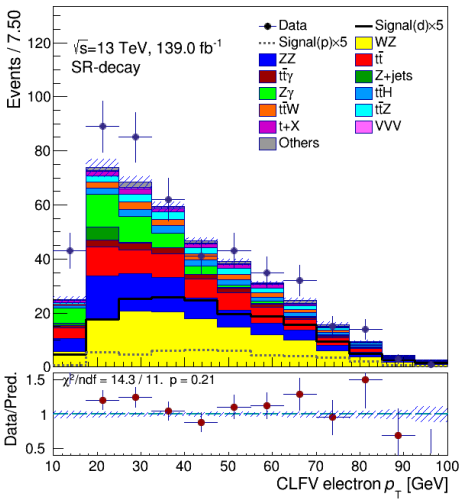
By averaging the loss function over the test sample we obtain the test, or *generalisation*, error, while the average over the training sample is the *training error*. An increasingly complex model, which uses the training data more and is able to adapt to more complicated underlying structures, results in a decrease of the training error. At some intermediate model complexity, a minimum test error can be achieved. If the test error is large compared to the training error, then it is possible that the model has become over-sensitive to the particular fluctuations of the training sample, or is "over-fitted" to the



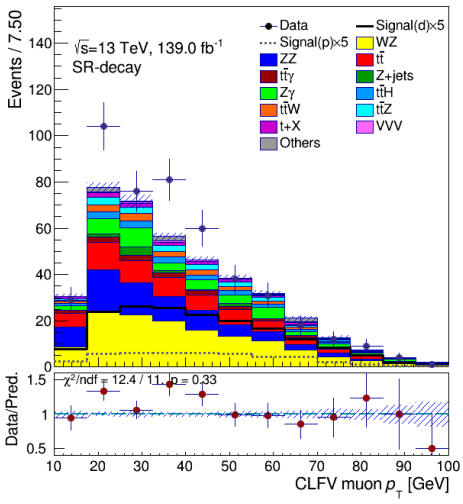
(a) CLFV top invariant mass



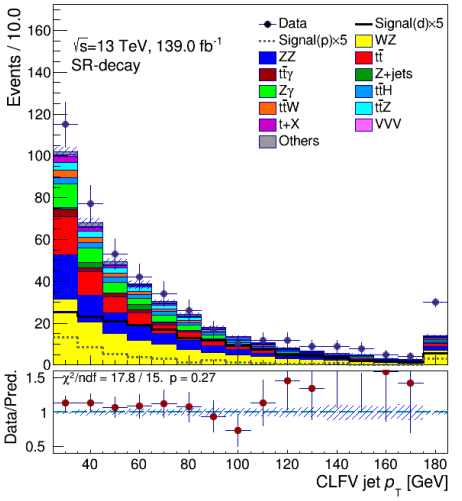
(b) SM top invariant mass



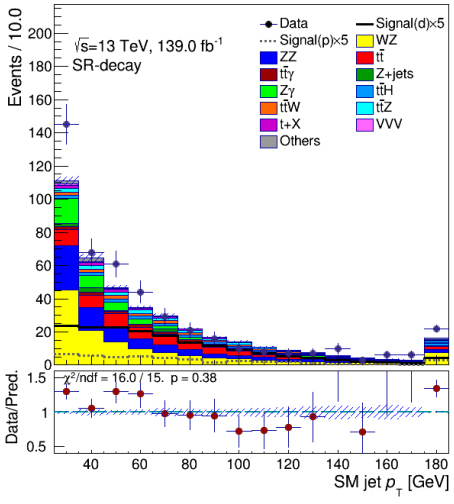
(c) CLFV electron p_{T}



(d) CLFV muon p_{T}



(e) CLFV light jet p_{T}



(f) SM b-jet p_{T}

Figure 5.15: MC prediction vs data in the SR-decay region for kinematic variables that are fed to the BDT. Statistical uncertainties are shown for both MC (etched area) and data (error bars). Overflow/underflow is added in the last/first bin. The ratio of data over MC prediction is given in the lower panel, along with its statistical uncertainty as error bars. The etched area in the lower panel represents the total MC uncertainty over the total MC prediction.

training sample, a situation called *overtraining*.

For a classification problem with qualitative, or "categorical", response $G \in 1, 2, ...K$, the loss function has to be modified, a typical choice being

$$L(G, \hat{G}(X)) = I(G \neq \hat{G}(X)) \quad (\text{"0-1" loss}) \,, \tag{5.8}$$

where I(condition) denotes the indicator function, equal to 1 when the condition is true and 0 otherwise.

In a simple scenario, a classification problem may often be a *two-class problem*. In such a case we may treat the outcome as a quantitative output. Then, for example, the predictions $\hat{y}$ could lie in $[0, 1]$, i.e. we predict the probability of each class, and we could then assign the class label to $\hat{G}$ according to whether $\hat{y} > 0.5$. This approach can also be extended to the general, K-class, problem.

Decision trees

In the following we focus on classification trees, even though the concept can be extended to regression problems [235, 236]. A decision tree utilises a sequence of *binary splits* of the data in order to categorise events into different classes. Each split corresponds to a *node* of the tree.

The growth of a decision tree is based on the process of defining the splitting criteria for each node. Starting from a root node, the initial split results in two subsets of training events that each go through the same algorithm for determining the next splitting iteration. This procedure is repeated until the whole tree is built, as illustrated in fig. 5.16. At each node, a decision is made concerning the variable and the corresponding "rectangular" cut value that provides the best separation between signal and background.

In the common scenario of a two-class problem, e.g. when $G \in \{0, 1\}$, corresponding to signal (S) and background (B) events, the splitting is usually based on a measure of the impurity of the node, called the *Gini index* and defined by

$$p(1 - p) \,, \tag{5.9}$$

where $p = S/(S + B)$ is the node purity. The training of the tree selects the variable and cut value that maximise the difference between the parent node index and the sum of the indices of the two daughter nodes, weighted by their relative fractions of events.

The depth of the tree is defined by the number of splits until the terminal nodes or *leaves* of the tree are reached. The maximum depth of a tree is usually pre-defined; otherwise, the terminal nodes are reached when no improvement in purity is seen.

Adaptive boosting

The boosting of a decision tree extends the tree concept from one tree to a collection of several trees which form a *forest*. The trees are all derived from the same training sample by a reweighting process applied to the events, and are finally combined into a single classifier, which is given by a weighted average of the individual decision trees [115, 235, 236].

An individual decision tree is a *weak* classifier, one whose error rate is only slightly better than random guessing. The purpose of boosting is to sequentially apply the weak classification algorithm to

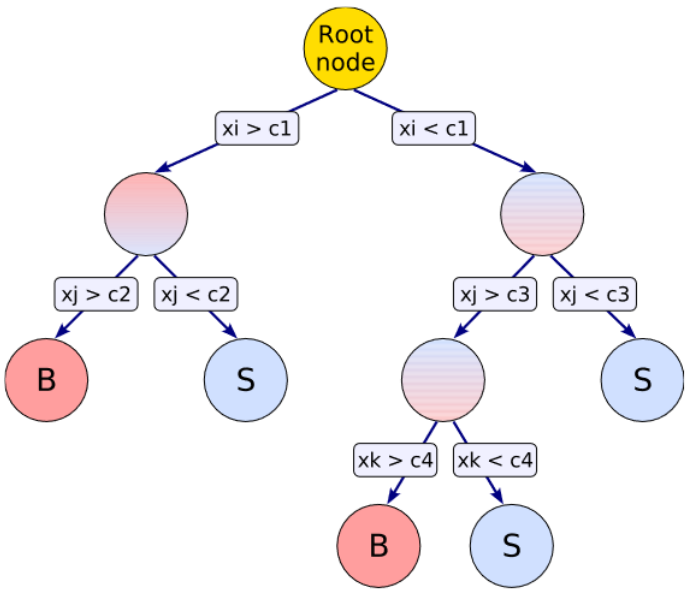


Figure 5.16: Sketch of a decision tree. The discriminating variables x_i are used at each node to perform binary splits of the data. Each split uses the variable that at this node gives the best separation between signal and background. Taken from [236].

repeatedly modified versions of the data, producing a sequence of weak classifiers.

The most popular boosting algorithm is the so-called AdaBoost [237].The basic idea is that events which were misclassified during the training of a decision tree are given a higher event weight in the training of the following tree.

A *ranking* of the BDT input variables is obtained by counting how often the variables are used to split decision tree nodes, and by weighting each split occurrence by the squared separation gain achieved and by the number of events in the node.

5.6.2 Implementation

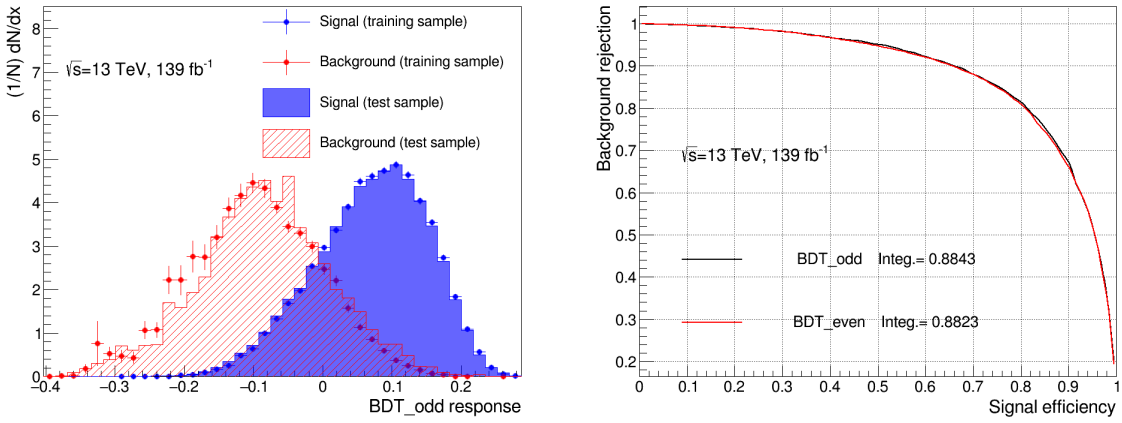
As described in the previous section (section 5.5), the reconstruction of the two ("CLFV" and "SM") top quarks of the decay signal final state, gives us discriminating variables, such as the CLFV top invariant mass, which provide a good separation of the decay signal from the background. The reconstruction relies on matching a light-quark jet with an OSSF lepton pair together with an m_t constraint. In the production signal case, the OSSF dilepton arises from the CLFV vertex without an accompanying jet, and not as a result of a top quark decay, hence the production signal does not benefit from such a procedure.

The BDT is implemented in the TMVA [236] package embedded in the ROOT [238] framework. The boosting is carried out using the AdaBoost algorithm, with a learning rate[16] `AdaBoostBeta` equal to 0.4, on a forest of 1500 trees, each with a maximum depth of three. The granularity in the variable space, which is scanned per split, `nCuts`, is chosen to be equal to 40. The set of input variables supplied to the BDT is given in table 5.12.

[16]The learning rate modifier β acts exponentially on the boost weight.

The BDTs are trained with MC simulated data in the SR-decay region. In order to control over-training and exploit all the MC data available, two BDTs have been trained. One is labelled "BDT even" and is trained over the even-numbered entries of the signal and background samples, and tested over the odd-numbered ones, while for "BDT odd" the opposite is done. In this manner the BDT output is tested and evaluated using a set of events which is disjoint from the events on which it is trained.

The output of the BDT algorithm is expressed as a probability distribution for each output class (signal S and background B). In that way, a plot like fig. 5.17a is obtained, where the probability of S or B is plotted at various BDT response (or score) values. The compatibility between the training and test distributions suggests that no over-training has occurred. The performance of a BDT is usually measured by the *receiver operating characteristic* (ROC) curve integral. The ROC curve is constructed by evaluating the signal efficiency and the background rejection (i.e. one minus the background efficiency) at various thresholds of the BDT score. The signal efficiency corresponds to the number of signal events above the threshold divided by the total number of signal events (the background efficiency is calculated similarly). ROC integral values closer to one suggest a model with better separation power. The ROC integral for the "BDT odd" ("BDT even") training is 0.884 (0.882), as shown in fig. 5.17b.



(a) (b)

Figure 5.17: (a) Over-training check for the odd-trained BDT. The filled area represents the BDT score distribution of the test sample, while the data points show the distribution of the training sample. A good agreement between the two distributions verifies that the discriminator was not biased by peculiarities in the training sample, due to statistical fluctuations. (b) The ROC curve (background rejection versus signal efficiency) for the cross-trained BDTs.

The set of input variables provided to the BDT has been selected in relation to the discriminator performance. Initially a set of "core" variables exhibiting clear separation (CLFV top mass, leptons' p_{T}, b-jet multiplicity) or physical motivation (H_{T}, $E_{\mathrm{T}}^{\mathrm{miss}}$, $m_{\ell\ell}$ of OS or SS same-flavour lepton pairs) were identified. To further optimise the BDT performance, several other sets of variables were added iteratively. Variables carrying no useful information according to the "method-specific" [236] ranking, and not increasing the ROC curve integral, were dropped. This leads to the selection of the variables listed in table 5.12. The list is sorted according to the method-specific ranking provided by TMVA.

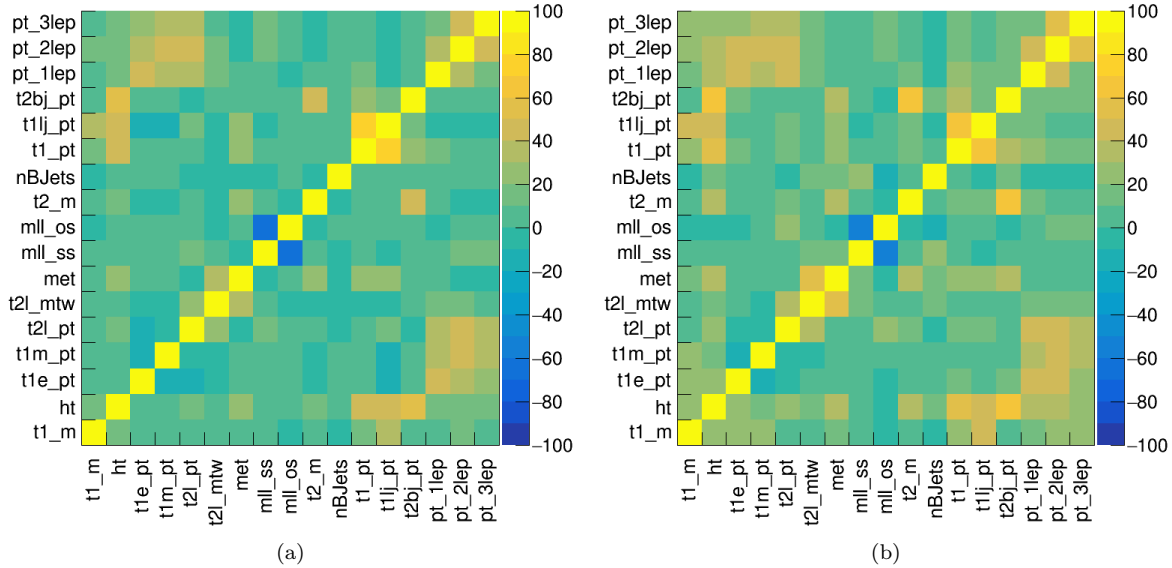
Figure 5.18 displays the linear correlations between the input variables.

Different numbers of trees have been tested: smaller forests yield a worse performance, in terms of the ROC integral, while increasing the number of trees leads to a *saturation*, beyond which no improvement in performance is obtained. Other options were also considered, including alternative boosting algorithms, different values of `nCuts` and `AdaBoostBeta` and the tree depth, before concluding on the final configuration.

The resulting BDT score distributions for signal and background processes, as well as for data, in the SR-decay region are shown in fig. 5.19. The BDT score will be used as a variable in the SR-decay region as part of the fitting process, described in section 5.9.

Variable	Label	Importance (10^{-2})
`t1_m`	CLFV top mass	7.387
`met`	Missing transverse momentum	7.056
`mll_os`	OSSF lepton-pair invariant mass	6.790
`ht`	Scalar sum of the p_{T} of all jets and leptons in the event	6.640
`t2l_mtw`	W transverse mass associated to the SM top lepton	6.547
`mll_ss`	SSSF lepton pair invariant mass	5.985
`t2bj_pt`	p_{T} of the b-jet associated with the SM top	5.727
`t1lj_pt`	p_{T} of the light jet associated with the CLFV top	5.727
`t1_pt`	p_{T} of the CLFV top	5.524
`t2l_pt`	p_{T} of the lepton associated to the SM decay	5.519
`pt_3lep`	p_{T} of the third leading lepton	5.502
`nBJets`	b-jet multiplicity (DL1r 77% working point)	5.425
`pt_1lep`	p_{T} of the leading lepton	5.421
`t2_m`	SM top invariant mass	5.400
`t1e_pt`	p_{T} of the electron associated to the CLFV decay	5.324
`pt_2lep`	p_{T} of the sub-leading lepton	5.235
`t1m_pt`	p_{T} of the muon associated to the CLFV decay	4.859

Table 5.12: The input variables used in the multivariate BDT analysis, listed according to the method-specific ranking provided by TMVA (for the "BDT odd" training).



(a) (b)

Figure 5.18: Correlation of the input variables for the signal (a) and background (b) classes.

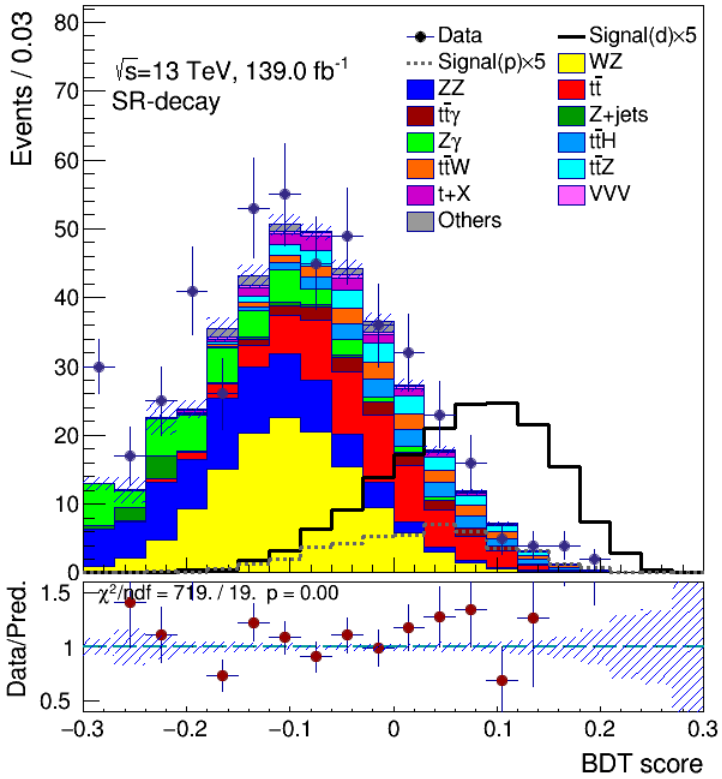


Figure 5.19: BDT discriminant distribution of signal, background and data in the SR-decay region. Only statistical uncertainties are included. The lower panel shows the ratio of data over total background (bkg.) prediction, the hatched area representing the total bkg. uncertainty divided by the bkg. prediction.

5.7 Systematic Uncertainties

This section describes the systematic uncertainties considered in the analysis. The systematic uncertainties are included in the full likelihood as nuisance parameters (NPs) as described in section 5.8 and section 5.9. The uncertainties are separated into experimental-related and theory-related categories. Additional uncertainties, e.g. the uncertainty originating from having only a finite number of simulated events, are also summarised below.

5.7.1 Experimental uncertainties

The following experimental uncertainties are considered:

- **Luminosity and pileup uncertainties**: The combined uncertainty on the integrated luminosity collected in the years 2015–2018 is 1.7% [239]. The simulated samples are reweighted by a scale factor to account for differences in the pileup distributions between MC and data. An uncertainty on this "pileup scale factor" is obtained by rescaling the pileup distribution around its nominal value, within its uncertainty.

- **Electron and muon uncertainties**:

  - The efficiencies of lepton reconstruction, identification and isolation and the trigger efficiency differ between data and MC. Scale factors, as measured using tag-and-probe methods in Z, W and J/ψ decays [189, 191, 201], are applied to MC events to correct for the discrepancies, as described in section 5.3.10. The uncertainties from the tag-and-probe measurements are propagated to the uncertainties on the lepton scale factors and are considered in the analysis. The uncertainties are propagated as a single NP for each of the identification, isolation, reconstruction and trigger efficiencies.

  - The electron (muon) energy (momentum) scale and resolution are compared between simulation and data using $Z \to \ell^+\ell^-$ events [201], and also $W \to \ell\nu$ events [189, 240]. Observed discrepancies between data and MC are corrected, and the residual uncertainties from the data measurements are considered in the analysis. Uncertainties on the momentum scale and the resolution of muons originating from the inner detector and the muon spectrometer are considered and varied separately.

- **Jet uncertainties**:

  - Scale factors are used to account for differences between the Jet Vertex Tagging (JVT) efficiencies in simulation and data. Events with $Z \to \mu^+\mu^-$+jets are used to compare the efficiencies [241]. Another systematic uncertainty is derived by comparing different MC generators for the Z+jets simulation. The statistical uncertainty of the derived scale factors is also taken into account.

  - The Jet Energy Scale (JES) is measured in LHC collision data, simulation and test-beam data. Firstly, a MC-based calibration is applied, then data taken at $\sqrt{s} = 13$ TeV are

used [242] to calibrate the residual uncertainty in JES. Events with a vector boson and additional jets are used to calibrate jets in the central region. Dijet events are exploited to calibrate forward jets against the jets in the central region of the detector. Multi-jet events are used to calibrate high p_{T} jets. Uncertainties related to all the steps are considered, and grouped based on a *Category reduction* scheme. This results in a set of 31 nuisance parameters, each with an up and a down variation, which can have different jet p_{T} and η dependencies.

– The Jet Energy Resolution (JER) is measured in dijet events from the width of the dijet asymmetry in various p_{T} and η bins, using similar techniques to those presented in [243]. The fits performed result in about 120 NPs, which are then reduced by a *PseudoData* reduction scheme into twelve NPs and a residual data-to-MC term, giving a total of thirteen NPs for JER.

- **Flavour-tagging uncertainties**: Correction factors are applied to the simulated event samples to compensate for differences between data and simulation in the b-tagging efficiencies for b, c and light-quark jets. The correction for b-jets is derived from $t\bar{t}$ events with final states containing two leptons [208]. The corrections are consistent with unity with uncertainties at the level of a few percent over most of the jet p_{T} range. The scale factor for the c-mistag-efficiency is derived using $t\bar{t}$ events with single-lepton final states [244]. Finally, the mistagging efficiency for light flavours is measured in dijet events [245]. The uncertainties from each measurement are propagated to the analysis via nine NPs for b-efficiency, four NPs for c-efficiency and four NPs for light-flavour efficiency.

- $\boldsymbol{E_{\mathrm{T}}^{\mathrm{miss}}}$ **uncertainties**: The systematic variations of the $E_{\mathrm{T}}^{\mathrm{miss}}$ terms representing the individual physics objects are propagated to the calculation of $E_{\mathrm{T}}^{\mathrm{miss}}$, and are thus included within the relevant object uncertainties. Additionally, scale and resolution variations of the $E_{\mathrm{T}}^{\mathrm{miss}}$ track soft-term are considered in the analysis.

5.7.2 Theoretical uncertainties

For the signal samples, a number of theoretical (or "modelling") uncertainties are evaluated:

- Independent variations of the renormalisation and factorisation scales (μ_R and μ_F, respectively) are considered to account for missing higher order corrections in the matrix element calculation. Events are reweighted using internal (provided within the MC sample itself, i.e. not requiring a separately generated sample) weights in which each of μ_R and μ_F are varied by factors of two around their original values.

- Uncertainties on the PDF set used in the generation of the sample are estimated using a set of 100 variations. The standard deviation of this set is used to define a symmetrised uncertainty on the signal sample.

- To estimate the uncertainty relating to the choice of parton shower and hadronisation model, the nominal signal samples generated with MadGraph5_aMC@NLO2.9.5+Pythia 8.306 are

compared to samples generated with MadGraph5_aMC@NLO2.9.5+Herwig 7.1.6.

The largest "non-prompt" background comes from dileptonic $t\bar{t}$ events with a non-prompt lepton. The nominal $t\bar{t}$ sample is produced using Powheg Box v2 interfaced to Pythia 8 for parton showering. For $t\bar{t}$, the following modelling uncertainties are considered.

- To estimate the size of the uncertainty due to the matching of the NLO matrix element calculation to the parton shower, a sample generated by Powheg + Herwig 7.1.3 is compared to a sample generated by MC@NLO + Herwig 7.1.3, which uses a different matching method. This comparison is chosen to avoid being sensitive to differences in the matrix element corrections (MEC) applied to the top decay, which are considered in the Powheg to Pythia matching scheme used in the nominal sample but not in the MC@NLO scheme.
- To estimate the uncertainty relating to the choice of parton shower and hadronisation model, the Powheg +Pythia 8 sample is compared to a Powheg +Herwig 7.1.3 sample.
- Scale variations are considered in the same manner as for the signal sample, using weights in the nominal sample.
- The uncertainties due to the PDF are calculated using the PDF4LHC15 prescription [246], resulting in 30 eigenvariations, each assigned a nuisance parameter.
- The uncertainty arising from the value of α_s is estimated using the nominal PDF set evaluated with $\alpha_s(M_Z) = 0.117$ and $\alpha_s(M_Z) = 0.119$.
- The uncertainty arising from the modelling of ISR is evaluated using internal weights corresponding to variations of α_s in the initial-state shower.
- The uncertainty arising from the choice of the h_{damp} parameter within Powheg is evaluated using an alternative Powheg +Pythia 8 $t\bar{t}$ sample in which $h_{damp} = 3.0m_{\text{top}}$ (the default setting is $h_{damp} = 1.5m_{\text{top}}$).
- Modelling of FSR is evaluated using internal weights corresponding to variations of μ_R for QCD emissions in the final-state shower.

For $t\bar{t}\gamma$, tW, $t\bar{t}V$ and $t\bar{t}H$ processes, scale variations and PDF uncertainties are considered in the same manner as in the $t\bar{t}$ sample. Additionally, for the tW single-top process:

- The uncertainties arising from the modelling of ISR and FSR are evaluated as in the $t\bar{t}$ sample.
- The nominal tW sample uses the diagram removal scheme [161] to account for overlap between tW and $t\bar{t}$ diagrams. In addition to removing doubly-resonant $t\bar{t}$-like diagrams, this method removes interference terms between tW and $t\bar{t}$. To understand the sensitivity to the interference, the nominal sample is compared to a Powheg +Pythia 8 tW sample using the diagram subtraction (DS) scheme (see also section 5.2.4.2). The difference between these two samples is taken as an additional uncertainty intended to encode the size of the interference effects.

For the $t\bar{t}W$ process:

- A combined NLO matching and parton shower and hadronisation model uncertainty is evaluated separately for the QCD and the dedicated EW samples. For the QCD component, the nominal Sherpa 2.2.10 sample is compared to a MadGraph5_aMC@NLO + Pythia8 sample with an FxFx merging scheme [247]. To evaluate the generator uncertainty on the subleading (NLO) electroweak (EW) $t\bar{t}W$ production [248] component, the Sherpa 2.2.10 EW sample is compared to a LO MadGraph5_aMC@NLO + Pythia8 sample which also models the subleading EW effects.

For the $t\bar{t}Z$ process:

- The ISR variations are evaluated using alternative samples generated with variations of α_s in the initial-state shower.
- Parton shower variations are calculated using samples produced with Herwig for parton shower and hadronisation modelling. A set of Herwig 7.0.4 $t\bar{t}Z$ samples are used and are compared to a set of MadGraph5_aMC@NLO + Pythia 8 samples produced with the same version of MadGraph5_aMC@NLO.

For the $t\bar{t}H$ process:

- ISR and FSR variations are evaluated in the same manner as for the $t\bar{t}$ process.
- Parton shower variations are evaluated using samples produced with Herwig for parton shower and hadronisation modelling.

For the modelling of $Z+\text{jets}(+\gamma)$, WZ and ZZ processes, the μ_{R} and μ_{F} scale variations are calculated in a slightly different manner than for $t\bar{t}$. Seven variations of μ_R and μ_F by factors of two are evaluated, considering cases where these are changed both independently and simultaneously but excluding variations of μ_R and μ_F in opposite directions. Then the maximum up and down variation is selected for the scale uncertainty. Additionally, for WZ and ZZ processes:

- The PDF set uncertainty is evaluated using the NNPDF3.0 NNLO PDF set.
- The uncertainty arising from the value of α_s is estimated using the nominal PDF set evaluated with $\alpha_s(M_Z) = 0.117$ and $\alpha_s(M_Z) = 0.119$.

5.7.3 Cross-section uncertainties

Normalisation-only cross-section uncertainties are included on many background processes, as detailed in table 5.13. Cross-section uncertainties of 12% and 10%, respectively, are applied to the $t\bar{t}Z$ and $t\bar{t}H$ processes, based on calculations reported in [168]. The cross-sections for these processes are consistent with more recent calculations documented in [249], for which the theoretical uncertainties are seen to be reduced. A 50% cross-section uncertainty is also applied to the $t\bar{t}W$ process to account for significant differences between the theoretical predictions and the measured cross-sections in data [247, 250].

For the single-top process, the cross-section uncertainty on tW production is 5.3% [251], while for tZq it is 14% [252–254]. Thus in regions where the inclusive $t + X$ content is dominated by tZq, we

apply a 14% uncertainty, while for regions where tW is the dominant component (specifically, control regions targeting events with two prompt and one non-prompt lepton), we apply a conservative 6% uncertainty.

For the $t\bar{t}\gamma$ process, a cross-section uncertainty of 15% is applied [255]. A 50% normalisation uncertainty is applied to backgrounds in the "Other SM" category, for which the largest contributions are $t\bar{t}WW$, $t\bar{t}t\bar{t}$, $ll\nu\nu jj$ and $t\bar{t}t$. The backgrounds that are normalised by the fit (WZ, $Z+\text{jets}(+\gamma)$, $t\bar{t}$), are not assigned cross-section uncertainties[17].

Process	Normalisation uncertainty
$t\bar{t}\gamma$	15% [255], [254]
$t\bar{t}H$	10% [250]
$t\bar{t}Z$	12% [250]
$t\bar{t}W$	50% [252], [256]
tW	5.3% [251]
tZq	14% [252]
ZZ	30% [257]
VVV	50% [258]
Other SM	50%

Table 5.13: Summary of the cross-section (normalisation) uncertainties considered for various processes. These uncertainties come from theoretical calculations or are taken as conservative uncertainties on small background contributions. The references point to the sources of these values where applicable, and are often links to other ATLAS analyses, which have been published using the same uncertainty values, rather than the original studies motivating the values.

5.7.4 Other uncertainties

The statistical uncertainty originating from the finite number of simulated events in the MC samples is considered in the analysis as a systematic uncertainty (the so called *gamma* nuisance parameters), as described in section 5.8. Uncertainties relating to the data-driven fake lepton estimations are also considered. The non-prompt and fake e/μ estimation is performed simultaneously with the limit-setting procedure so that all of the above sources of uncertainty are accounted for.

[17] A 40% normalisation uncertainty is applied to the marginal component of $Z+\text{jets}$ and $t\bar{t}$ backgrounds which do not originate from heavy flavour hadronic decays (from photon conversion in the $Z\gamma$ case). This fraction of events is not subject to a normalisation factor obtained from the fit, as described in section 5.9.

5.8 The profile likelihood ratio test statistic

In this section we describe the maximum likelihood fit to the data, used to obtain a limit on the signal cross-section. As is often the case in the analysis of particle physics data, one has to have a measure (or "test statistic") of the confidence in rejecting a "signal-plus-background" hypothesis (or "alternative" hypothesis) against a "background-only" hypothesis (or "null hypothesis"). Hence we start by describing the common frequentist approach[18] to the likelihood ratio as a test statistic.

5.8.1 The likelihood function

Starting with a set of statistically independent measurements $\boldsymbol{x}$ of a variable x (such as the mass of a particle or another kinematic variable), it is typical to bin the measurements into a histogram $\boldsymbol{n} = n_1, ..., n_N$ of N bins.

The expectation value count for each bin can be expressed as

$$E[n_i] = \mu S_i + B_i \ . \tag{5.10}$$

Here μ is a signal strength parameter (also called parameter of interest (POI)), defined such that $\mu = 0$ corresponds to the absence of a signal (background-only hypothesis) and $\mu = 1$ gives the signal rate S_i expected from the nominal signal hypothesis. B_i represents the total background count predicted by our model.

While the histograms with simulated signal and background events effectively describe a distribution, the statistical model for such a binned distribution[19] is a series of counting experiments that can be described with a Poisson distribution ($\mathcal{P}$) for each bin. The binned likelihood function (LF) is the product of the Poisson probabilities for all bins [259]

$$\mathcal{L}(\boldsymbol{n}|\mu) = \prod_{i=1}^{N} \mathcal{P}(n_i|\mu S_i + B_i) \ . \tag{5.11}$$

Our knowledge of the expectation values for the signal and background, S_i and B_i, (i.e our assumed model) is subject to systematic uncertainties, which are encoded into the LF as *nuisance parameters* (NPs). The S_i and B_i become functions of the NPs by including an additive term (e.g. $B_i + \theta\Delta B_i$, where θ is the NP and ΔB_i is the associated 1σ systematic variation) for each NP to represent the $\pm 1\sigma$ effect of the systematic uncertainty on the nominal prediction.

Since we already have a prediction of the size of the systematic effect, we can incorporate constraints into the LF for each NP. This is done by adding *Gaussian* constraint terms (also known as *penalty* terms), usually with a mean value of 0 and variance equal to 1[20] [261]. The constraint terms reflect our knowledge of the systematic uncertainties, and prevent the respective nuisance parameters from floating

[18] The frequentist probability ("probability of the data given an assumed hypothesis") is used in particle physics more often than the Bayes approach ("probability of a hypothesis, given the data") which needs a *prior* probability assigned to the hypotheses.

[19] The histogram which models a sample is often referred to as a "template".

[20] Alternatively, NPs may enter the LF as multiplicative terms to the expectation values S_i and B_i; then the Gaussian mean value will be 1, and the variance σ^2 will represent the effect of the systematic variation on the nominal expectation values S_i and B_i [260].

to larger values, or even allow them to diminish during the fit to the data (in contrast to "free-floating parameters" mentioned further in this section).

The LF then becomes

$$\mathcal{L}(\boldsymbol{n}|\mu,\boldsymbol{\theta}) = \prod_{i=1}^{N} \mathcal{P}(n_i|\mu S_i(\boldsymbol{\theta}) + B_i(\boldsymbol{\theta})) \times \prod_{j\in syst} \mathcal{G}(\theta_j|0,1) \ , \tag{5.12}$$

where $\boldsymbol{\theta}$ is the set of nuisance parameters associated to the systematic uncertainties, and $\mathcal{G}(\theta_j|\theta_j^0, \Delta\theta_j)$ is the Gaussian distribution of the parameter θ_j, with mean θ_j^0 and standard deviation $\Delta\theta_j$.

Each NP may affect the nominal prediction of the background and the signal by shape (i.e. a different impact per bin), by normalisation or both. Each of these components has a different interpolation strategy. The shape component uses a bin-by-bin linear interpolation, while the normalisation component uses exponential interpolation [262, 263]. The exponential interpolation is motivated by preventing the normalisation from having unphysical negative values. Exponential interpolation with a Gaussian constraint term leads to a log-normal distribution for the normalisation component.

The background B is commonly split into multiple categories, describing different background sources. Certain backgrounds are not expected to be modelled well, and so are subject to large systematic uncertainties. For such cases, subsidiary measurements can be made in *control regions* (CRs) in which certain sources of background are enhanced, while the expected signal is minimal (in contrast to *signal regions* (SRs) which are designed for maximum signal acceptance and background suppression), providing further information on the background normalisation and sometimes also on its shape. Similar to the definition of the signal strength, *normalisation factors* (NFs) on specific backgrounds, which are *free-floating* (i.e. without a constraint term), can be applied as additional nuisance parameters, giving

$$B_l(k,\boldsymbol{\theta}) = kB_l(\boldsymbol{\theta}) \ , \tag{5.13}$$

where the index l denotes the background that we chose to normalise by use of the NF k. By the implementation of CRs in the LF (essentially by incrementing the bin count in the LF formula[21]) we may constrain the NPs related to the backgrounds targeted.

Nuisance parameters that describe the effect of limited Monte Carlo statistics (it is common to have simulated samples of reweighted events) can be incorporated into the LF. These nuisance parameters are also called *gamma* factors, or simply *gammas*. A single "MC-stat" NP is assigned to the total prediction $(S_i + B_i)$ in each bin, and it enters the LF with a Poisson constraint term [262].

The "best fit values" for the POI μ and the NPs $\boldsymbol{\theta}$, including the NFs, from our measurements $\boldsymbol{x}$, are called the maximum likelihood estimators (MLEs) $\hat{\mu}$, $\hat{\boldsymbol{\theta}}$, and are obtained by finding the global maximum of the LF in the space of M parameters (μ, $\boldsymbol{\theta}$).

Since it is more convenient to work with sums instead of products, as well as finding minima instead of maxima, it is common to work with the negative log-likelihood (NLL) function, $-\ln\mathcal{L}$. Then, the

[21] All regions (SRs and CRs) have to be chosen with no statistical correlations between them (orthogonal) so that we can simply take the likelihood product of all regions.

MLEs for all parameters can be found by setting the derivatives of $-\ln\mathcal{L}$ w.r.t. every parameter equal to zero, i.e.

$$-\frac{\partial \ln\mathcal{L}}{\partial\theta_j} = 0\ ,\ \text{for}\ j = 0, ..., M-1\ , \tag{5.14}$$

where we have considered $\theta_0 = \mu$.

In the *asymptotic limit*, that is when the number of measurements goes to infinity, the MLEs are consistent (converge to their true values), unbiased (their expectation values are equal to the true values) and efficient (their variance $V[\boldsymbol{\theta}]$ is "small") [115]. In the asymptotic limit, the NLL "profiled" as a function of a single parameter has a Gaussian shape. In practice, for finite statistics, a Gaussian function approximates the shape around the minimum. The variance of the MLE for a parameter can then be estimated from the point where the NLL increases by 1/2 from its minimum value.

Alternatively, for finite statistics, and assuming that the Gaussian approximation is valid, the covariance matrix of all MLEs, $V_{ij} = \text{cov}[\hat{\theta}_i, \hat{\theta}_j]$, can be estimated from the inverse of the matrix of the second derivatives of NLL w.r.t. all parameters (Hessian matrix $\mathcal{H}$) evaluated at their best-fit values:

$$V_{ij}^{-1} = \frac{\partial^2 \ln\mathcal{L}(\hat{\theta}_i, \hat{\theta}_j)}{\partial\theta_i \partial\theta_j} = \mathcal{H}\ . \tag{5.15}$$

From the covariance matrix of the MLEs, we can obtain the variance of the MLE of the POI, $\sigma^2 = V_{00}$ (using index 0 for the POI), and of any other NP θ, $\sigma^2 = V_{jj}$.

5.8.2 The test statistic

For a simple analysis with no NPs, defining the likelihood ratio as

$$\frac{\mathcal{L}(\boldsymbol{x}|H_{S+B})}{\mathcal{L}(\boldsymbol{x}|H_B)}\ , \tag{5.16}$$

where H_{S+B} is the "signal-plus-background" ($\mu > 0$) hypothesis tested against a "background-only" ($\mu = 0$) hypothesis, H_B, is a reasonable choice as a test statistic [264].

With an increasing number of NPs, the calculations concerning discovery and limit setting (see next sections 5.8.3 and 5.8.4) become increasingly complicated. A better option, is to replace the likelihood ratio test statistic with the profile likelihood ratio test statistic,

$$\lambda(\mu) = \frac{\mathcal{L}(\mu, \hat{\hat{\boldsymbol{\theta}}})}{\mathcal{L}(\hat{\mu}, \hat{\boldsymbol{\theta}})}\ , \tag{5.17}$$

where $\hat{\hat{\boldsymbol{\theta}}}$ is the *conditional* MLE set of NPs, which maximize the LF for a given value of μ (hence "profiled" as a function of μ). The denominator is the LF at its global maximum ($\hat{\mu}$ and $\hat{\theta}$ are the MLEs of μ and θ). From its definition, λ lies in the range $0 \leq \lambda \leq 1$, with values close to 1 suggesting good agreement between data and the hypothesised value of μ.

In physical theories we are commonly interested in the case where the rate of signal events is non-negative, and thus $\mu \geq 0$. However, it is often convenient to allow $\mu < 0$; $\hat{\mu} < 0$ indicates a deficit

of signal-like events with respect to the expected background and the boundary at $\mu = 0$ complicates the asymptotic distributions. In [259] a constraint on the test statistic itself, rather than a formal $\mu \geq 0$ boundary, is used by replacing the MLEs $\hat{\mu}$ and $\hat{\boldsymbol{\theta}}$ with 0 and the conditional MLE of $\boldsymbol{\theta}$ at $\mu = 0$ respectively, when $\hat{\mu} < 0$:

$$\lambda(\mu) = \begin{cases} \dfrac{\mathcal{L}(\mu, \hat{\hat{\boldsymbol{\theta}}})}{\mathcal{L}(\hat{\mu}, \hat{\boldsymbol{\theta}})} & \hat{\mu} \geq 0 \\ \dfrac{\mathcal{L}(\mu, \hat{\hat{\boldsymbol{\theta}}})}{\mathcal{L}(0, \hat{\hat{\boldsymbol{\theta}}}(0))} & \hat{\mu} < 0 \, . \end{cases} \tag{5.18}$$

The use of the test statistic of eq. (5.18) is implied in the following.

As with the LF, it is more convenient to work with logarithms, so we define the quantity

$$q_\mu = -2ln\lambda(\mu) \, . \tag{5.19}$$

Higher values of q_μ represent increasing incompatibility between the data and a hypothesised signal strength μ.

To quantify the level of disagreement between data and the hypothesis, we compute the p-value,

$$p_\mu = \int_{q_{\mu,obs}}^{\infty} f(q_\mu|\mu) dq_\mu \, , \tag{5.20}$$

where $q_{\mu,obs}$ is the value of the test statistic q_μ observed from the data and $f(q_\mu|\mu)$ is the probability density function (pdf) of q_μ under the assumption of signal strength μ. Small p-values will then lead one to reject the null hypothesis. According to *Wilks theorem* [265], following the *Wald approximation* [266], f follows a χ^2 distribution in the asymptotic, or large sample, limit.

The *significance* corresponding to a p-value can be defined as the number of standard deviations Z at which a Gaussian random variable of zero mean would give a one-sided tail area equal to p. The significance Z is then related to the p-value by

$$Z = \Phi^{-1}(1 - p) \, , \tag{5.21}$$

where Φ^{-1} is the inverse of the cumulative distribution for a unit Gaussian.

5.8.3 Discovery

For discovery, we are interested in rejecting the null hypothesis ($\mu = 0$), and thus we use the test statistic

$$q_0 = \begin{cases} -2ln\lambda(0) & \hat{\mu} \geq 0 \\ 0 & \hat{\mu} < 0 \, . \end{cases} \tag{5.22}$$

Thus, the p-value quantifying the level of compatibility between data and the null hypothesis is

$$p_0 = \int_{q_{0,obs}}^{\infty} f(q_0|0) dq_0 \, . \tag{5.23}$$

The significance $Z\sigma$ of a measurement is computed by using eq. (5.21). The standard discovery convention is 5σ, corresponding to $p_0 = 2.87 \times 10^{-7}$.

5.8.4 Setting limits

For limit setting, we consider the alternative hypothesis of some non-zero value of μ and try to reject it. The test statistic is modified to:

$$q_\mu = \begin{cases} -2ln\lambda(\mu) & \hat{\mu} \leq \mu \\ 0 & \hat{\mu} > \mu \ . \end{cases} \tag{5.24}$$

Data with upward fluctuations in $\hat{\mu}$ are not counted when computing the p-value, because they would be compatible with some larger value of μ.

The p-value is computed using eq. (5.20) for all values of μ, and the set of μ values for which the p-value is greater than or equal to a fixed value 1-CL form a confidence interval for μ, where typically a *confidence level* CL = 95% is taken. The upper end of this interval μ_{up} is the upper limit, i.e. $\mu \leq \mu_{up}$ at 95% CL is the largest value of μ for which the p-value is at least 0.05.

Assuming the asymptotic approximation (Wald [266]), the significance simplifies to

$$Z_\mu = \Phi^{-1}(1 - p_\mu) = \sqrt{q_\mu} \ . \tag{5.25}$$

The CL_s method

In practice, the CL_s method [267], also referred to as the *modified frequentist approach*, is most often used for setting confidence levels instead of the previously mentioned method (often called CL_{sb}), by replacing the p-value with

$$p'_\mu = \frac{p_\mu}{1 - p_b} \ , \tag{5.26}$$

where p_b is the p-value under the background-only ("b-only", i.e. $\mu = 0$) hypothesis. In that way, the p-value is penalised by being divided by $1 - p_b$. If the two distributions $f(q|b)$ and $f(q|s+b)$, illustrated in fig. 5.20, are widely separated, then $1 - p_b$ is close to unity, and the exclusion limit based on CL_s is similar to that obtained from the usual p-value. However, if one has little sensitivity to the signal model, then the two distributions are close, $1 - p_b$ becomes small, and thus p'_μ increases. In that sense, the CL_s technique protects us against excluding signals with very low cross sections, to which the experiment is not sensitive.

5.8.5 Expected significance and upper limit

The sensitivity of an experiment can be characterized by the expected, or more precisely the median, significance and the limit coming from a large number of similar experiments.

The expected upper limit is the upper limit one would expect to obtain if the background-only hypothesis is true. The median can be computed from an ensemble of pseudo-experiments, created by repeating a generation of pseudo-data in the b-only hypothesis, by a randomisation of the nuisance parameters [262], usually referred to as a "toy Monte Carlo" approach.

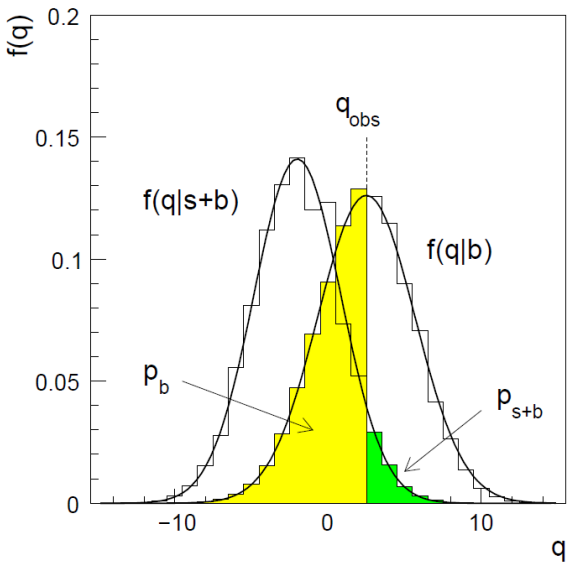


Figure 5.20: Distributions of the test statistic q under the $\mu = 0$ (background-only) and $\mu = 1$ (signal+background) hypotheses. From [259].

The estimation of the median significance and limit may be approximated by replacing the ensemble of simulated datasets by a single representative one, usually referred to as the *Asimov* dataset[22]. The Asimov dataset is such that the data are equal to their expectation values under the assumption of signal strength μ':

$$n_{i,A} = \mu' S_i + B_i \ . \tag{5.27}$$

In practice, it is common to set μ'=0, so that the Asimov dataset is exactly the expected background. For upper limit setting, and under the asymptotic approximation, the median significance for μ reduces to the simple formula:

$$\text{med}[Z_\mu|\mu' = 0] = \sqrt{q_{\mu,A}} \tag{5.28}$$

where $q_{\mu,A}$ is the test statistic obtained from the Asimov dataset.

[22]The name "Asimov" is related to the science fiction short story *Franchise* by Isaac Asimov [268], published in 1955, where the single most representative voter is selected by a super-computer to replace the entire electorate in future United States elections.

5.9 Results

5.9.1 Statistical analysis framework

A binned maximum likelihood fit is used to extract the signal strength of the CLFV top quark production and decay processes. In the absence of an observed signal, exclusion limits can be set on the cross section of these processes and on the branching ratio of the top quark decay to the $e\mu q$ final state. Exclusion limits are extracted at 95% confidence level using the modified frequentist approach under the asymptotic approximation.

The fit is performed by use of the TRExFitter [263] framework (v4.16). TRExFitter builds statistical models in the HistFactory [269] format, in the context of RooFit [270], the framework in which the statistical modelling is expressed, and RooStats [271], which provides additional tools for interval estimation and hypothesis testing (both packages of the Root [272, 273] data analysis framework).

A vital part of the fit is the minimisation of the negative logarithm of the likelihood (NLL). This is done internally by calling Minuit2 [274–276]. The minimisation and estimation of parameters are done in multiple steps. The first step uses the Migrad algorithm to find the minimum of the likelihood function. The second step includes using the Hesse method, which provides a more precise estimate of the uncertainties than Migrad, but reports only symmetric uncertainties, by construction. The last (optional) step runs the Minos algorithm to get the most precise estimate of the uncertainties of the parameters by scanning the likelihood, and in general may result in asymmetric uncertainties [263]. The application of Minos increases the running time of the fit considerably, and, since only minor changes in the parameter errors are obtained, it is only applied to the signal strength parameter.

The minimisation procedure is not only responsible for finding the minimum, but it also estimates the correlation of the input parameters (POI, normalisation factors and systematic uncertainty NPs), by calculating the Hessian matrix of the NLL.

5.9.1.1 Systematic uncertainties treatment

The systematic uncertainties resulting from instrumental effects and from the modelling of signal and background processes in MC simulations, described in section 5.7, are included in the fit as nuisance parameters.

A *pruning* procedure is applied to remove nuisance parameters that have only a small impact. This is done to simplify the fit model, to reduce the computational time needed to run the fit, and to ensure its stability. For each nuisance parameter in the fit, both the shape and normalisation terms are included in the likelihood. Pruning is applied per sample and per region, as well as being applied separately to the shape and normalisation terms of an uncertainty. The pruning threshold is set to 0.3% for both shape and normalisation, so that any systematic uncertainty for a given background which results in a change in the normalisation or shape of that background by less than 0.3% is dropped.

The pruning on the normalisation and shape components is performed as follows. For each region, the $\pm 1\sigma$ distributions (or "up/down" variations) corresponding to a specific systematic uncertainty are

considered. For the normalisation component, if any of the up/down variations result in a change to the overall yield, compared to the nominal yield, which is larger than the threshold, then that normalisation component is kept. For the shape component, each of the up/down variations is first normalised to the nominal sample. Then, if any bin in an up/down variation differs from the corresponding bin in the nominal distribution by more than the threshold, then that shape component is kept. Reducing the pruning threshold from 0.3% to 0.2%, and to 0.1%, has been tested, and found to give the same results, in terms of limit setting. A table exemplifying the pruning of nuisance parameters is shown in fig. 5.21.

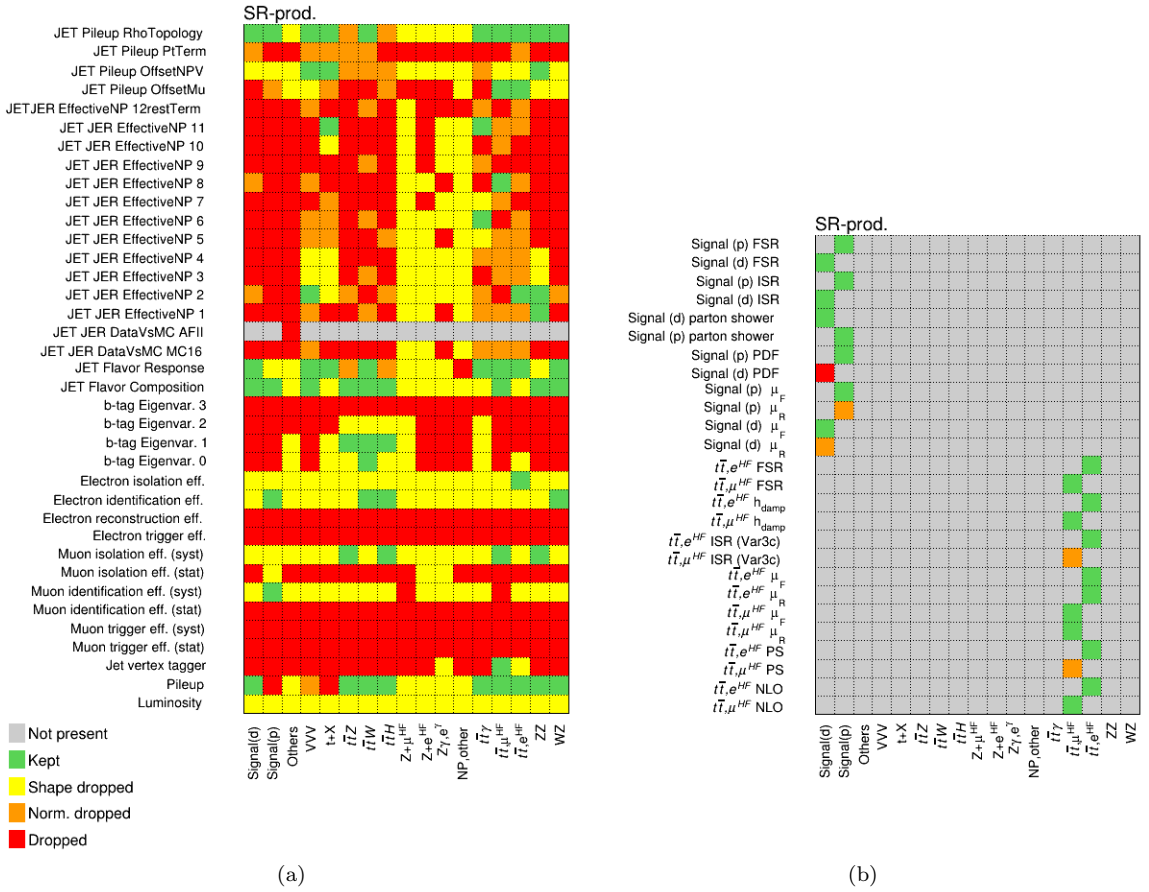
Additionally, a *smoothing* procedure (by use of the relevant TRExFitter tools) is applied to each of the systematic uncertainty variation templates to reduce the effect of any fluctuations resulting from limited MC statistics. Finally, all systematic uncertainties are *symmetrised* using the `TWOSIDED` option in TRExFitter, which calculates half the difference between the up and down variations for each bin and symmetrises them. For uncertainties where only one variation is provided (e.g. when alternative samples are compared to the nominal for an estimate of the PS systematic uncertainty), the downward variation is assumed to be a mirrored version of the upward variation (using the `ONESIDED` option). Plots showing the ± 1 standard deviation (σ) variations of modelling (or "theory") systematic uncertainties (often called "red-blue" plots) affecting $t\bar{t}$, WZ, $Z+\text{jets}(+\gamma)$ and $t\bar{t}W$ processes, are shown in figs. 5.22 to 5.24 (for the definition of the backgrounds normalised by the fit, see the following paragraph). Additional plots of modelling systematic uncertainties for $t\bar{t}$, $t\bar{t}Z$, ZZ and signal are compiled in Appendix B.1.

5.9.1.2 Background estimation

The estimations of the $t\bar{t}$ and $Z+\text{jets}$ background processes containing non-prompt leptons (from heavy flavour hadronic decays), the $Z\gamma$ background process containing electrons from γ conversions, and the dominant, prompt, WZ background, are performed as part of the fit by including both SRs and CRs in the fit. This is done by including a free-floating normalisation factor (NF) for each one of them. The NFs are symbolised by the letter "k" in the following, hence the NFs included are: $k(Z+\text{HF}\mu)$, $k(Z+\text{HF}e)$, $k(t\bar{t},\text{HF}\mu)$, $k(t\bar{t},\text{HF}e)$, $k(Z\gamma(\gamma \to e^+e^-))$ and $k(WZ)$. The minor components of $Z+\text{jets}$ and $t\bar{t}$ backgrounds which do not originate from heavy flavour hadronic decays, and of $Z\gamma$ not from photon conversions, are collected into a single *NP, other* category (where *NP* here stands for non-prompt), which is not subject to a NF. The non-prompt lepton background estimation has been broken down by source ($t\bar{t}$ and $Z+\text{jets}$) as well as by type (muon or electron) to accommodate differences in their modelling. The VRs serve as "spectator regions", in which to validate the background estimations. The parameter-of-interest is the signal strength modifier, μ_{CLFV}, which is a normalisation factor applied to the signal sample, while the background NFs form additional nuisance parameters.

5.9.1.3 Limit calculation for the CLFV process

The expected limit on the CLFV process is given by TRExFitter in terms of μ_{CLFV}, which can be translated into a limit on the CLFV cross-section. This cross-section is the linear sum of those of the decay and production signal processes, both of which depend on the square of the value of the Wilson coefficients (WCs) which give rise to the processes:



(a) (b)

Figure 5.21: Pruning of some of the "instrumental" lepton-, jet- and b-tag-related nuisance parameters (a), and of the most important signal and $t\bar{t}$ "modelling" nuisance parameters (b). The NPs which are dropped are marked with red colour, while those which have the shape or normalisation component dropped are denoted with yellow or orange respectively. The NPs denoted with green have both the shape and normalisation component kept.

$$\sigma_{\text{CLFV}} = \sigma(qg \to t\ell\ell') + \sigma(pp \to t(\to \ell\ell' q)\bar{t}) \tag{5.29}$$

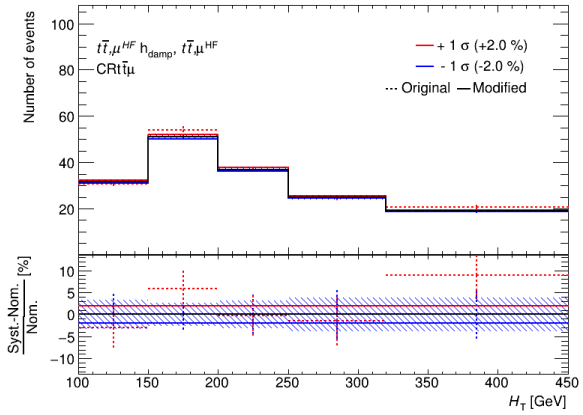
and

$$\sigma_{\text{CLFV}} \propto |C|^2. \tag{5.30}$$

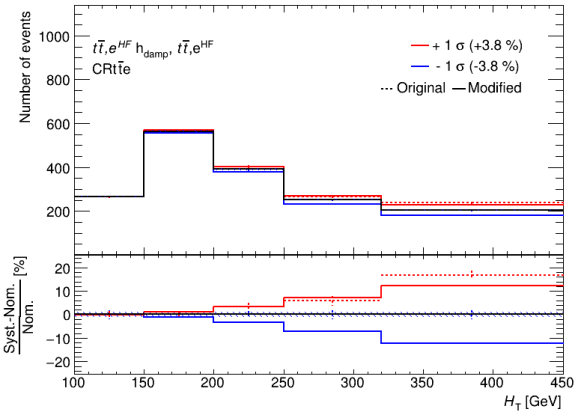
For the starting (pre-fit) cross section of the signal processes, it it is assumed that all of the Wilson coefficients discussed in section 2.4 are equal to the same value of 0.3. In this case the limit on the square of the modulus of the Wilson coefficient value, $|C^{95\%CL}|^2$, is given by the ratio of the limit on the cross section, $\sigma^{95\%CL}$, to the starting cross section obtained when the value of the Wilson coefficients is 0.3, $\sigma^{(|C|=0.3)}$, adjusted by the square of the "starting" WC value:

$$|C^{95\%CL}|^2 = \frac{\sigma^{95\%CL}}{\sigma^{(|C|=0.3)}} \times (0.3)^2 \ . \tag{5.31}$$

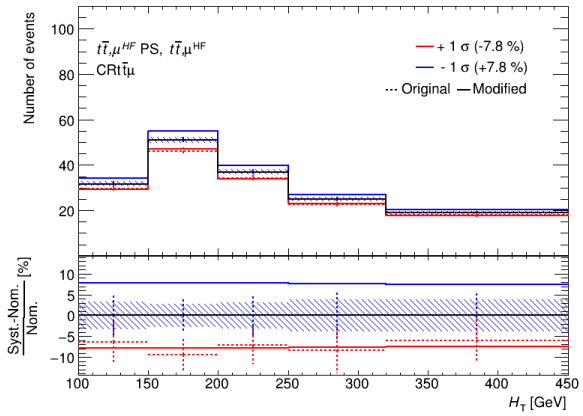
The limit on the modulus of the Wilson coefficient is then obtained as the square root of this. From the limit on the Wilson coefficient it is straightforward to derive a limit on the branching ratio of the top to the CLFV final state using eq. (2.4) and dividing by the SM top quark width, under the assumption



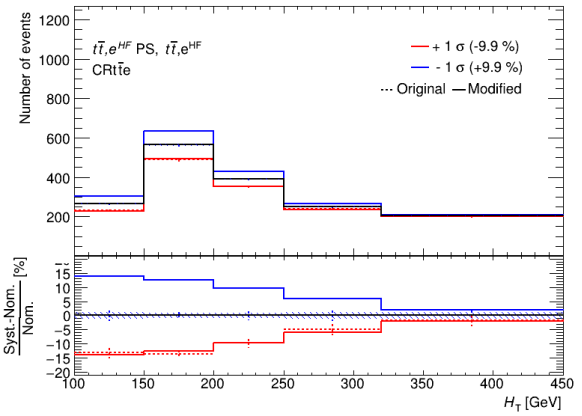
(a) $t\bar{t}$,HFμ h_{damp} uncertainty.



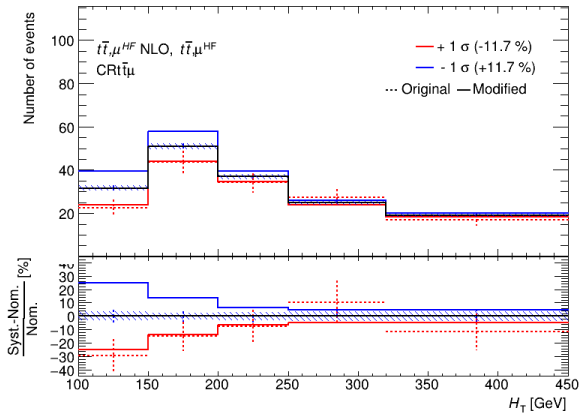
(b) $t\bar{t}$,HFe h_{damp} uncertainty.



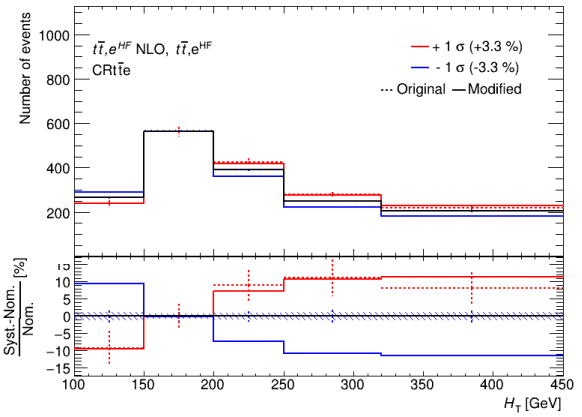
(c) $t\bar{t}$,HFμ parton shower uncertainty.



(d) $t\bar{t}$,HFe parton shower uncertainty.

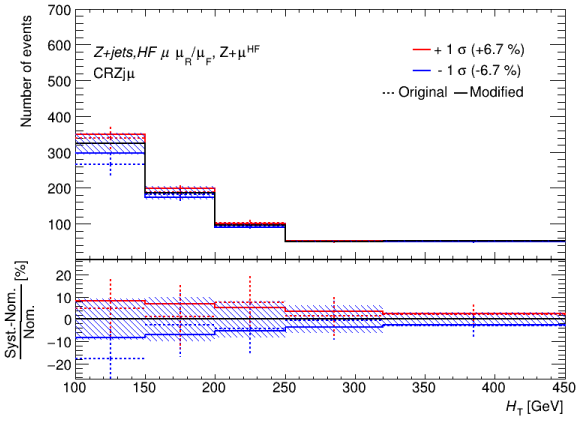


(e) $t\bar{t}$,HFμ NLO matching uncertainty.

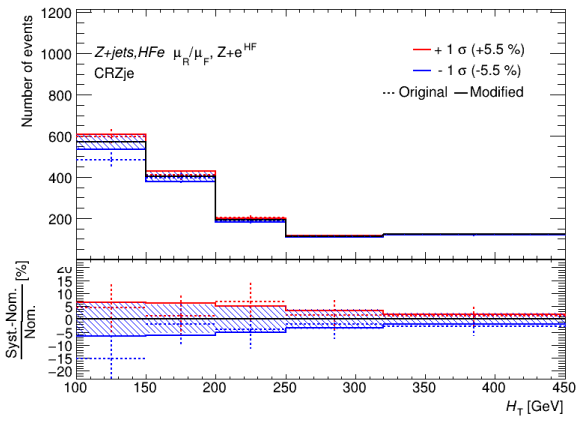


(f) $t\bar{t}$,HFe NLO matching uncertainty.

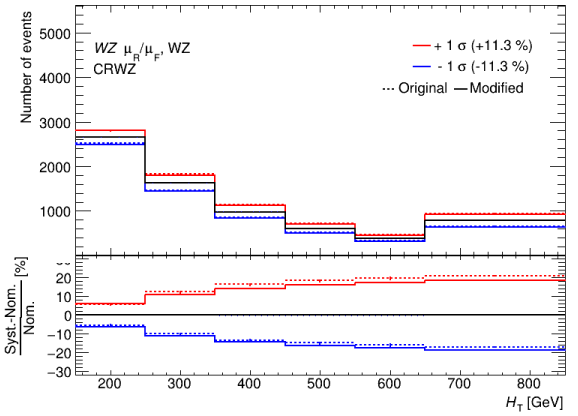
Figure 5.22: Plots in the CR$t\bar{t}$ μ/e regions showing the effects of $\pm1\sigma$ variations of (a,b) h_{damp}, (c,d) parton shower, (e,f) NLO matching modelling systematic uncertainties on $t\bar{t}$, HFμ/e processes, before (original) and after (modified) the application of smoothing and symmetrisation by TRExFitter. The continuous (original) and dashed (modified) lines refer to the up/down (red/blue) variations. The black histogram line, in the upper panel, corresponds to the nominal sample. The symmetrisation used for these systematics is `ONESIDED`, i.e. the "down" variation is added as a mirror image of the "up" variation. Hence the modified (blue) line is not accompanied by a dashed line. The hatched area represents the statistical uncertainty (divided by the MC prediction in the ratio pad).



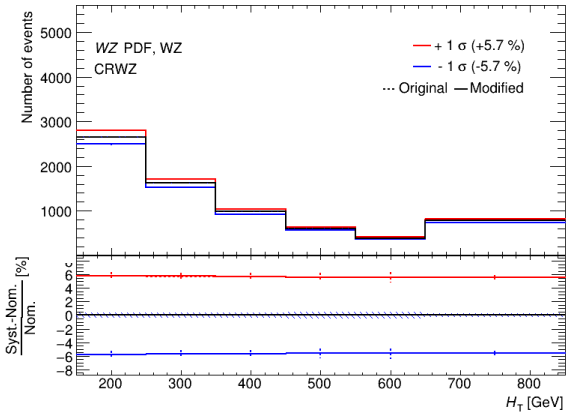
(a) Z+HFμ scale uncertainty.



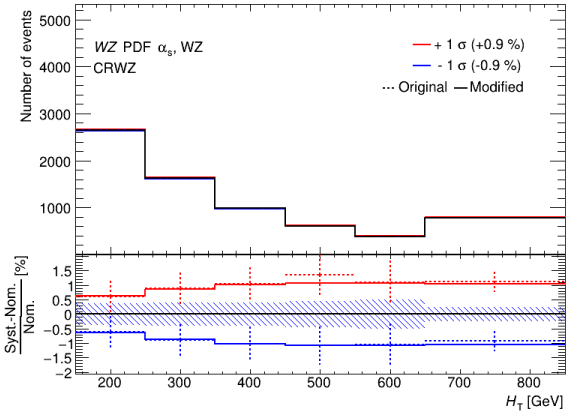
(b) Z+HFe scale uncertainty.



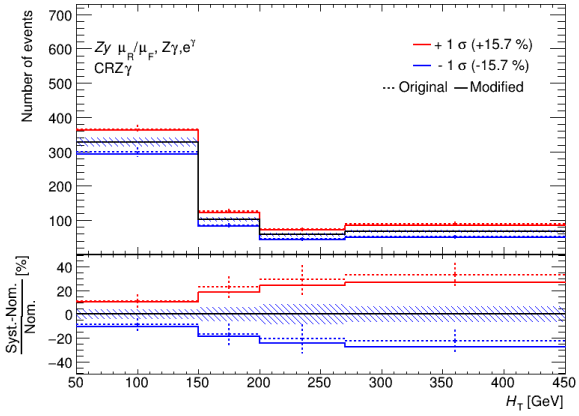
(c) WZ scale uncertainty.



(d) WZ PDF set uncertainty.

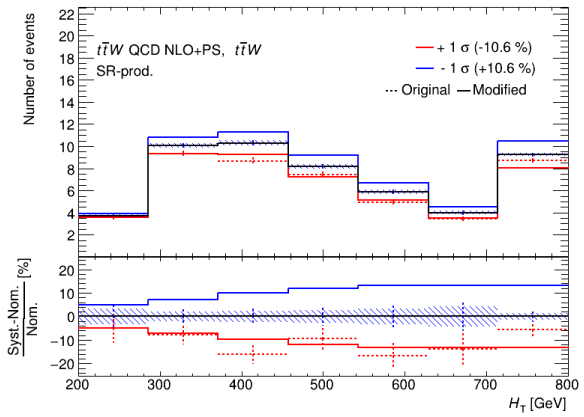


(e) WZ α_s uncertainty.

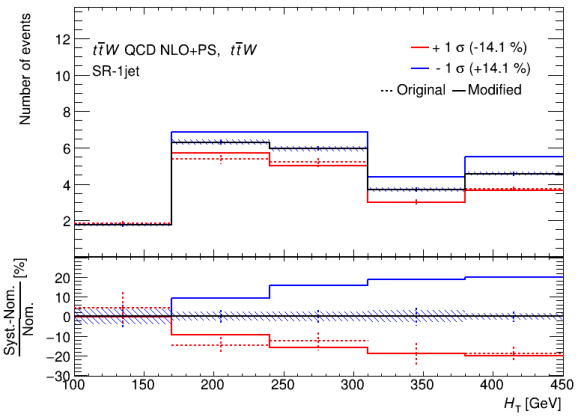


(f) $Z\gamma$ ($\gamma \to e^+e^-$) scale uncertainty.

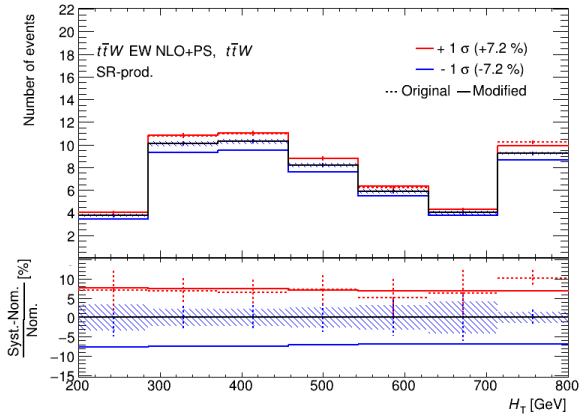
Figure 5.23: Plots in the CRZjμ/e, CRWZ and CR$Z\gamma$ regions showing the effects of $\pm 1\sigma$ variations of the (μ_R/μ_F) scale modelling nuisance parameter, on the backgrounds targeted. For WZ, the PDF set and strong coupling constant variation systematics are also shown. The continuous (original) and dashed (modified) lines refer to the up/down (red/blue) variations. The black histogram line, in the upper panel, corresponds to the nominal sample. The variations are shown both before (original) and after (modified) the application of smoothing and symmetrisation by TRExFitter. The hatched area represents the statistical uncertainty (divided by the MC prediction in the ratio pad).



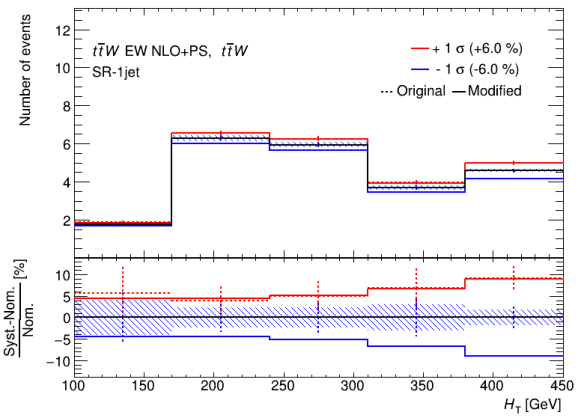
(a) $t\bar{t}W$ ME+PS uncertainty (QCD) in SR-prod.



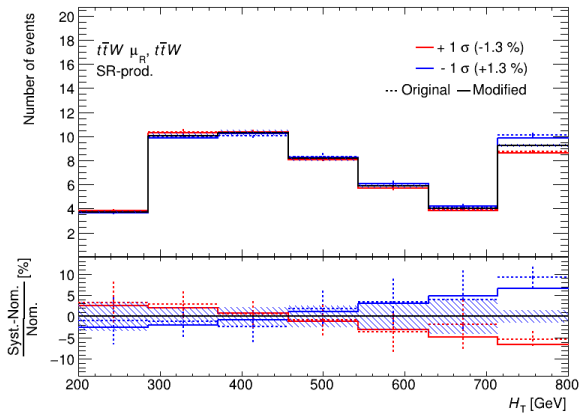
(b) $t\bar{t}W$ ME+PS uncertainty (QCD) in SR-1jet.



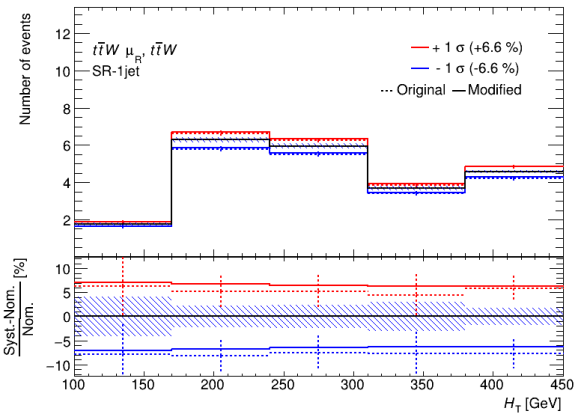
(c) $t\bar{t}W$ ME+PS uncertainty (EW) in SR-prod.



(d) $t\bar{t}W$ ME+PS uncertainty (EW) in SR-1jet.



(e) $t\bar{t}W$ μ_{R} uncertainty in SR-prod.



(f) $t\bar{t}W$ μ_{R} uncertainty in SR-1jet.

Figure 5.24: Plots in the signal regions where $t\bar{t}W$ background is mostly concentrated, showing the effects of $\pm 1\sigma$ variations of (a,b) the combined ME+PS for QCD processes, (c.d) the combined ME+PS for EW processes, (e,f) the renormalisation scale modelling nuisance parameter before and after the application of smoothing and symmetrisation by TRExFitter. The continuous (original) and dashed (modified) lines refer to the up/down (red/blue) variations. The black histogram line, in the upper panel, corresponds to the nominal sample. The symmetrisation used for ME+PS systematics is `ONESIDED`, i.e. the "down" variation is added as a mirror image of the "up" variation. Hence the modified (blue) line is not accompanied by a dashed line. The hatched area represents the statistical uncertainty (divided by the MC prediction in the ratio pad).

of the same coupling in the decay and production vertices.

5.9.1.4 Fitted distributions and binning

In the SR-decay region, the fitted distribution is the BDT score, which has been shown to provide good separation between decay-signal and background. In all other regions, the fitted variable is the H_T (scalar sum of p_{T} of leptons and jets), as it is representative of the energy of the final state of each process. In the SRs, this translates into a good separation between production-signal and backgrounds.

The binning has been chosen by taking into account that a smaller number of bins simplifies the fit, improving its stability and reducing the running time, and allows for a sufficient number of events per bin. On the other hand, further reducing the bin number would result in a loss in sensitivity. In the SR-decay region, the number and size of the bins are determined automatically by the algorithm *Transformation D* of TRExFitter.

5.9.2 Fit using Asimov data

For the purposes of testing and validating the analysis strategy, a first fit is run with the SRs *blinded*, that is no real data are used in the SRs. This is to avoid biases in the analysis design coming from looking at the data distribution or basing the selection of events on the significance of the final results.

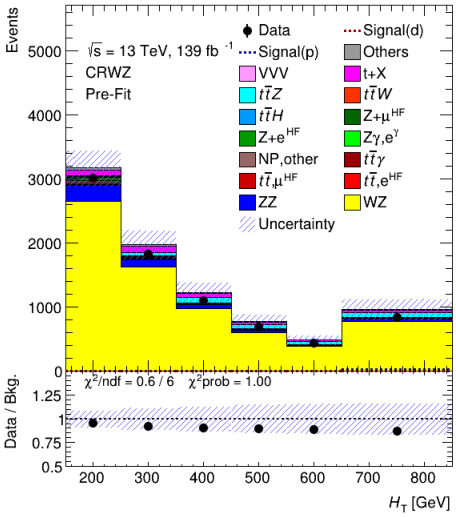
Hence, the fit is first run using a *mixed* data and Asimov data sample; real data are used in the CRs, while Asimov data (represented by the MC expectation, as discussed in section 5.8.5) are used in the SRs. No signal is included in the construction of the Asimov dataset ($\mu_{CLFV,A} = 0$). TRExFitter handles this fit in two stages, first performing a fit in the CRs. This uses the data to extract a value for the NFs as well as for each of the other nuisance parameters. A "modified" Asimov data set is then constructed for each of the SRs using the values of the nuisance parameters found in the first step. From here, a fit is performed in both regions using the modified Asimov dataset in the SRs and data in the CRs. Pre- and post-fit plots of the analysis regions included in this fit are shown in figs. 5.25 to 5.27.

The normalisation factors obtained by the fit are shown in table 5.14. The large $k(Z\gamma(\gamma \to e^+e^-))$ can be explained by the large data/MC difference in CR$Z\gamma$ (fig. 5.25c).

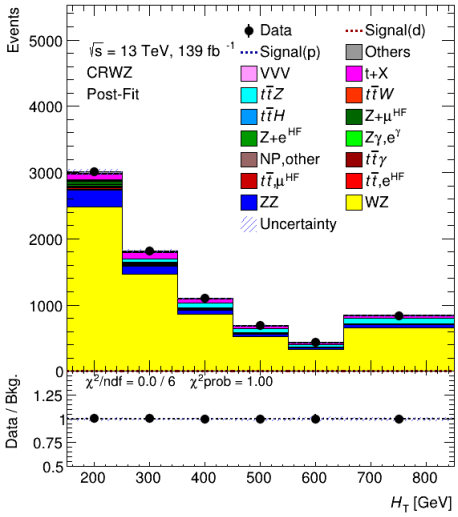
μ_{CLFV}	$k(WZ)$	$k(Z\gamma(\gamma \to e^+e^-))$	$k(Z+HF\mu)$	$k(Z+HFe)$	$k(t\bar{t},HF\mu)$	$k(t\bar{t},HFe)$
0.000 ± 0.006	0.964 ± 0.078	1.31 ± 0.18	1.00 ± 0.11	0.97 ± 0.10	1.04 ± 0.16	1.03 ± 0.14

Table 5.14: Table showing the normalisation factors obtained from the fit using Asimov data in SRs.

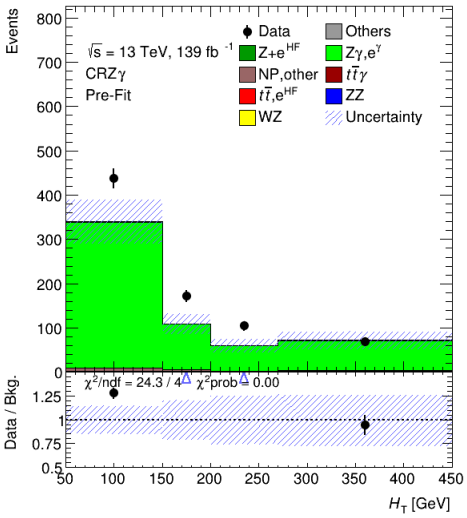
The parameters extracted from the fit are applied to the VRs, which are shown in fig. 5.28. Summary plots are given in figs. 5.29 and 5.30. Pre- and post-fit yields in all analysis regions are collected in tables 5.15 and 5.16.



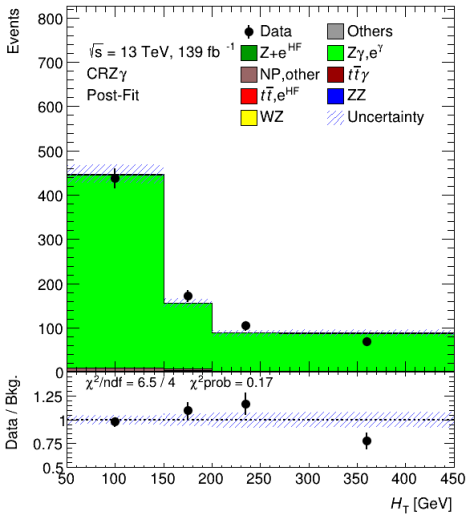
(a) Pre-fit H_{T} distribution in CRWZ.



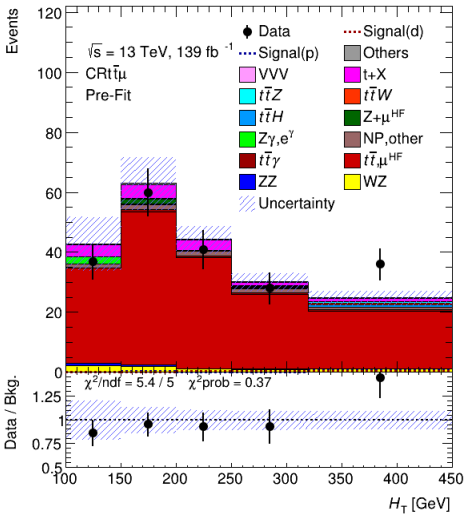
(b) Post-fit H_{T} distribution in CRWZ.



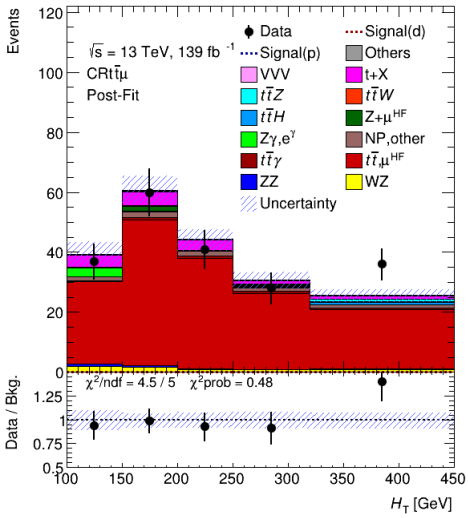
(c) Pre-fit H_{T} distribution in CR$Z\gamma$.



(d) Post-fit H_{T} distribution in CR$Z\gamma$.

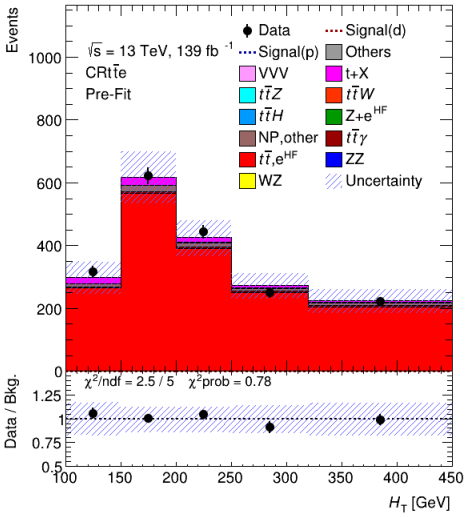


(e) Pre-fit H_{T} distribution in CR$t\bar{t}\mu$.

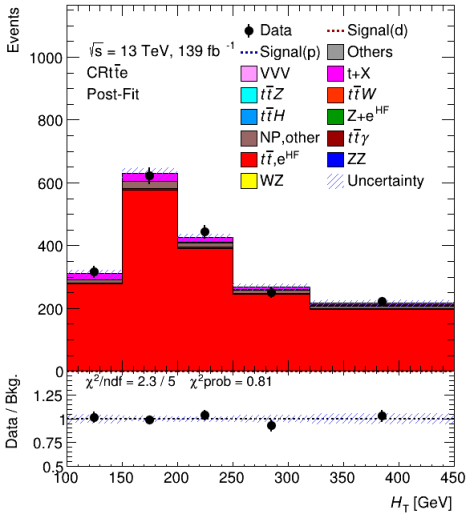


(f) Post-fit H_{T} distribution in CR$t\bar{t}\mu$.

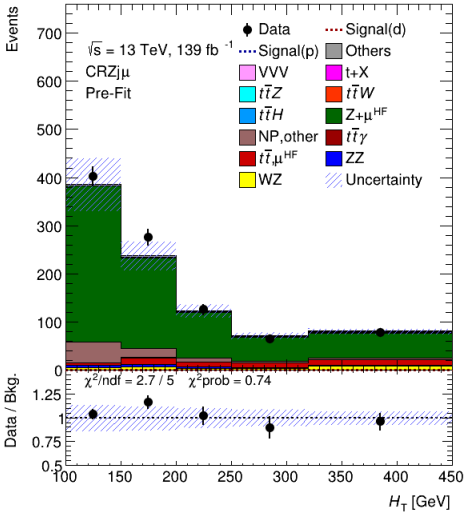
Figure 5.25: Pre-fit (figs. 5.25a, 5.25c and 5.25e) and post-fit (figs. 5.25b, 5.25d and 5.25f) plots of CRWZ, CR$Z\gamma$ and CR$t\bar{t}\mu$, for the fit using Asimov data in SRs.



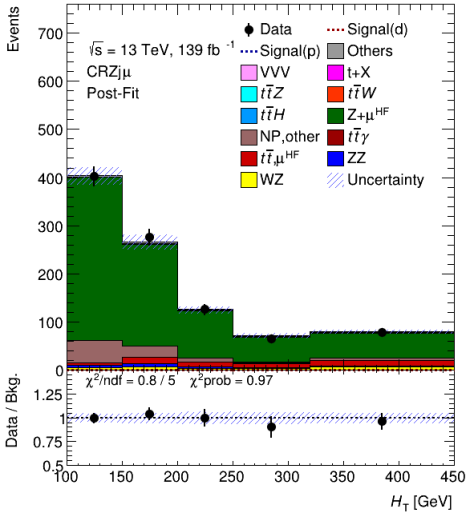
(a) Pre-fit H_{T} distribution in CR$t\bar{t}e$.



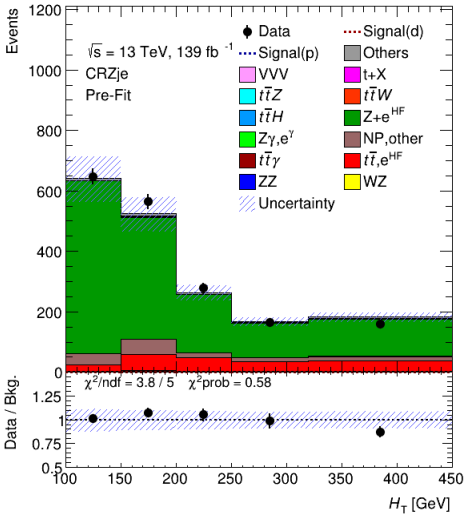
(b) Post-fit H_{T} distribution in CR$t\bar{t}e$.



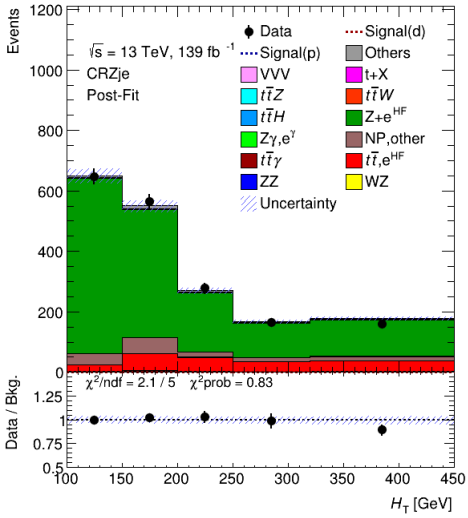
(c) Pre-fit H_{T} distribution in CR$Z\mathrm{j}\mu$.



(d) Post-fit H_{T} distribution in CR$Z\mathrm{j}\mu$.

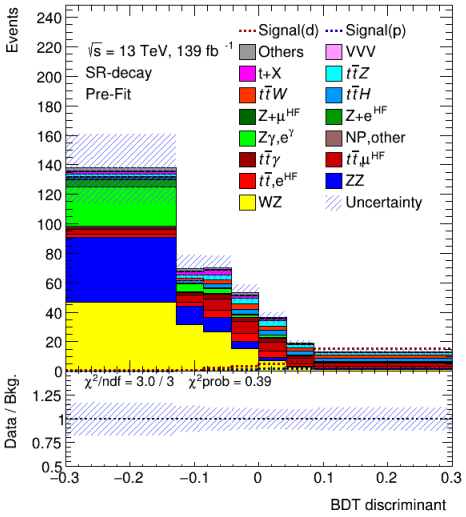


(e) Pre-fit H_{T} distribution in CR$Z\mathrm{j}e$.

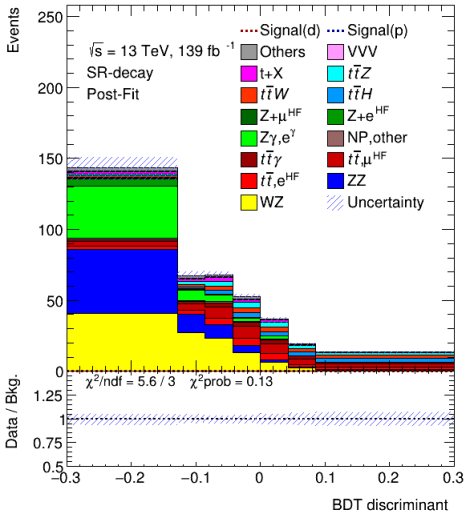


(f) Post-fit H_{T} distribution in CR$Z\mathrm{j}e$.

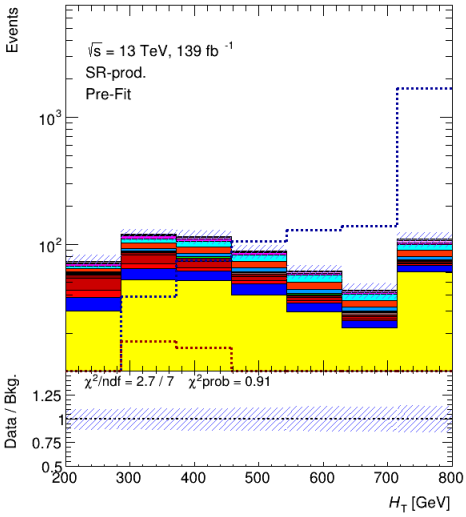
Figure 5.26: Pre-fit (figs. 5.26a, 5.26c and 5.26e) and post-fit (figs. 5.26b, 5.26d and 5.26f) plots of CR$t\bar{t}e$, CR$Z\mathrm{j}\mu$, CR$Z\mathrm{j}e$, for the fit using Asimov data in SRs.



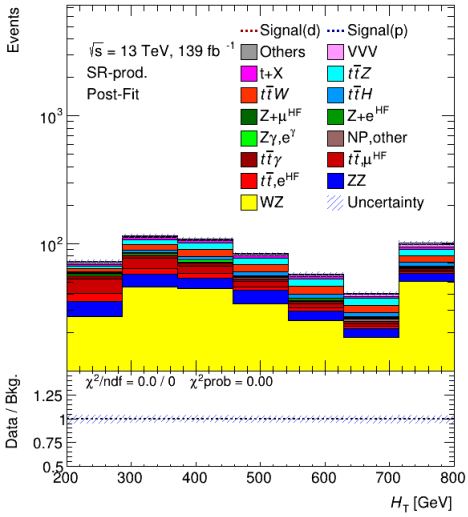
(a) Pre-fit BDT score distribution in SR-decay.



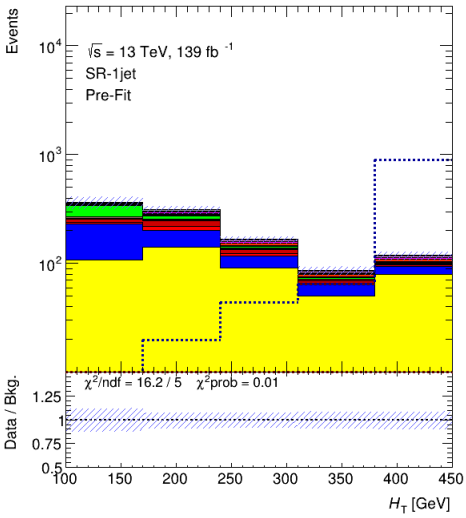
(b) Post-fit BDT score distribution in SR-decay.



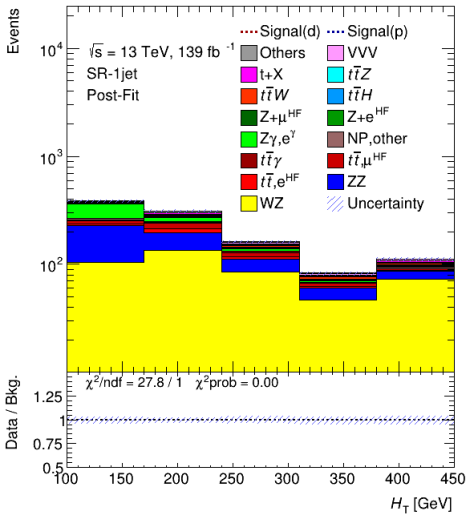
(c) Pre-fit H_{T} distribution in SR-production.



(d) Post-fit H_{T} distribution in SR-production.

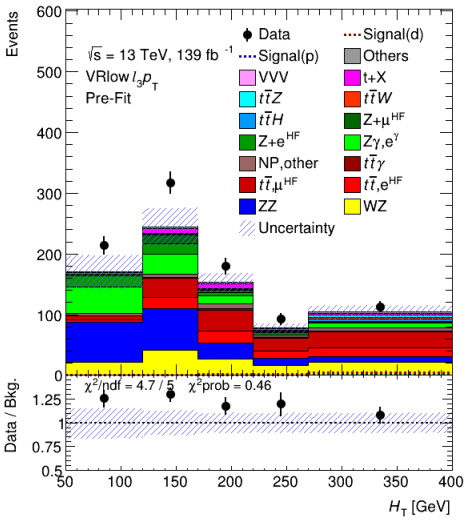


(e) Pre-fit H_{T} distribution in SR-1jet.

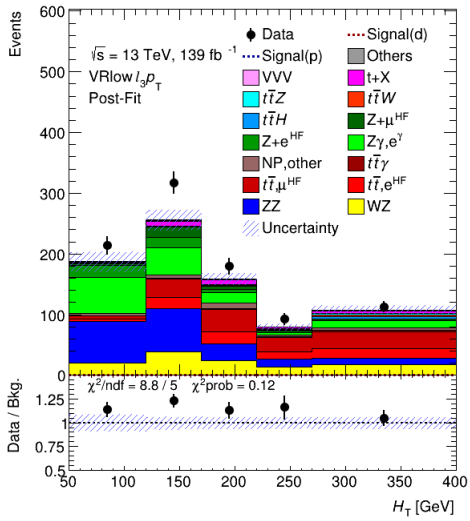


(f) Post-fit H_{T} distribution in SR-1jet.

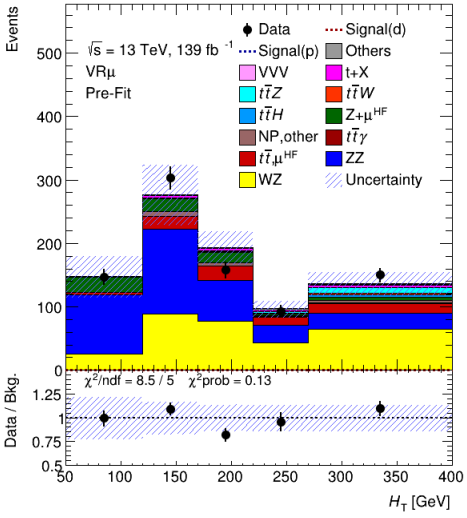
Figure 5.27: Pre-fit (figs. 5.27a, 5.27c and 5.27e) and post-fit (figs. 5.27b, 5.27d and 5.27f) plots of SR-decay, SR-production, SR-1jet (in logarithmic scale), for the fit using Asimov data in SRs. Some legends are suppressed, for the signal distributions to be apparent.



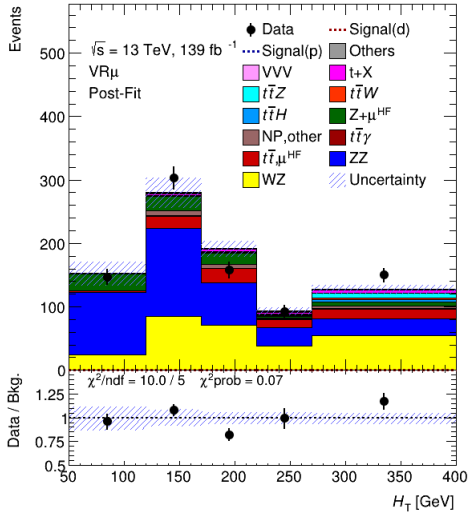
(a) Pre-fit H_{T} distribution in VR$low\ell_3p_{\mathrm{T}}$.



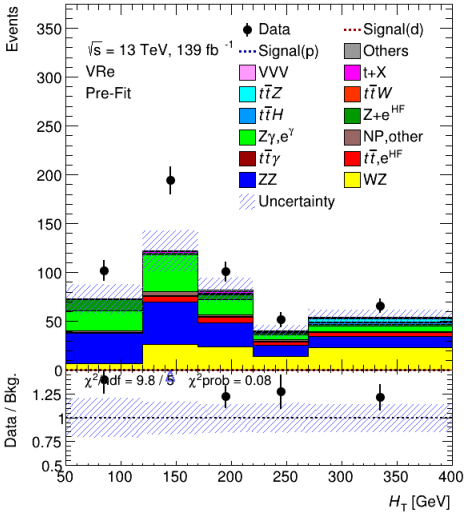
(b) Post-fit H_{T} distribution in VR$low\ell_3p_{\mathrm{T}}$.



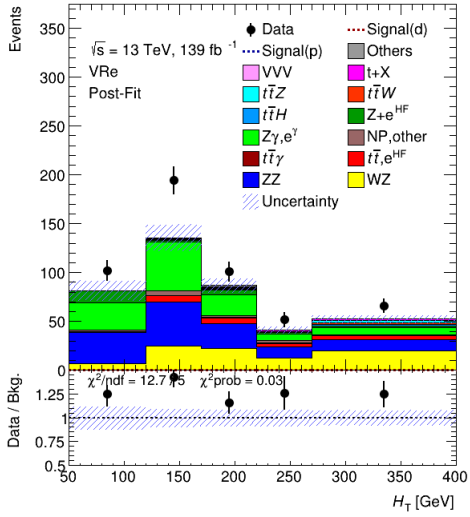
(c) Pre-fit H_{T} distribution in VRμ.



(d) Post-fit H_{T} distribution in VRμ.

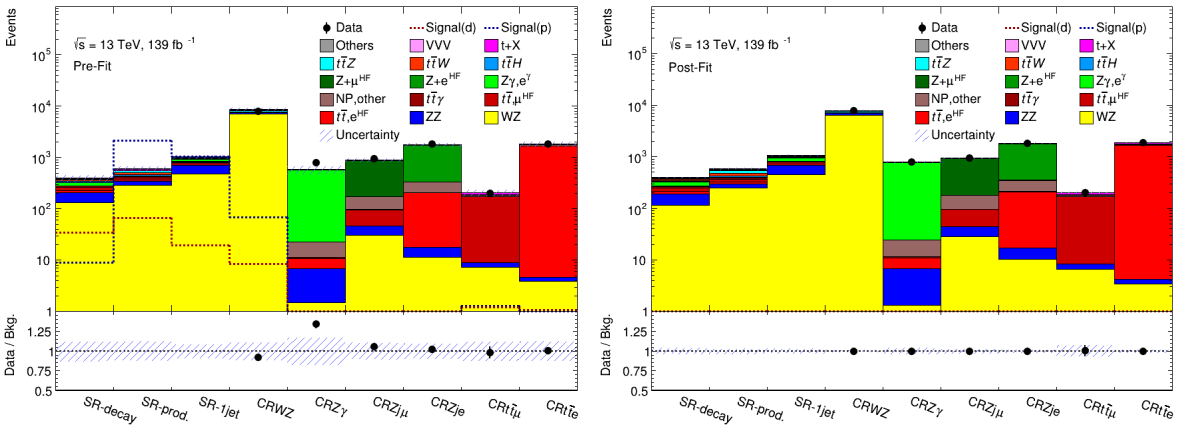


(e) Pre-fit H_{T} distribution in VRe.



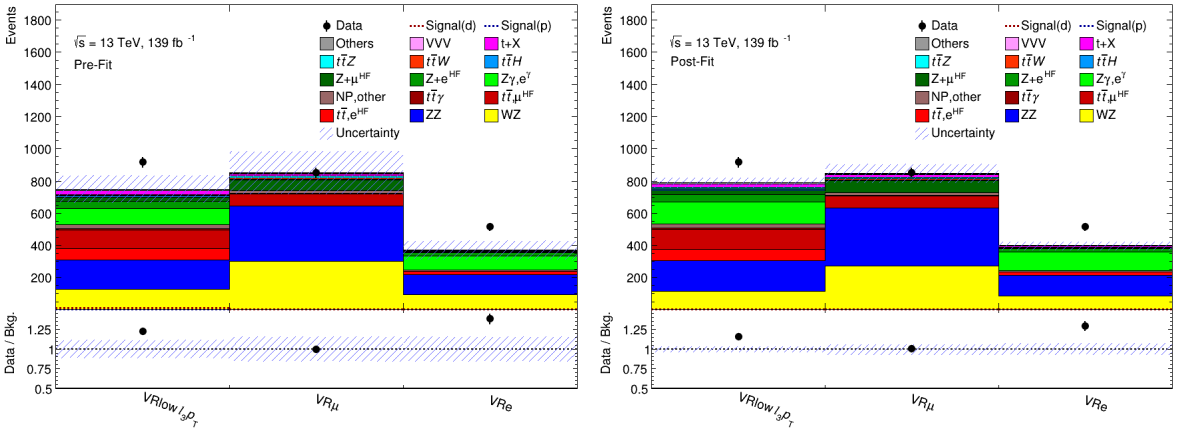
(f) Post-fit H_{T} distribution in VRe.

Figure 5.28: Pre-fit (figs. 5.28a, 5.28c and 5.28e) and post-fit (figs. 5.28b, 5.28d and 5.28f) plots of VR$low\ell_3p_{\mathrm{T}}$, VRμ, VRe, for the fit using Asimov data in SRs.



(a) (b)

Figure 5.29: Pre/post-fit summary plot of SRs and CRs using the fit with Asimov data in the SRs.



(a) (b)

Figure 5.30: Pre/post-fit summary plot of VRs using the fit with Asimov data in the SRs.

	SR-decay	SR-prod.	SR-1jet	VR$low\ell_3 p_{\mathrm{T}}$	VRμ	VRe	CRWZ	CR$Z\gamma$	CRZjμ	CRZje	CR$t\bar{t}\mu$	CR$t\bar{t}e$
Signal (d)	34.1 ± 4.2	66.8 ± 2.9	19.4 ± 4.1	8.9 ± 0.6	2.0 ± 0.3	0.9 ± 0.1	8.4 ± 0.9	0.0	0.08 ± 0.03	0.05 ± 0.02	1.2 ± 0.3	1.1 ± 0.1
Signal (p)	9.0 ± 1.2	2165 ± 183	1031 ± 84	3.4 ± 1.0	1.8 ± 0.3	0.8 ± 0.2	67.6 ± 6.3	0.0	0.8 ± 0.2	0.5 ± 0.4	1.3 ± 0.5	1.0 ± 0.5
Others	7.0 ± 4.3	12.1 ± 6.4	16.1 ± 8.9	8.9 ± 4.9	5.5 ± 3.4	3.3 ± 2.2	66 ± 34	3.9 ± 2.0	10.6 ± 5.4	27 ± 14	1.1 ± 0.7	1.4 ± 0.7
VVV	3.3 ± 1.7	14.9 ± 7.5	18.7 ± 9.4	1.8 ± 0.9	2.1 ± 1.1	0.8 ± 0.4	32 ± 16	0.0	0.1 ± 0.1	0.0	0.2 ± 0.1	0.2 ± 0.1
$t + X$	11.2 ± 2.2	22.8 ± 3.6	23.4 ± 4.8	22.3 ± 3.5	14.6 ± 3.4	4.8 ± 1.2	345 ± 50	0.0	6.5 ± 1.5	9.6 ± 3.7	14.6 ± 4.5	70 ± 11
$t\bar{t}Z$	18.0 ± 4.6	51.5 ± 9.8	8.6 ± 2.6	6.6 ± 2.6	12.2 ± 4.0	4.6 ± 2.0	345 ± 55	0.0	1.7 ± 0.3	2.9 ± 0.7	0.9 ± 0.3	2.3 ± 1.2
$t\bar{t}W$	15.9 ± 8.3	51 ± 27	22.3 ± 11.8	5.4 ± 2.8	5.7 ± 3.0	2.2 ± 1.1	14.3 ± 7.2	0.0	0.1 ± 0.1	0.2 ± 0.2	1.0 ± 0.5	2.5 ± 1.3
$t\bar{t}H$	17.1 ± 2.0	23.4 ± 2.6	3.7 ± 0.6	5.8 ± 0.7	5.6 ± 0.6	2.2 ± 0.3	10.0 ± 1.1	0.0	0.09 ± 0.02	0.16 ± 0.05	1.2 ± 0.1	2.3 ± 0.4
Z+HFμ	1.1 ± 0.9	1.2 ± 1.3	15.1 ± 5.7	24.7 ± 13.8	67 ± 24	0.0	83.1 ± 13.5	0.0	708 ± 79	0.0	1.8 ± 1.5	0.0
Z+HFe	5.6 ± 4.4	4.2 ± 1.3	8.1 ± 6.7	45 ± 12	0.0	21.1 ± 8.7	32.1 ± 8.4	2.6 ± 2.1	0.0	1403 ± 140	0.0	2.1 ± 1.9
$Z\gamma, e^{\gamma}$	39.0 ± 11.2	6.7 ± 1.7	97.6 ± 16.1	101 ± 20	0.0	85 ± 17	17.8 ± 4.3	556 ± 101	0.0	2.4 ± 1.5	2.8 ± 1.8	0.0
NP,other	2.7 ± 3.4	4.6 ± 2.2	15.2 ± 8.1	25 ± 12	17.0 ± 9.8	8.3 ± 5.6	35 ± 17	11.6 ± 7.5	77 ± 34	128 ± 55	6.4 ± 2.6	63 ± 25
$t\bar{t}\gamma$	14.0 ± 2.4	17.9 ± 3.0	13.2 ± 2.3	7.4 ± 1.3	2.3 ± 0.5	2.8 ± 0.6	6.5 ± 1.1	0.5 ± 0.2	0.6 ± 0.2	2.2 ± 0.5	2.8 ± 0.5	19.2 ± 3.3
$t\bar{t}$,HFμ	36.7 ± 3.9	39.9 ± 8.6	63.1 ± 8.6	120 ± 10	73.9 ± 8.6	0.0	31.5 ± 10.7	0.0	50.1 ± 8.0	0.0	164 ± 24	0.0
$t\bar{t}$,HFe	22.3 ± 5.4	22.1 ± 2.9	38.4 ± 9.1	67.0 ± 6.5	0.0	20.0 ± 3.3	15.0 ± 3.8	3.9 ± 0.7	0.0	187 ± 20	0.0	1675 ± 216
ZZ	74 ± 27	54 ± 19	232 ± 72	184 ± 60	347 ± 112	123 ± 40	532 ± 180	5.3 ± 1.9	15.6 ± 5.0	6.4 ± 2.1	1.8 ± 0.6	0.8 ± 0.3
WZ	132 ± 27	286 ± 57	472 ± 38	127 ± 16	299 ± 37	95 ± 13	7034 ± 946	1.5 ± 0.3	30.5 ± 3.9	11.4 ± 1.8	7.2 ± 1.0	3.8 ± 0.7
Tot. Bkg.	400 ± 51	612 ± 73	1047 ± 94	750 ± 84	852 ± 131	373 ± 57	8598 ± 994	585 ± 103	900 ± 93	1780 ± 171	205 ± 26	1841 ± 223
Data	463	611	1238	917	851	515	7906	784	948	1815	202	1855

Table 5.15: Pre-fit yields in the analysis regions. The signal processes are normalised using Wilson coefficient values equal to 0.3. The errors quoted include MC statistical and systematic uncertainties.

	SR-decay	SR-prod.	SR-1jet	VR$low\ell_3 p_{\mathrm{T}}$	VRμ	VRe	CRWZ	CR$Z\gamma$	CRZjμ	CRZje	CR$t\bar{t}\mu$	CR$t\bar{t}e$
Signal (d)	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Signal (p)	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Others	6.4 ± 3.5	11.2 ± 5.8	14.8 ± 7.9	8.0 ± 4.2	5.2 ± 2.8	3.0 ± 1.6	61 ± 32	3.6 ± 1.8	10.0 ± 5.1	25 ± 13	1.0 ± 0.5	1.3 ± 0.7
VVV	3.4 ± 1.7	15.1 ± 7.4	18.6 ± 9.1	1.8 ± 0.9	2.2 ± 1.1	0.8 ± 0.4	32 ± 16	0.0	0.13 ± 0.07	0.0	0.2 ± 0.1	0.2 ± 0.1
$t+X$	11.4 ± 2.1	22.9 ± 3.5	23.0 ± 4.5	22.5 ± 3.1	14.7 ± 3.3	4.8 ± 1.1	344 ± 49	0.0	6.7 ± 1.3	9.8 ± 3.6	14.6 ± 4.3	71 ± 11
$t\bar{t}Z$	17.1 ± 4.2	49.7 ± 9.0	7.9 ± 2.2	6.1 ± 2.4	11.5 ± 3.7	4.2 ± 1.8	335 ± 51	0.0	1.6 ± 0.3	2.8 ± 0.7	0.9 ± 0.3	2.0 ± 1.1
$t\bar{t}W$	16.5 ± 6.5	53 ± 21	22.6 ± 9.1	5.6 ± 2.2	5.9 ± 2.5	2.3 ± 0.9	14.7 ± 5.6	0.0	0.10 ± 0.07	0.2 ± 0.1	1.1 ± 0.4	2.5 ± 1.0
$t\bar{t}H$	17.2 ± 1.9	23.5 ± 2.5	3.6 ± 0.5	5.8 ± 0.7	5.6 ± 0.6	2.2 ± 0.3	10.0 ± 1.1	0.0	0.09 ± 0.02	0.16 ± 0.05	1.3 ± 0.1	2.3 ± 0.4
Z+HFμ	1.3 ± 0.5	1.5 ± 0.8	15.2 ± 2.4	26.7 ± 9.8	74.2 ± 16.1	0.0	87 ± 11	0.0	750 ± 44	0.0	1.9 ± 0.9	0.0
Z+HFe	5.8 ± 3.1	4.3 ± 0.6	9.8 ± 4.9	46.5 ± 6.0	0.0	22.3 ± 3.6	32.3 ± 6.3	2.4 ± 0.5	0.0	1429 ± 67	0.0	2.2 ± 1.7
$Z\gamma$,e^{γ}	53.2 ± 4.8	9.0 ± 0.7	132.0 ± 9.2	136 ± 10	0.0	114.2 ± 9.8	24.2 ± 2.9	754 ± 29	0.0	2.9 ± 1.3	3.7 ± 0.3	0.0
NP,other	4.2 ± 4.4	4.7 ± 1.8	15.0 ± 6.4	25.6 ± 9.6	18.5 ± 7.0	8.9 ± 4.0	38 ± 16	12.6 ± 5.3	85 ± 32	136 ± 52	6.8 ± 2.5	66 ± 24
$t\bar{t}\gamma$	14.1 ± 2.2	18.0 ± 2.8	13.0 ± 2.1	7.4 ± 1.2	2.4 ± 0.4	2.8 ± 0.5	6.5 ± 1.0	0.5 ± 0.2	0.6 ± 0.1	2.2 ± 0.4	2.8 ± 0.4	19.3 ± 3.2
$t\bar{t}$,HFμ	37.7 ± 5.9	45 ± 11	62.6 ± 7.9	123 ± 14	74.5 ± 8.9	0.0	37 ± 12	0.0	50.1 ± 6.0	0.0	158 ± 15	0.0
$t\bar{t}$,HFe	21.9 ± 2.5	22.5 ± 1.8	37.7 ± 5.3	69.6 ± 7.1	0.0	20.7 ± 2.1	15.1 ± 2.1	4.2 ± 0.7	0.0	191.7 ± 9.6	0.0	1683 ± 51
ZZ	77 ± 14	55 ± 10	237 ± 34	191 ± 29	359 ± 54	126 ± 20	552 ± 95	5.5 ± 1.0	16.5 ± 2.5	6.7 ± 1.1	1.8 ± 0.3	0.8 ± 0.1
WZ	114.0 ± 5.2	244.3 ± 9.5	443 ± 18	115.8 ± 4.4	273 ± 12	86.1 ± 5.2	6316 ± 156	1.3 ± 0.1	28.1 ± 1.2	10.3 ± 0.6	6.5 ± 0.3	3.5 ± 0.2
Tot.Bkg.	401 ± 14	580 ± 19	1056 ± 29	791 ± 37	847 ± 60	398 ± 30	7906 ± 89	784 ± 28	949 ± 31	1817 ± 43	201 ± 14	1854 ± 43
Data	–	–	–	917	851	515	7906	784	948	1815	202	1855

Table 5.16: Post-fit yields from the mixed data and Asimov fit. The errors quoted include MC statistical and systematic uncertainties.

5.9.2.1 Correlations, constraints and pulls of nuisance parameters

The correlation coefficients between the nuisance parameters and the POI are determined automatically from the respective covariance matrix during the minimisation of the negative log likelihood function, and are shown in fig. 5.31. Nuisance parameters which correspond to uncertainties on background sources with similar shapes are expected to be correlated.

Systematic (i.e. constrained) uncertainty nuisance parameters and normalisation factors are used in conjunction to correct for data/MC differences. Correlation of a NP with a free-floating normalisation factor, implies a NP whose variation strongly affects the value and uncertainty of that NF. Conversely, the NF may result in a reduction of the post-fit uncertainty of the NP. It can be seen from fig. 5.31, that the systematic uncertainty NPs of the backgrounds that we choose to normalise via NFs, have large (anti-)correlations with the respective NFs. E.g. the WZ scale ($\mu_{\mathrm{R}}/\mu_{\mathrm{F}}$) and PDF uncertainties, and $k(WZ)$ are adjusted during the fit to correct the data/MC difference seen in CRWZ, and are anticorrelated. The ZZ background is concentrated in the same bins as WZ, so the NPs related to ZZ, namely the cross-section and scale uncertainty NPs, appear anticorrelated to each other and correlated to the WZ scale NP. A significant correlation between a nuisance parameter and the parameter-of-interest, μ_{CLFV}, indicates that varying this parameter has a large impact on the expected signal strength[23].

The *pull* of a NP is the difference between its *best-fit-value* and the pre-fit value, divided by the pre-fit uncertainty:

$$\mathrm{pull}(\theta) = \frac{\hat{\theta} - \theta_0}{\Delta\theta} \,. \tag{5.32}$$

The nuisance parameter pull plot can be seen in fig. 5.32. The uncertainty shown in the plot is the "post-fit" uncertainty, i.e. the uncertainty on the best-fit-value[24].

In general large (post-fit) constraints are not seen, i.e. the uncertainty per NP in fig. 5.32 is close to one. However, some NPs are constrained. For example the NPs corresponding to the $t\bar{t}$ NLO matching modelling systematic uncertainties are constrained by the fit (particularly for the HFe component, which uses the higher statistics CR$t\bar{t}e$ region). This implies that the fit is able to reduce the sizes of these uncertainties based on information extracted from the regions provided, and most importantly from the dedicated control regions, which target the specific backgrounds[25]. The $t\bar{t}W$ cross-section uncertainty is also constrained; given its high initial (or pre-fit) value (50%), the fit uses the data, mainly in the SR-production region where $t\bar{t}W$ has its main contribution compared to other regions, to reduce the pre-fit uncertainty. The fit is also able to constrain the large (30%) pre-fit ZZ cross-section uncertainty, by using the high-statistics bins in CRWZ and, mainly, the SRs.

Also, in general, only a small number of pulls is observed. A pull appears in fig. 5.32 as a departure from the nominal pre-fit value, which is zero as discussed in section 5.8.1. In such cases, the fit adjusts

[23] The correlation matrix is compatible with the one obtained when running the fit with Asimov data in all regions (SRs and CRs), shown in Appendix B.2.

[24] We note that when referring to pulls, we are talking about nuisance parameters which have a pre-fit uncertainty defined (i.e. they enter the LF with a constraint term), and thus no NFs (nor the POI) nor gammas appear in fig. 5.32.

[25] The constraints are compatible with the ones seen in the case of a fit that uses Asimov data in all regions (SRs and CRs), shown in Appendix B.2.

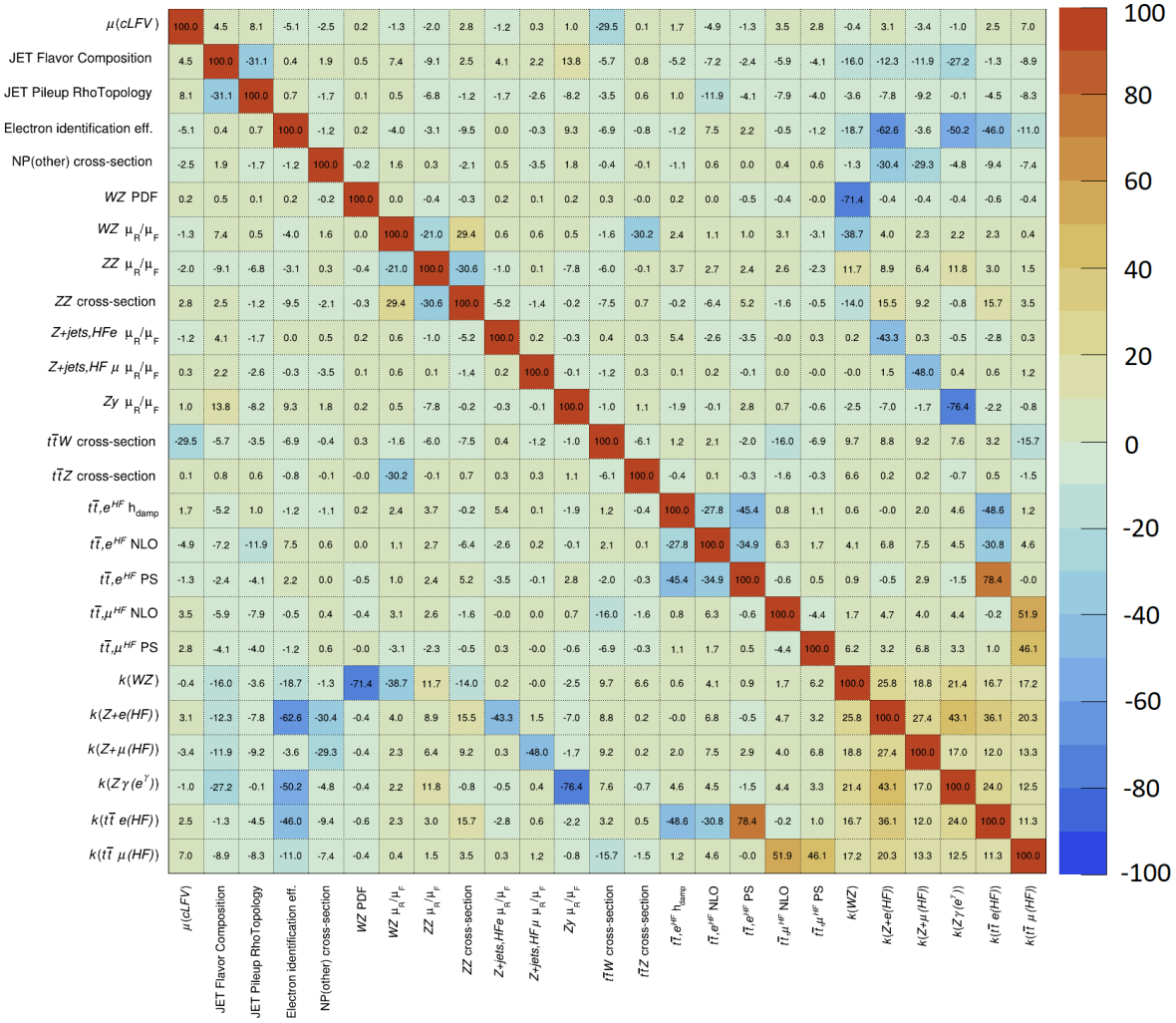
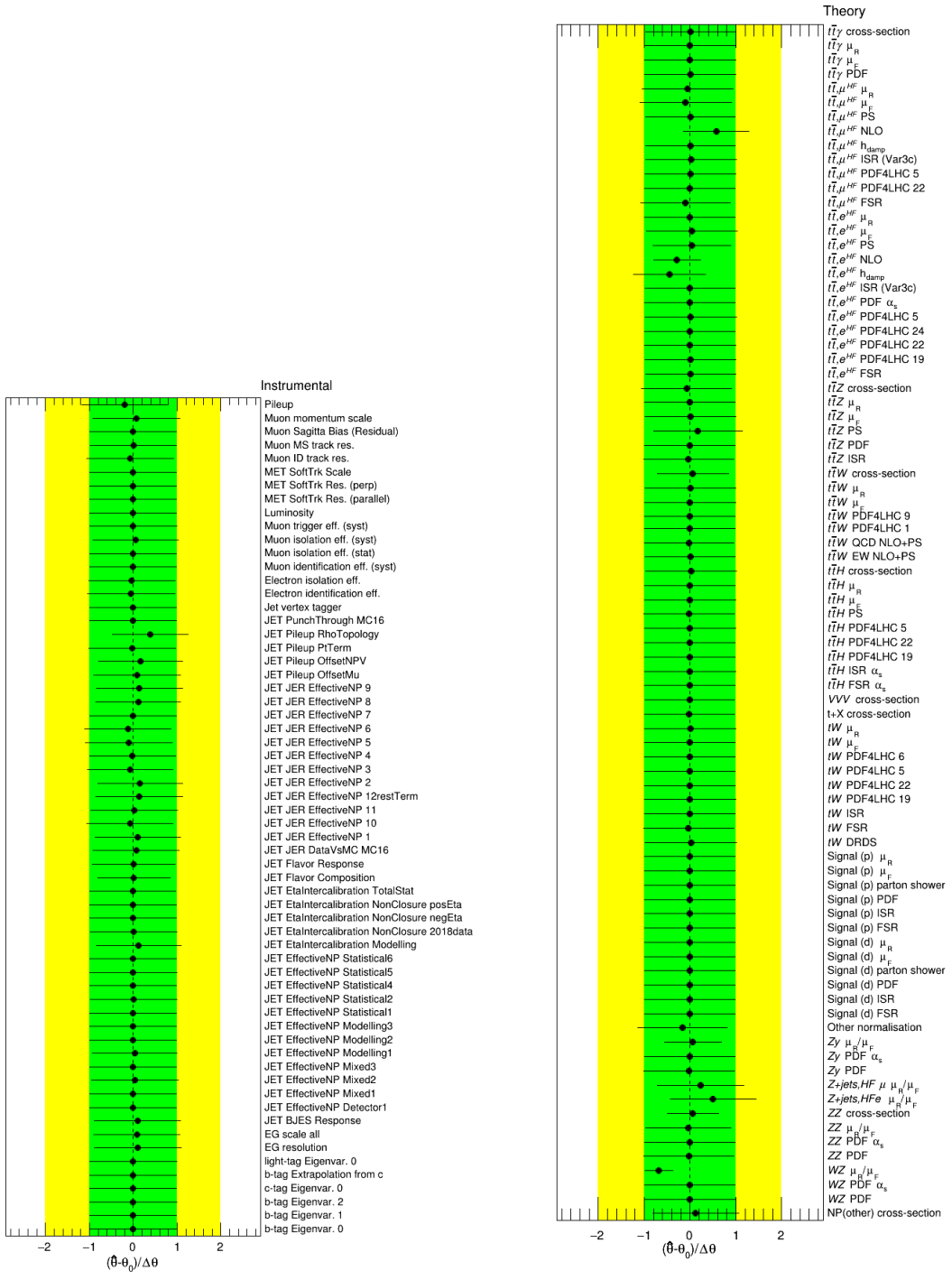


Figure 5.31: Correlation matrix (in %) of the NPs obtained from the fit using Asimov data in the SRs. A threshold of 25% has been set (so that NPs with lower correlation values do not appear) in order to make the plot more visible.

the value of the NP in order to compensate for relatively large differences between the data and MC prediction. For example, the $t\bar{t}$, HFμ NLO matching NP (shown in fig. 5.22e) is pulled to a larger value, in an attempt to decrease the level of that background. This is related to the data/MC discrepancy observed fig. 5.25e, where MC prediction lies above the data in the low H_{T} bins. The scale (μ_{R}, μ_{F}) NPs act in the opposite direction, while the respective NF, $k(t\bar{t}, \mathrm{HF}\mu)$ slightly increases the total normalisation. For $t\bar{t}$, HFe, the NLO matching and the h_{damp} NPs are pulled to smaller values, again used by the fit to correct for data/MC discrepancies mainly in the control region CR$t\bar{t}e$.

Pulls are also seen for the largest (scale) systematic uncertainties on $Z+$jets. Plots showing $\pm1\sigma$ variations of each of these nuisance parameters are shown in figs. 5.23a and 5.23b. As in the $t\bar{t}$ case, these are highly correlated with the respective NFs, $k(Z + \mathrm{HF}e/\mu)$, so the fit adjusts their values in order to compensate for the data/MC differences in the respective CRs. The scale uncertainty on $Z\gamma$ is constrained by the fit, using the high-purity CR$Z\gamma$ region.

The largest systematic uncertainty, μ_R/μ_F, for WZ is constrained as well as being strongly pulled. This can be explained as this uncertainty has a large shape component, as can be seen in fig. 5.23c,



(a) (b)

Figure 5.32: Nuisance parameter pull plot obtained from the fit using Asimov data in the SRs, showing how the values of the nuisance parameters and their uncertainties changed relative to those before the fit. The error bars correspond to the post-fit uncertainty, $\Delta\hat{\theta}$, which is to be compared to the pre-fit $\Delta\theta = 1$ uncertainty. To improve visibility, the $t\bar{t}W$ PDF set NPs, for which no constraint or pull is observed, are omitted.

increasing with H_{T}. Since a data/MC slope exists in the H_{T} distribution fitted in CRWZ (fig. 5.25a), this scale uncertainty is correlated with $k(WZ)$, as seen in fig. 5.31, and its slope is used by the fit to correct the data/MC discrepancy in CRWZ.

The result in the VRs is significantly affected by the NF values obtained by the fit. The $Z\gamma$ normalisation is increased by the fit, driven by CR$Z\gamma$, and is able to improve the data/MC discrepancy in VRe (fig. 5.37e). The data/MC agreement in the VRs improves, even though VRe already has a large pre-fit discrepancy which cannot be compensated fully.

5.9.3 Unblinded fit

Since the "mixed" fit using data (in CRs) and Asimov data (in SRs) discussed in the previous section (referred to as the "Asimov" fit in the following) is stable and the pulls on the nuisance parameters can be understood as a means to compensate the data/MC discrepancies in the CRs, the statistical analysis strategy is validated. This justifies proceeding to an *unblinding*[26] of the SRs, i.e. the use of real data in the SRs.

Pre- and post-fit plots of the analysis regions included in the fit are shown in figs. 5.33 to 5.36; fig. 5.35 and fig. 5.36 show the same pre/post-fit plots in the signal regions, but with different scale, logarithmic and linear, in order to better display the signal contribution (in the former) and the background contributions (in the latter).

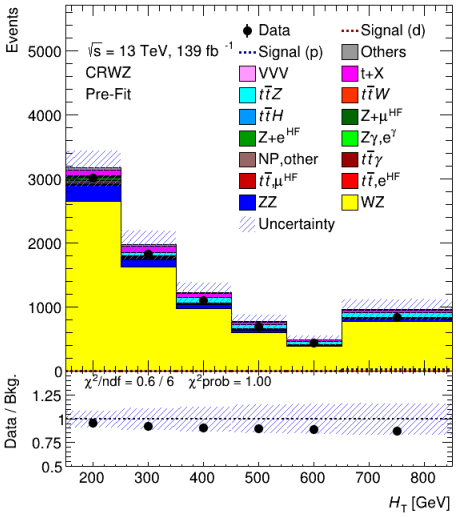
The normalisation factors obtained by the fit are shown in table 5.17. The signal strength parameter obtained from the fit is compatible with zero and hence the absence of signal. The slightly negative value, which is allowed by the fit as discussed in section 5.8, can be explained by an excess in the MC prediction of the background compared to data in the last, and most signal-sensitive, bin of the SR-production region (fig. 5.36c).

μ_{CLFV}	$k(WZ)$	$k(Z\gamma(\gamma \to e^+e^-))$	$k(Z+HF\mu)$	$k(Z+HFe)$	$k(t\bar{t},HF\mu)$	$k(t\bar{t},HFe)$
-0.0023 ± 0.0061	0.872 ± 0.074	1.32 ± 0.18	0.96 ± 0.10	0.94 ± 0.10	1.02 ± 0.16	1.01 ± 0.14

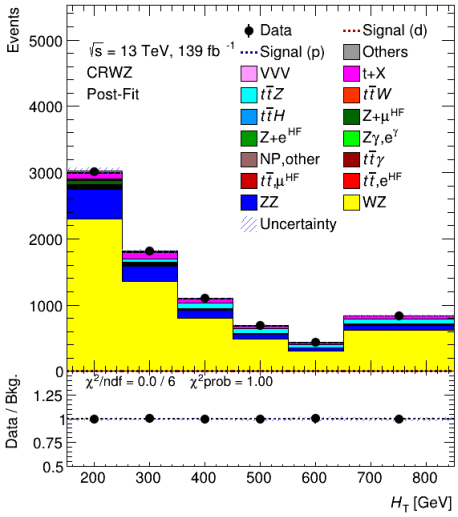
Table 5.17: Table showing the normalisation factors obtained from the unblinded fit.

The parameters extracted from the fit are also applied to the VRs, resulting in the plots shown in fig. 5.37. Summary plots are collected in figs. 5.38 and 5.39, and post-fit yields by process in each analysis region are collected in table 5.18 (the pre-fit yields have already been presented in table 5.15).

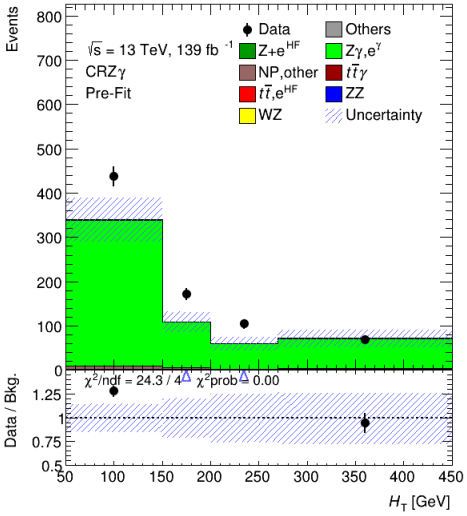
[26] This is an *unofficial* unblinding for the purposes of this thesis. The results are not yet published.



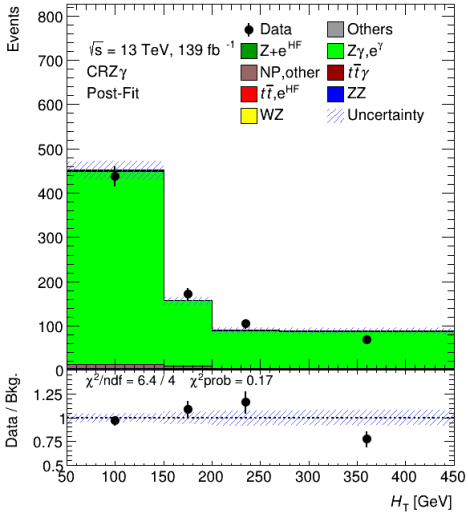
(a) Pre-fit H_{T} distribution in CRWZ.



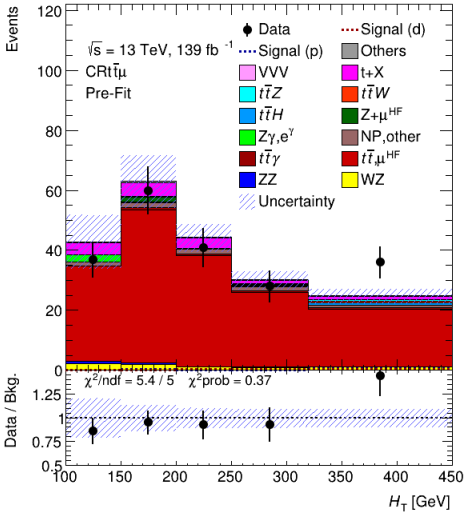
(b) Post-fit H_{T} distribution in CRWZ.



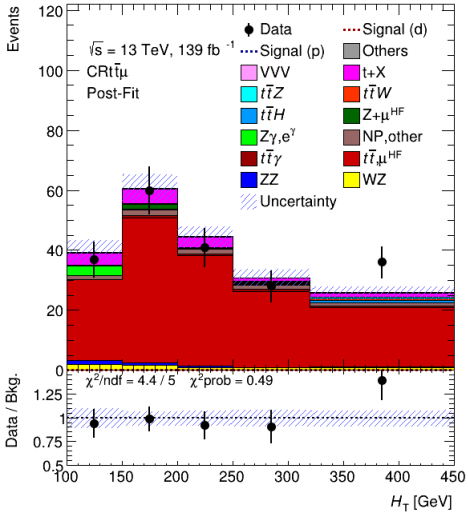
(c) Pre-fit H_{T} distribution in CR$Z\gamma$.



(d) Post-fit H_{T} distribution in CR$Z\gamma$.

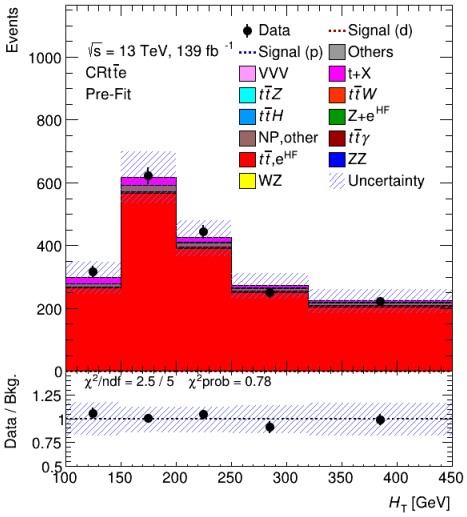


(e) Pre-fit H_{T} distribution in CR$t\bar{t}\mu$.

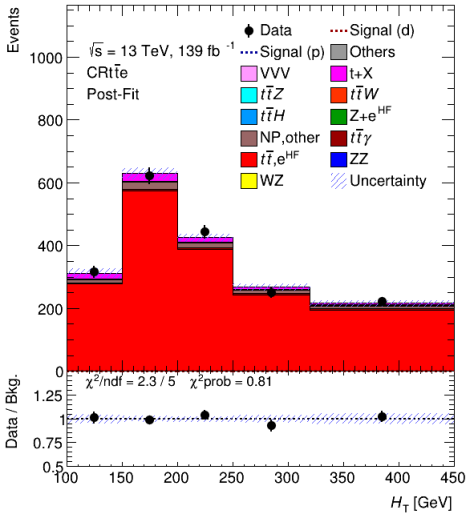


(f) Post-fit H_{T} distribution in CR$t\bar{t}\mu$.

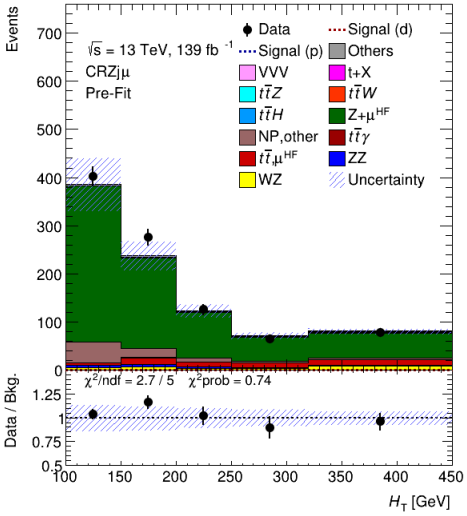
Figure 5.33: Pre-fit (figs. 5.33a, 5.33c and 5.33e) and post-fit (figs. 5.33b, 5.33d and 5.33f) plots of CRWZ, CR$Z\gamma$ and CR$t\bar{t}\mu$, for the unblinded fit.



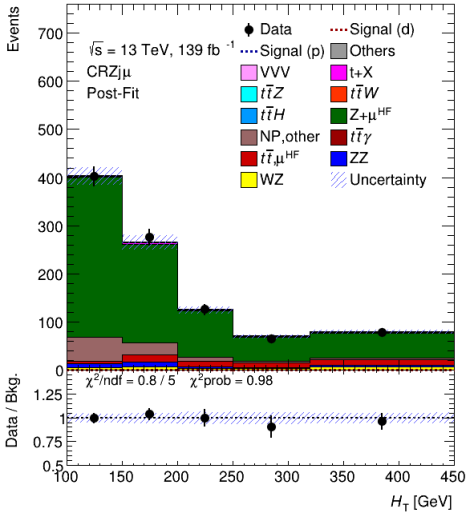
(a) Pre-fit H_{T} distribution in CR$t\bar{t}e$.



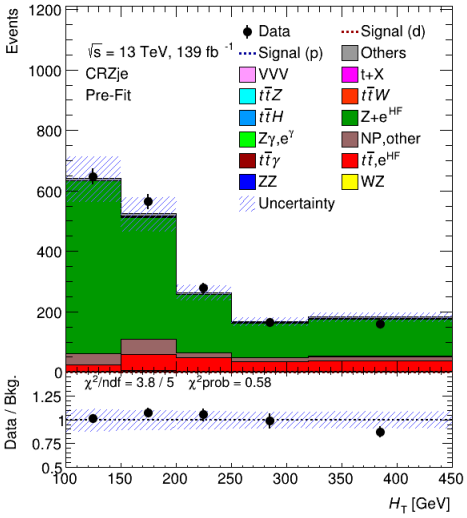
(b) Post-fit H_{T} distribution in CR$t\bar{t}e$.



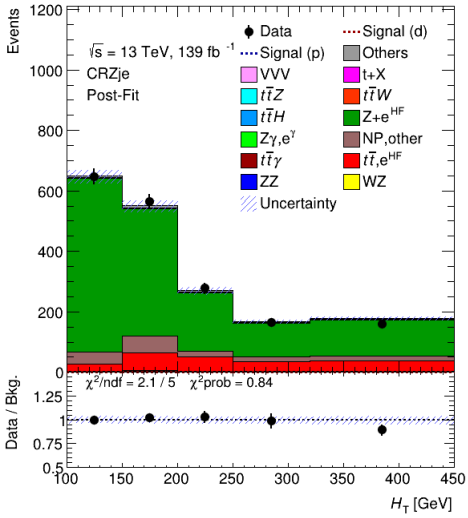
(c) Pre-fit H_{T} distribution in CR$Z\mathrm{j}\mu$.



(d) Post-fit H_{T} distribution in CR$Z\mathrm{j}\mu$.

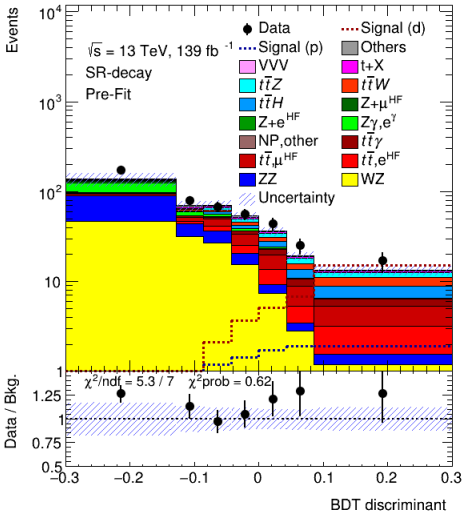


(e) Pre-fit H_{T} distribution in CR$Z\mathrm{j}e$.

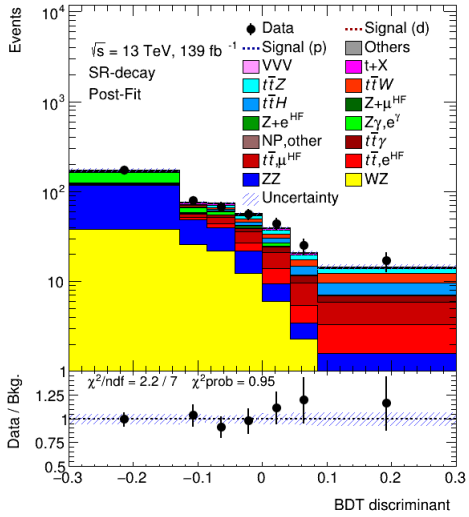


(f) Post-fit H_{T} distribution in CR$Z\mathrm{j}e$.

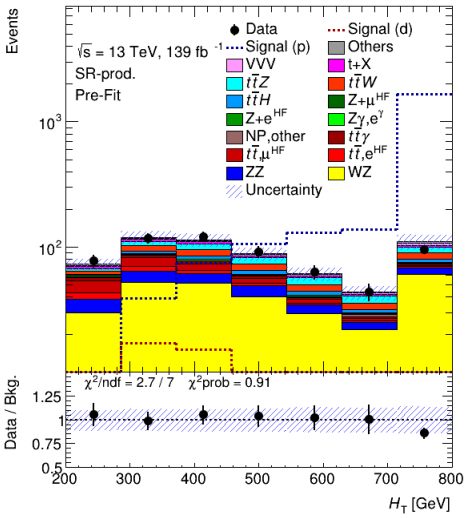
Figure 5.34: Pre-fit (figs. 5.34a, 5.34c and 5.34e) and post-fit (figs. 5.34b, 5.34d and 5.34f) plots of CR$t\bar{t}e$, CR$Z\mathrm{j}\mu$, CR$Z\mathrm{j}e$, for the unblinded fit.



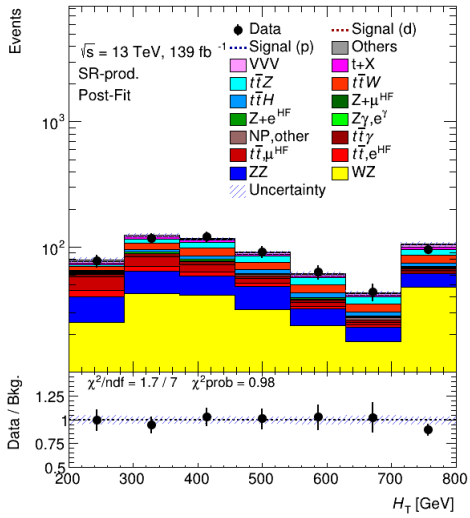
(a) Pre-fit BDT score distribution in SR-decay.



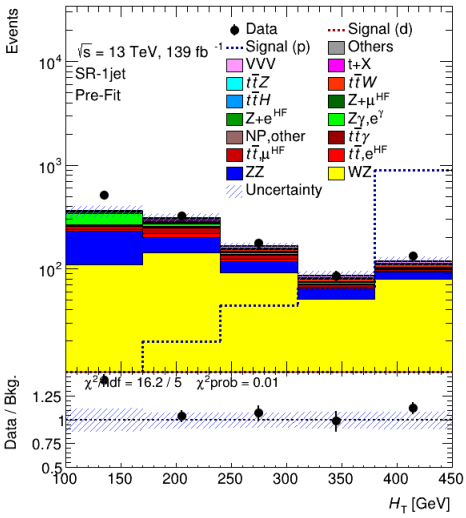
(b) Post-fit BDT score distribution in SR-decay.



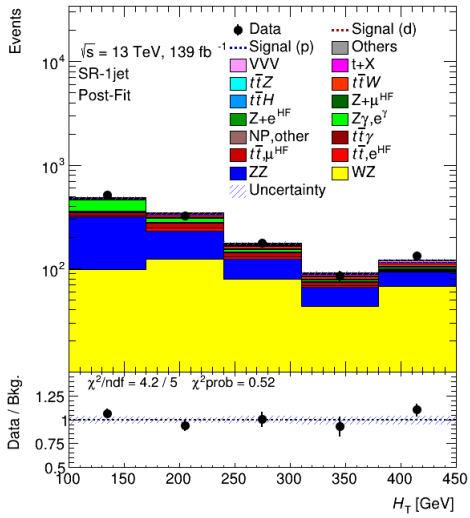
(c) Pre-fit H_{T} distribution in SR-production.



(d) Post-fit H_{T} distribution in SR-production.

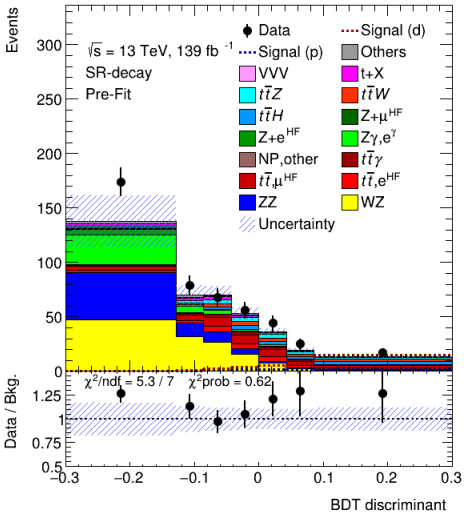


(e) Pre-fit H_{T} distribution in SR-1jet.

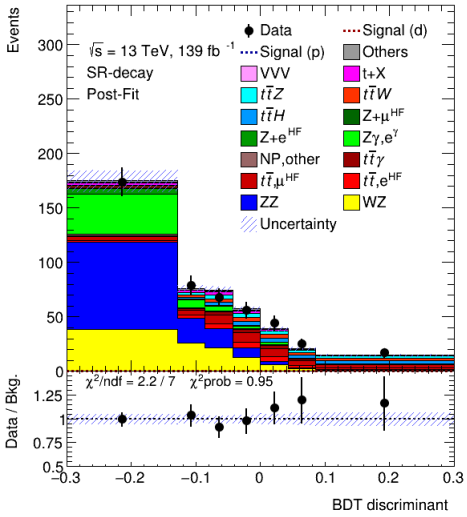


(f) Post-fit H_{T} distribution in SR-1jet.

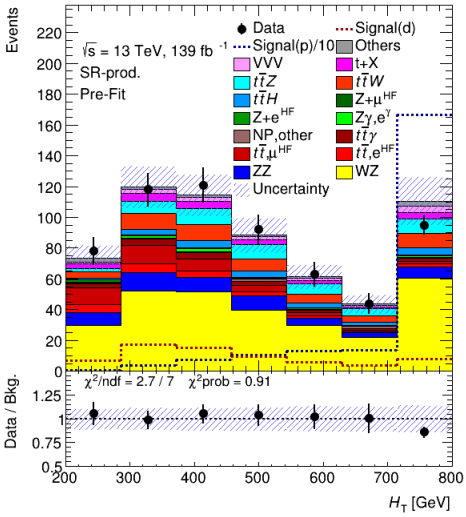
Figure 5.35: Pre-fit (figs. 5.35a, 5.35c and 5.35e) and post-fit (figs. 5.35b, 5.35d and 5.35f) plots of SR-decay, SR-production, SR-1jet (in logarithmic scale), for the unblinded fit.



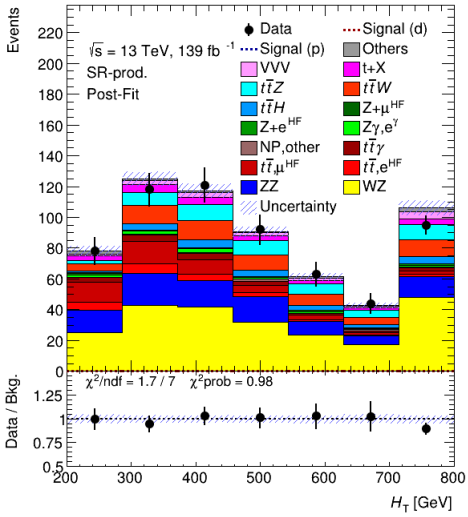
(a) Pre-fit BDT score distribution in SR-decay.



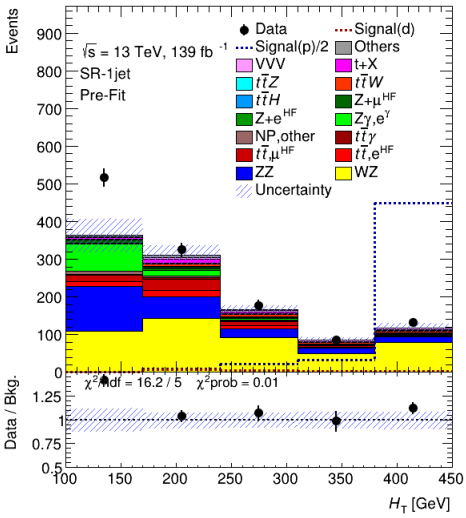
(b) Post-fit BDT score distribution in SR-decay.



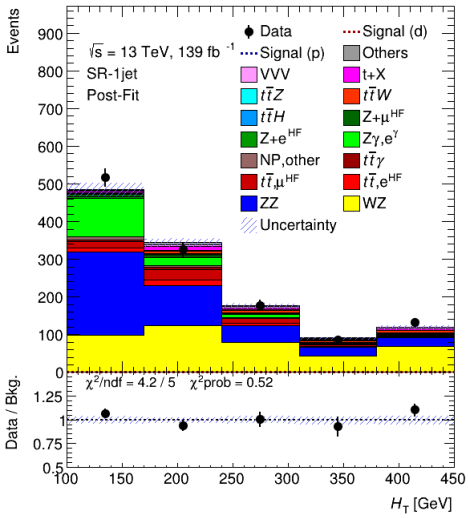
(c) Pre-fit H_{T} distribution in SR-production.



(d) Post-fit H_{T} distribution in SR-production.

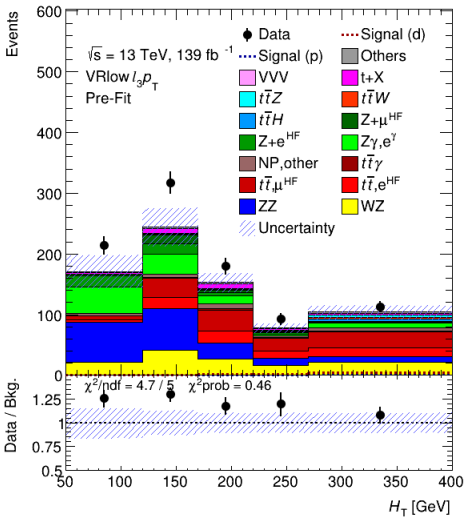


(e) Pre-fit H_{T} distribution in SR-1jet.

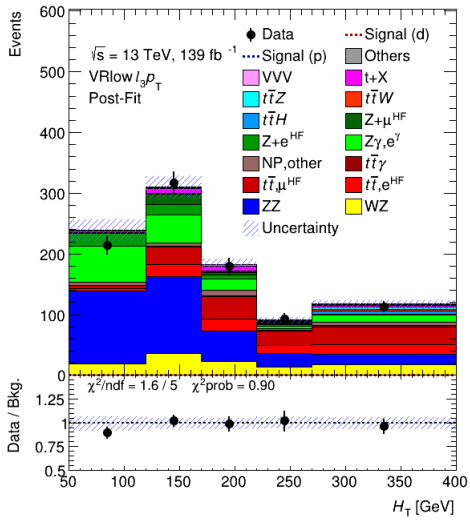


(f) Post-fit H_{T} distribution in SR-1jet.

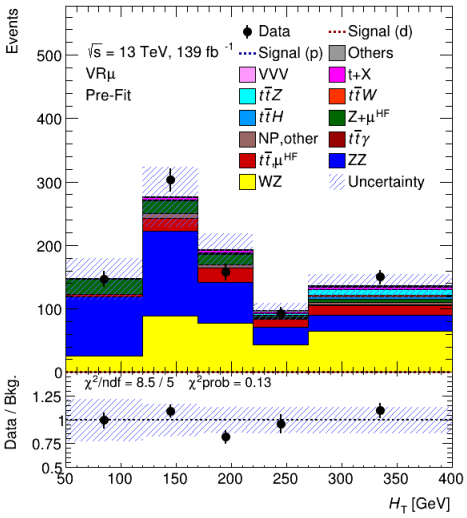
Figure 5.36: Pre-fit (figs. 5.36a, 5.36c and 5.36e) and post-fit (figs. 5.36b, 5.36d and 5.36f) plots of SR-decay, SR-production, SR-1jet (in linear scale), for the unblinded fit. The production-signal has been scaled in the pre-fit plots, in SR-prod. and SR-1jet, for appearance purposes.



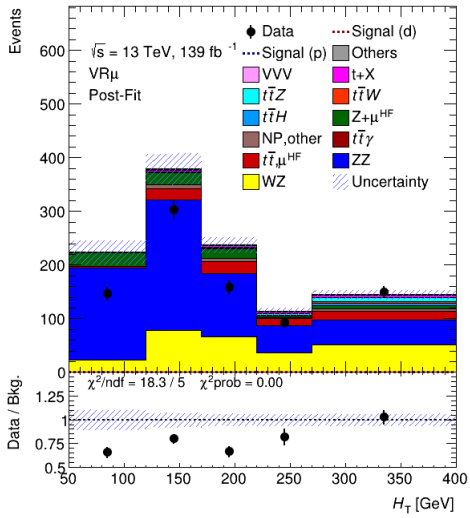
(a) Pre-fit H_{T} distribution in VR$low\ell_3p_{\mathrm{T}}$.



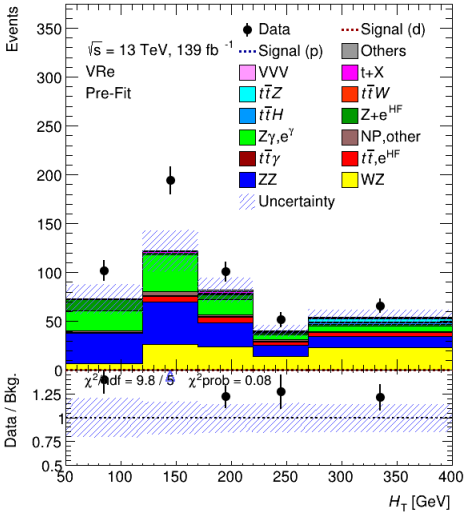
(b) Post-fit H_{T} distribution in VR$low\ell_3p_{\mathrm{T}}$.



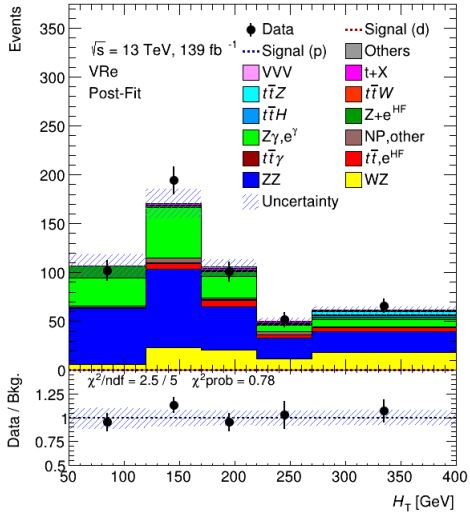
(c) Pre-fit H_{T} distribution in VRμ.



(d) Post-fit H_{T} distribution in VRμ.

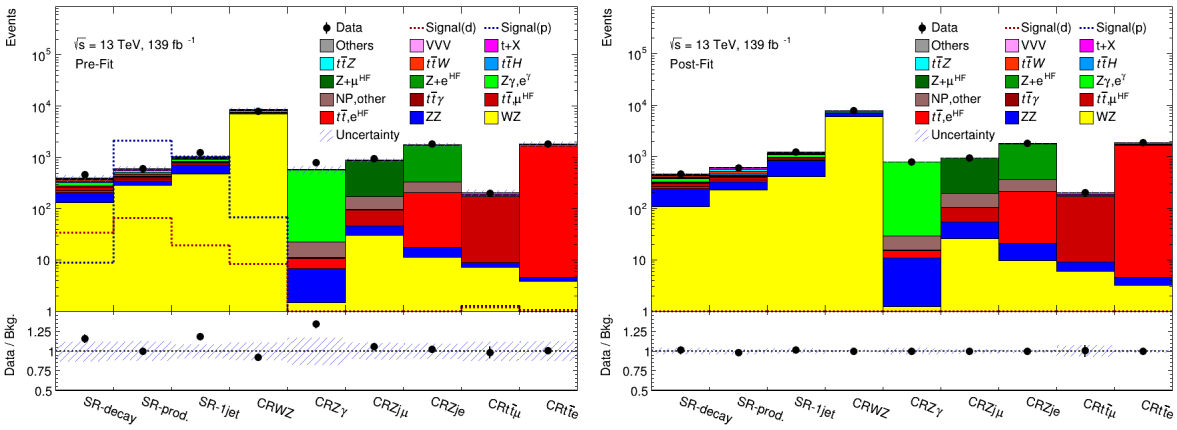


(e) Pre-fit H_{T} distribution in VRe.



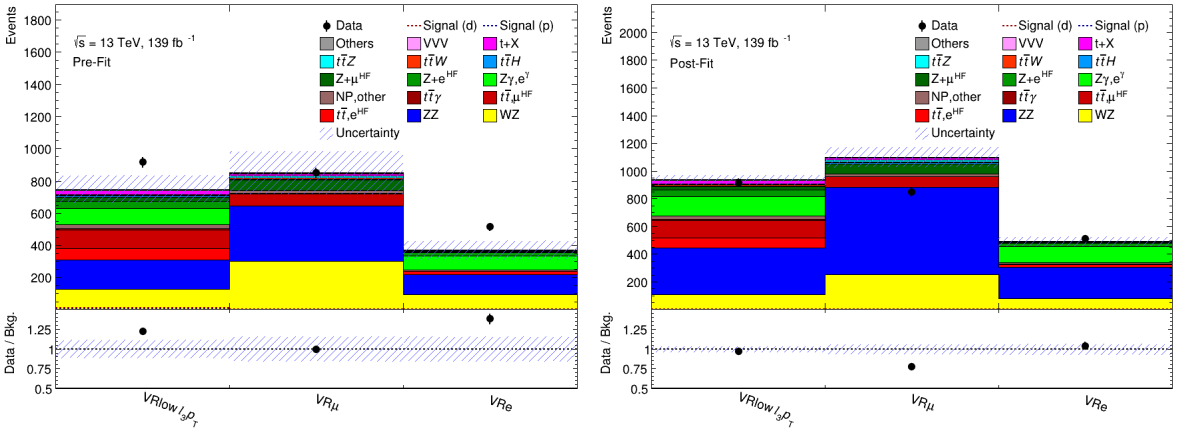
(f) Post-fit H_{T} distribution in VRe.

Figure 5.37: Pre-fit (figs. 5.37a, 5.37c and 5.37e) and post-fit (figs. 5.37b, 5.37d and 5.37f) plots of VR$low\ell_3p_{\mathrm{T}}$, VRμ, VRe, for the unblinded fit.



(a) (b)

Figure 5.38: Pre/post-fit summary plot of SRs and CRs using the unblinded fit.



(a) (b)

Figure 5.39: Pre/post-fit summary plot of VRs using the unblinded fit.

	SR-decay	SR-prod.	SR-1jet	VR$low\ell_3 p_{\mathrm{T}}$	VRμ	VRe	CRWZ	CR$Z\gamma$	CRZjμ	CRZje	CR$t\bar{t}\mu$	CR$t\bar{t}e$
Signal (d)	-0.080 ± 0.003	-0.16 ± 0.01	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Signal (p)	0.0	-5.2 ± 0.3	-2.4 ± 0.1	0.0	0.0	0.0	-0.16 ± 0.01	0.0	0.0	0.0	0.0	0.0
Others	5.3 ± 3.1	9.3 ± 5.1	12.4 ± 7.0	6.8 ± 3.8	4.2 ± 2.4	2.6 ± 1.4	51 ± 28	3.0 ± 1.6	8.3 ± 4.5	21 ± 11	0.9 ± 0.5	1.1 ± 0.6
VVV	3.6 ± 1.7	16.2 ± 7.6	20.2 ± 9.4	2.0 ± 0.9	2.3 ± 1.1	0.8 ± 0.4	34 ± 16	0.0	0.1 ± 0.1	0.0	0.3 ± 0.1	0.2 ± 0.1
$t+X$	11.6 ± 2.2	23.1 ± 3.5	23.4 ± 4.5	23.3 ± 3.2	15.1 ± 3.4	4.9 ± 1.1	350 ± 50	0.0	6.7 ± 1.3	10.2 ± 3.7	14.8 ± 4.4	73 ± 11
$t\bar{t}Z$	17.8 ± 4.4	51.5 ± 9.3	8.3 ± 2.3	6.5 ± 2.5	12.0 ± 3.8	4.5 ± 1.9	346 ± 53	0.0	1.7 ± 0.3	2.9 ± 0.7	0.9 ± 0.3	2.1 ± 1.2
$t\bar{t}W$	19.5 ± 7.2	62 ± 23	27 ± 10	6.6 ± 2.5	6.6 ± 2.7	2.7 ± 1.0	17.0 ± 6.1	0.0	0.1 ± 0.1	0.2 ± 0.1	1.3 ± 0.5	2.9 ± 1.1
$t\bar{t}H$	17.8 ± 2.0	24.2 ± 2.6	3.8 ± 0.5	6.1 ± 0.7	5.8 ± 0.6	2.2 ± 0.3	10.3 ± 1.1	0.0	0.10 ± 0.02	0.16 ± 0.05	1.3 ± 0.1	2.4 ± 0.4
Z+HFμ	1.3 ± 0.5	1.6 ± 0.8	14.9 ± 2.4	25.4 ± 9.3	72 ± 16	0.0	85 ± 11	0.0	735 ± 44	0.0	1.9 ± 0.9	0.0
Z+HFe	5.5 ± 3.0	4.2 ± 0.5	10.3 ± 5.3	47.1 ± 6.0	0.0	22.1 ± 3.6	30.9 ± 6.1	2.3 ± 0.5	0.0	1419 ± 65	0.0	2.3 ± 1.8
$Z\gamma$,e^{γ}	53.3 ± 4.9	9.0 ± 0.7	136.5 ± 9.6	139 ± 10	0.0	118 ± 10	24.2 ± 2.9	754 ± 29	0.0	2.5 ± 1.2	3.8 ± 0.4	0.0
NP,other	4.6 ± 4.9	5.0 ± 1.8	16.5 ± 7.0	28 ± 10	19.9 ± 7.3	10.1 ± 4.4	40 ± 16	13.7 ± 5.6	91 ± 33	146 ± 54	7.3 ± 2.6	71 ± 26
$t\bar{t}\gamma$	14.6 ± 2.3	18.5 ± 2.9	13.4 ± 2.1	7.7 ± 1.2	2.4 ± 0.4	2.9 ± 0.5	6.7 ± 1.0	0.5 ± 0.2	0.6 ± 0.1	2.3 ± 0.4	2.9 ± 0.4	19.9 ± 3.2
$t\bar{t}$,HFμ	38.7 ± 6.2	47 ± 12	63.6 ± 8.0	125 ± 15	75.2 ± 9.1	0.0	38 ± 11	0.0	50.1 ± 6.1	0.0	157 ± 15	0.0
$t\bar{t}$,HFe	21.7 ± 2.4	22.4 ± 1.9	37.7 ± 5.2	70.3 ± 7.2	0.0	20.8 ± 2.2	15.1 ± 2.1	4.2 ± 0.7	0.0	191.4 ± 9.8	0.0	1675 ± 51
ZZ	135 ± 18	98 ± 13	420 ± 41	338 ± 34	629 ± 67	223 ± 26	969 ± 123	9.6 ± 1.3	28.9 ± 3.1	11.7 ± 1.3	3.1 ± 0.4	1.4 ± 0.2
WZ	107.4 ± 5.6	230 ± 10	413 ± 20	109.3 ± 4.7	255 ± 12	81.1 ± 5.1	5892 ± 174	1.2 ± 0.1	26.2 ± 1.2	9.6 ± 0.6	6.1 ± 0.3	3.3 ± 0.2
Tot.Bkg.	458 ± 16	622 ± 21	1220 ± 33	940 ± 41	1100 ± 72	495 ± 36	7908 ± 89	789 ± 28	949 ± 31	1817 ± 43	201 ± 14	1854 ± 43
Data	463	611	1238	917	851	515	7906	784	948	1815	202	1855

Table 5.18: Post-fit yields from the unblinded fit. The errors quoted include MC statistical and systematic uncertainties.

5.9.3.1 Correlations, constraints and pulls of nuisance parameters

The correlation matrix obtained from the unblinded fit, fig. 5.40, is compatible with the one obtained from the Asimov fit (fig. 5.31), apart from the ZZ NPs (particularly the ZZ cross-section NP) which appear more correlated to each other and to the WZ NF and NPs. This can be understood by the data/MC discrepancy observed in the first bins of SR-decay and SR-1jet regions, as discussed below.

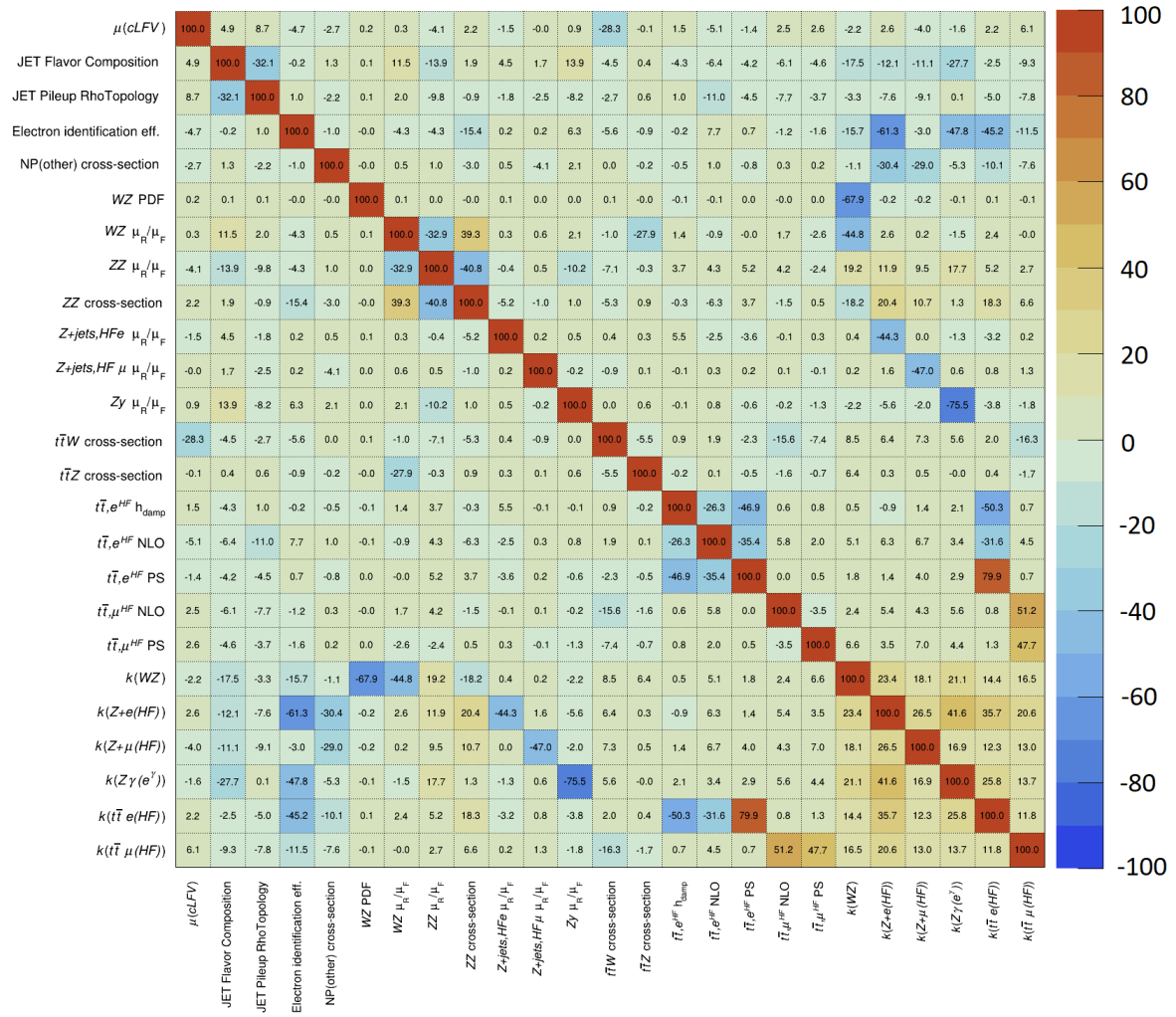
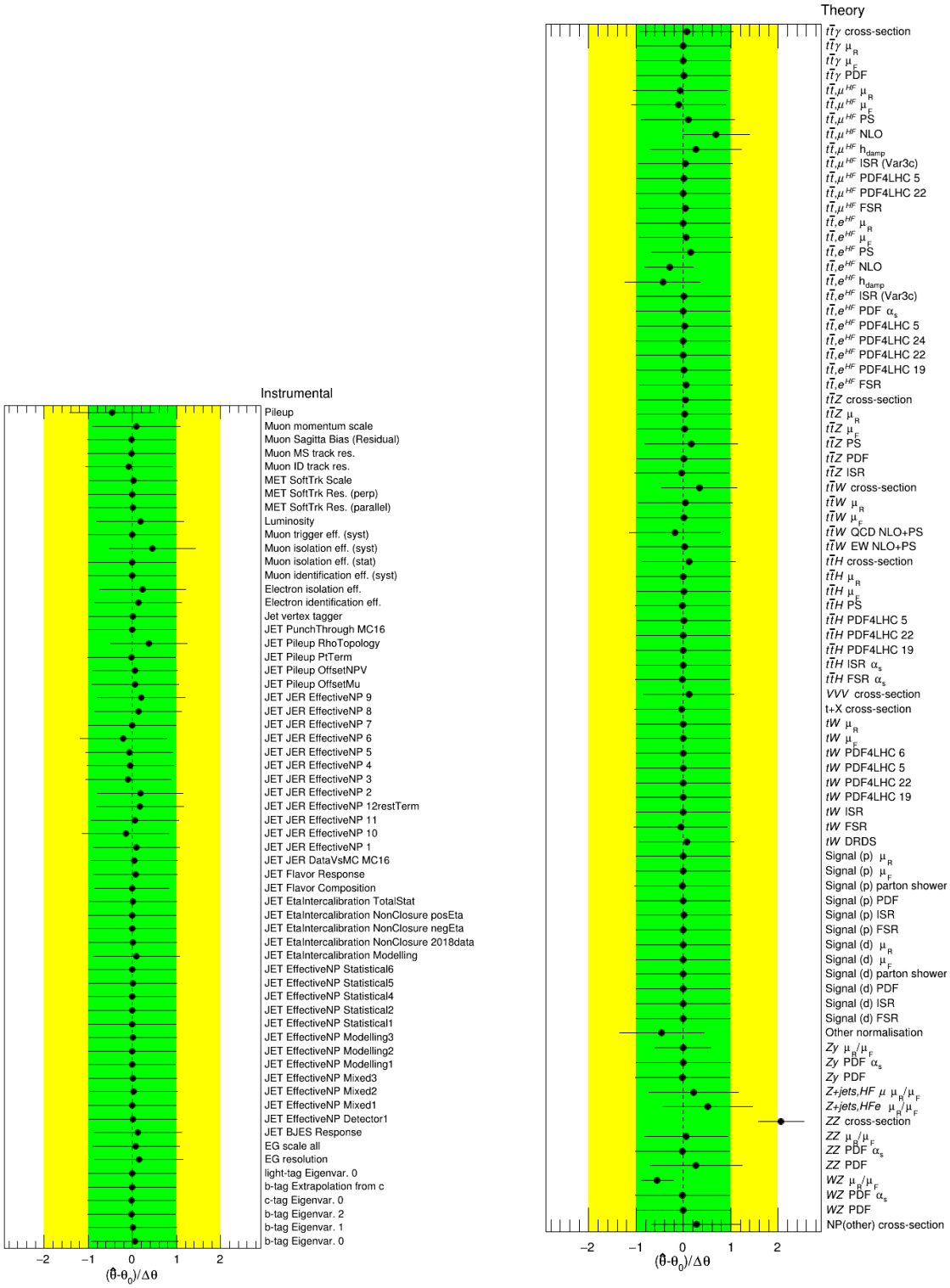


Figure 5.40: Correlation matrix (in %) of the NPs obtained from the unblinded fit. A threshold of 25% has been set (so that NPs with lower correlation values do not appear) in order to make the plot more visible.

The NP pull plot for the unblinded fit (fig. 5.41), preserves the pulls and constraints that are seen in the Asimov fit (fig. 5.32), which are focused on NPs that are used by the fit to match the MC prediction to the data in the control regions.

Additional pulls are observed on the normalisation uncertainty (cross-section NP) of $t\bar{t}W$ and *Other* backgrounds. The fit adjusts their contribution in the SRs, particularly SR-production, to improve the data/MC agreement.

A notable additional pull is observed on the ZZ background cross-section. For that background,



(a) (b)

Figure 5.41: Nuisance parameter pull plot obtained from the unblinded fit, showing how the values of the nuisance parameters and their uncertainties changed relative to those before the fit. The error bars correspond to the post-fit uncertainty, $\Delta\hat{\theta}$, which is to be compared to the pre-fit $\Delta\theta = 1$ uncertainty. To improve visibility, the $t\bar{t}W$ PDF set NPs, for which no constraint or pull is observed, are omitted.

the large disagreement between data and MC in the first bins of SR-decay and SR-1jet (fig. 5.36a and fig. 5.36e), which are signal-depleted and essentially act as "control regions", is compensated by a considerable increase (hence the large positive pull in fig. 5.41) in the ZZ background normalisation. Since other backgrounds present in these bins, namely WZ and $Z\gamma$, are already constrained in terms of normalisation by their respective control regions, the ZZ background becomes responsible for fixing the remaining data/MC disagreement (for a discussion refer to section 5.9.5).

5.9.4 Ranking of nuisance parameters

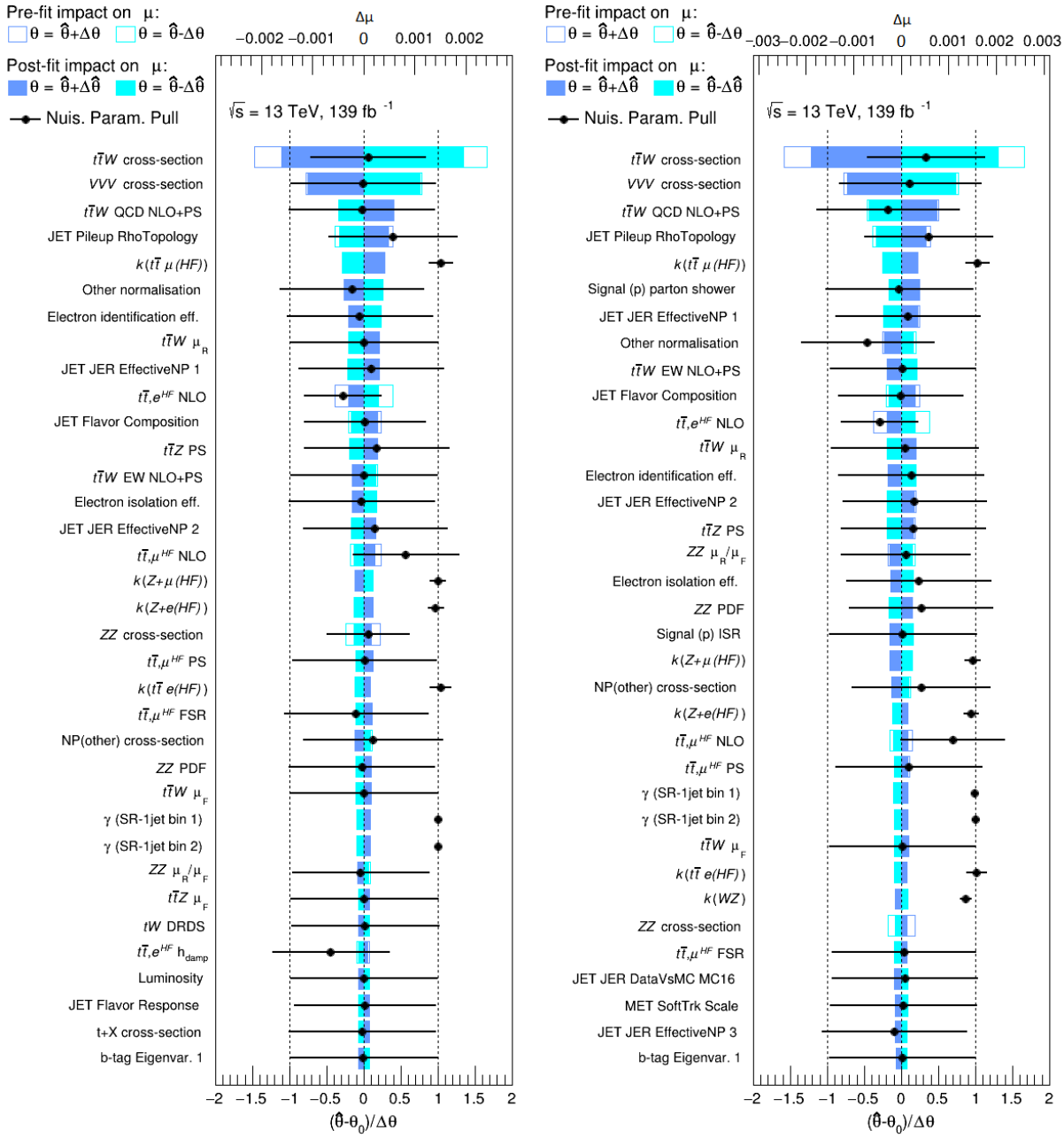
The NP ranking is based on the impact a specific NP has on the estimated value of the POI, given by the shift of the POI value between the nominal fit and another fit in which that nuisance parameter is fixed to a value of $\hat{\theta} \pm x$. The value $\hat{\theta}$ is the MLE (or "post-fit" value) of the NP, and:

- $x = \Delta\theta = 1$ for a *pre-fit impact* estimation. For unconstrained nuisance parameters the pre-fit impact is not defined.
- $x = \Delta\hat{\theta}$ for a *post-fit impact* estimation. $\Delta\hat{\theta}$ denotes the uncertainty on $\hat{\theta}$. Since NPs can be constrained by the fit, the post-fit impact may be smaller than the pre-fit impact.

The ranking of nuisance parameters obtained by the Asimov fit, fig. 5.42a, and the unblinded fit, fig. 5.42b, is consistent with the sizes of the correlations of each nuisance parameter with μ_{CLFV}.

As seen in fig. 5.42, the modelling of the $t\bar{t}W$ and VVV backgrounds are leading contributions. These backgrounds have a considerable presence in the upper, and signal-sensitive, bins of the SR-production region. Also, as discussed in section 5.7.3, a large 50% normalisation uncertainty is assigned to the cross-sections of these processes to account for the differences observed between expected and measured contributions in other ATLAS multi-lepton analyses. The combined generator uncertainty on the $t\bar{t}W$ background also has a large impact on μ_{CLFV} with a normalisation component of more than 10% (for the QCD component) in the most sensitive SRs (SR-prod. and SR-1jet) (as shown in figs. 5.24a and 5.24b). The cross sections of other (SM) processes, which are assigned a conservative 50% normalisation uncertainty as mentioned in section 5.7.3, also feature in the ranking plot. $k(t\bar{t}, \mathrm{HF}\mu)$ is the highest ranked NF; $t\bar{t}$ is the largest non-prompt background in the SRs, and CR$t\bar{t}\mu$ offers a lower-statistics subsidiary measurement of $t\bar{t}$, HFμ compared to the electron counterpart. The electron identification efficiency also ranks highly, while the modelling of $t\bar{t}$, HFe background enters with the NLO (ME to PS) matching correction uncertainty (the effect shown in fig. 5.22f).

The production-signal parton shower uncertainty, with as high as a 6% pre-fit normalisation change in the SR-production region, enters as the sixth highest ranked NP for the unblinded fit. In the Asimov fit, the Asimov data in the SRs match the predicted background (modified by the CR-fit, as discussed in section 5.9.2), so no signal is expected, and hence the signal uncertainties are not expected to have any large impact on the measured signal strength. The production-signal channel offers more events than the decay-signal, and hence dictates the sensitivity for the limit setting.



(a) NP ranking plot, for the Asimov fit.

(b) NP ranking plot, for the unblinded fit.

Figure 5.42: Ranking of NPs obtained from the fit using Asimov data in the SRs (a) and the unblinded fit (b). They correspond to the sources of systematic uncertainty with the largest impact on the fitted signal strength μ_{CLFV}. The points, drawn according to the bottom axis, show the deviation of each of the fitted NPs from its nominal value (zero for NPs related to systematic uncertainties and one for NFs and *gammas*) in units of the pre-fit standard deviation $\Delta\theta$; the ZZ cross section in (b) is pulled further than the range, as was seen in fig. 5.41. The error bars show the post-fit uncertainties $\Delta\hat{\theta}$, which have a size close to one if the data do not provide any further constraint on that uncertainty. Conversely, an error bar smaller than one indicates a reduction with respect to the original uncertainty. The hollow blue bar represents the pre-fit impact on μ_{CLFV} and the filled area its post-fit impact. The NPs are sorted according to their post-fit impact (according to the top axis).

5.9.5 Discussion of fit results

Even though the pulls in the diboson-related NPs are large and can be considered as a way for the fit to match the data with the MC prediction in all regions (SRs and CRs), they are not expected to affect the final limit setting on the signal process significantly. This is because the most signal-sensitive bins are the last few bins in the SRs, where the diboson prediction is reduced compared to the first bins. Notably, in SR-prod. (fig. 5.36c), which is the most signal-sensitive region, and SR-1jet (fig. 5.36e), the second most sensitive region, the last bins present a reasonable (pre-fit) agreement between the data and the predicted background, and the presence of ZZ background is relatively small. So even if the ZZ normalisation is inflated by the fit, it does not significantly affect the signal-sensitive bins.

The considerable ZZ background increase spoils the VRμ pre-fit data/MC agreement (fig. 5.37c). The non-prompt backgrounds are well estimated by the subsidiary measurements in the dedicated and pure CRs.

There are further plans to increase the fit sensitivity and stability, as well as to diminish the large pulls. The application of a wider Z-mass window cut in the SRs, as well as an $E_{\mathrm{T}}^{\mathrm{miss}}$ cut to further suppress ZZ background, would reduce the diboson backgrounds. The resulting SR-prod. and SR-1jet regions could then be merged. By increasing the number of bins in the resulting SR, the final limit setting is expected to be enhanced.

5.9.6 Limits on the CLFV process

The expected upper limit on the signal strength for the CLFV processes is computed using the CL$_s$ method, and using the Asimov dataset in the SRs under the hypothesis of the absence of signal, with the fit described in section 5.9.2. The expected signal strength upper limit at 95% CL is found to be

$$\mu_{\mathrm{CLFV}} < 0.013^{+0.005}_{-0.004} \,, \tag{5.33}$$

where the errors quoted correspond to $\pm 1\sigma$ variations in the value of $\hat{\mu}$ estimated by the fit. This error band is a measure of how much the upper limit would be expected to vary due to fluctuations in the data, if the experiment were repeated many times [259].

Equation (5.33) translates into an expected upper limit on the cross section of the (production and decay combined) signal processes of

$$\sigma_{\mathrm{CLFV}} < 3.6^{+1.5}_{-1.0}\ \mathrm{fb} \quad (95\%\ \mathrm{CL}) \,. \tag{5.34}$$

Using the prescription of section 5.9.1.3, we translate the limit on the cross-section into the following expected upper limits on the value of the Wilson coefficients and the branching ratio (BR) of the CLFV top quark decay:

$$|C| < 0.034^{+0.006}_{-0.005} \quad (95\%\ \mathrm{CL}) \,, \tag{5.35}$$

$$\mathcal{B}(t \to \ell\ell' q)^{\mathrm{stat.+syst.}}_{\mathrm{exp.}} < 4.2^{+1.8}_{-1.2} \times 10^{-7} \quad (95\%\ \mathrm{CL}) \,. \tag{5.36}$$

The sensitivity to μ_{CLFV} comes predominantly from the production signal, which dominates in statistics

over the decay signal, and is driven mainly by the SR-production region. The translation into a limit on the branching ratio of the top quark CLFV decay is based on the assumption that the same Wilson coefficients are accountable for the CLFV in top quark production and decay processes, both treated as contact interactions in the EFT regime.

Considering only statistical uncertainties, the expected branching ratio upper limit becomes:

$$\mathcal{B}(t \to \ell\ell' q)^{\text{stat.}}_{\text{exp.}} < 3.6^{+1.5}_{-1.0} \times 10^{-7} \quad (95\%\ \text{CL}) \,, \tag{5.37}$$

which suggests that the inclusion of systematic uncertainties has a modest impact on the extracted limit, causing a 17% increase.

The observed upper limit is computed using the unblinded fit (i.e. considering real data in the SRs), and is found to be

$$\mathcal{B}(t \to \ell\ell' q)^{\text{stat.+syst.}}_{\text{obs.}} < 3.7 \times 10^{-7} \quad (95\%\ \text{CL}) \,. \tag{5.38}$$

We use the notation $\ell\ell'$ for the BR, since it can be regarded as "inclusive". This BR includes all 3 leptonic modes in the CLFV final state ($e\mu$, $e\tau$, $\mu\tau$). This analysis though is sensitive to electrons and muons only, so it makes sense to recompute the BR after removing all generated signal events in which a (leptonically decaying) τ lepton is present at the CLFV vertex. The removal of $\ell\tau q$ events is performed by use of the MC generator "truth" records, and results in a $\sim 20\%$ reduction of the signal yield, due to the removal of events in which final-state e or μ leptons arise from leptonically decaying τ-leptons. An "exclusive" e/μ (i.e. no τ) upper limit can then be obtained, by normalising the selected events to $\frac{1}{3}$ of the total, "τ-inclusive", cross section of the generated CLFV samples, under the assumption of lepton universality.

The "$e\mu$-exclusive" expected upper limit on the top CLFV decay branching ratio is found to be

$$\mathcal{B}(t \to e\mu q)^{\text{stat.+syst.}}_{\text{exp.}} < 1.7^{+0.7}_{-0.5} \times 10^{-7} \quad (95\%\ \text{CL}) \,, \tag{5.39}$$

and the observed one is

$$\mathcal{B}(t \to e\mu q)^{\text{stat.+syst.}}_{\text{obs.}} < 1.6 \times 10^{-7} \quad (95\%\ \text{CL}) \,. \tag{5.40}$$

The observed limits are slightly tighter, but in agreement with the expected ones.

The limits set on the BR improve the previous limits set by ATLAS, which used a subset of the same data, corresponding to 80 fb^{-1} of pp collisions collected by ATLAS, and only the decay-signal process, as reported in [2]:

$$\mathcal{B}(t \to \ell\ell' q)^{\text{stat.+syst.}}_{\text{obs.}} < 1.86 \times 10^{-5} \quad (95\%\ \text{CL}) \,, \tag{5.41}$$

$$\mathcal{B}(t \to e\mu q)^{\text{stat.+syst.}}_{\text{obs.}} < 6.6 \times 10^{-6} \quad (95\%\ \text{CL}) \,. \tag{5.42}$$

5.10 Planned EFT interpretation

In this section we discuss the strategy for obtaining upper limits on the Wilson coefficients responsible for generating the signal processes in the context of EFT, presented in section 2.4. The work presented here is ongoing, and only expected limits will be reported using the Asimov fit configuration of section 5.9.2, since areas for improving the fit by adding selection cuts in the SRs are envisaged.

The EFT operators describing the (two quark)-(two lepton) vertex are flavour-dependent; a separate operator describes the decays $t \to e\mu u$ and $t \to e\mu c$. Twelve signal MC samples, corresponding to the six EFT operators discussed in section 2.4 split by quark flavour, have been generated for the production signal. In the case of the decay channel there are six distinct EFT MC samples which are inclusive of quark flavour. These may be further split according to quark flavour (u/c) by use of the MC generator "truth" records. Table 5.19 shows the split of the decay-signal events with u- and c-quarks in the CLFV vertex for events entering any of the signal regions. It is expected that CLFV top decay will produce up and charm quarks in similar proportions, with a small difference due to acceptance effects. For some events the truth record did not provide information, hence these events are labelled as "no match". For production signal, the c-quark component is suppressed compared to the u-quark, since it involves only sea (and not valence) quarks of the incoming protons.

The cross section of the CLFV process for each of the aforementioned "EFT samples" is proportional to the square of the value of the relevant Wilson coefficient for a given value of the scale of new physics, Λ[27]. The fit is re-run with one EFT MC sample replacing the inclusive MC signal sample to find the limit on the cross section which can then be translated into a limit on the Wilson coefficient.

	Yield	Fraction of matched
No match	0.77±0.02	-
u-quark	56.93±0.16	0.520
c-quark	52.64±0.15	0.480

Table 5.19: Table showing quark flavour in the CLFV vertex for decay mode signal events in the inclusive sample, corresponding to events entering any of the SRs.

Expected upper limits at 95% C.L., using the CL_s method, have been derived for the Wilson coefficients, by performing fits using the "mixed" data (in CRs) and Asimov data (in SRs) fit discussed in section 5.9.2. These expected limits are shown in table 5.20. These *expected* limits may be compared with the indirect limits in table 2.3, to get an idea about the constraining power of the fit.

$i, j = e, \mu$	$c_{lq}^{-(ijk3)}$	$c_{eq}^{(ijk3)}$	$c_{lu}^{(ijk3)}$	$c_{eu}^{(ijk3)}$	$c_{lequ}^{1(ijk3)}$	$c_{lequ}^{1(ij3k)}$	$c_{lequ}^{3(ijk3)}$	$c_{lequ}^{3(ij3k)}$
ATLAS expected (u)	0.18	0.18	0.17	0.17	0.18	0.18	0.044	0.044
ATLAS expected (c)	0.76	0.76	0.72	0.72	0.79	0.79	0.18	0.18

Table 5.20: Table showing the expected upper limits at 95% C.L. on the Wilson coefficients corresponding to (two lepton)-(two quark), $e\mu\tau u$ and $e\mu\tau c$, EFT operators, which introduce CLFV top decay in the $e\mu$ channel. The lu, eu, lq, eq couplings are *vector* couplings, while $lequ_1$ is a *scalar* and $lequ_3$ a *tensor* coupling, referring to the Lorentz structure of the respective EFT operators reported in table 2.2.

[27] Λ is chosen to be 1 TeV.

By combining the separate EFT MC signal samples corresponding to the four vector (lu, lq, eq, eu) EFT operators, we can set limits on the inclusive vector EFT couplings; the results are shown in table 5.21.

	Vector $\|C\|$ limit
ATLAS expected (u)	0.088
ATLAS expected (c)	0.37

Table 5.21: Table showing the expected (at 95 % C.L.) upper limits on the Wilson coefficient corresponding to the vector, i.e. combined eu, lu, lq, eq, (two lepton)-(two quark) EFT operator which introduces CLFV top decay in the $e\mu$ channel.

Plans discussed in section 5.9.5, which include additional cuts in the SRs, as well as finer binning, may improve our sensitivity to the EFT coefficients. When the unblinded results are obtained, a comparison can be made with the limits tabulated in table 2.4, set by CMS [67].

Chapter 6

Conclusion and outlook

One of the key purposes of the ATLAS experiment is the exploration of new physics, beyond the Standard Model. The possibility of the violation of charged-lepton flavour in top-quark decay modes is explored in this thesis, which uses data collected by ATLAS, from proton–proton collisions at the LHC at a centre-of-mass energy of $\sqrt{s} = 13$ TeV, corresponding to an integrated luminosity of 139 fb^{-1}.

For this search, the signal considered comprises two processes, charged-lepton flavour violating top-quark decay and production, $t\bar{t} \to (\ell\ell^{'} q)(W(\to \ell\nu)b)$ and $qg \to t\ell\ell^{'}$, where ℓ and $\ell^{'}$ are of different flavours. The leptons considered in the selected final state are muons and electrons, $\ell = \{e, \mu\}$, allowing also those coming from τ leptonic decays.

The decay (production) target topology is thus characterised by the presence of three isolated charged leptons, at least two (one) jets, and missing transverse momentum from the undetected neutrino. The analysis controls the most important background containing prompt leptons, coming from WZ dibosonic production, as well as backgrounds with non-prompt leptons, by the design of dedicated control regions. The signal processes are defined by signal regions, specially designed for enhancing the signal acceptance.

No evidence for a signal was found, and an observed (expected) upper limit on the $t \to \ell\ell^{'} q$ branching ratio has been set at the 95% confidence level

$$\mathrm{BR}(t \to \ell\ell' q) < 3.7(4.2) \times 10^{-7} \, . \tag{6.1}$$

With the exclusion of leptonically decaying τ-leptons from the CLFV final states, the upper limit becomes

$$\mathrm{BR}(t \to e\mu q) < 1.6(1.7) \times 10^{-7} \, . \tag{6.2}$$

The CLFV interaction vertex can be interpreted, in a model-independent way, as a contact interaction generated by (2-lepton)–(2-quark) operators of the Standard Model Effective Field Theory (EFT). A strategy is discussed for setting an upper-limit on the coupling coefficients (Wilson coefficients) of the relevant EFT operators.

The limits set may be enhanced by refinement of the selection of the signal regions and their binning. Future refinement of the analysis techniques, related to physics processes such as $t\bar{t}W$, which have a high impact on the sensitivity of this analysis, is expected to reduce the systematic uncertainties.

Of particular interest for this analysis will be a model-dependent interpretation using a *leptoquark* model to generate the signal processes. In that context, this analysis can be extended to provide limits on the relevant leptoquark masses and coupling strengths.

The ATLAS upgrade demands are touched upon in this thesis, by the description of the test module required for the commissioning of the newest electronic devices, the feature extractors (FEXs), that are

now responsible for the identification of leptons, jets and other physics objects, within the triggering system. The FEX modules come with their own digital complexities, one being the careful mapping of the location of the energies delivered to them, encoded in their firmware.

The FEX modules must continue coping with the increasing triggering demands when the High Luminosity LHC (HL-LHC) period takes place (planned for after 2028). The integrated luminosity of pp collisions during HL-LHC operation is expected to reach up to 4000 fb^{-1}. Analyses searching for exotic phenomena, and which are limited by the available statistics, such as the present one, are expected to benefit greatly from the increasing amount of data, either for setting even tighter limits, or for the discovery of new physics.

...
in the critical moments I will be finding again
my former ascetic spirit."

C.P. Cavafy, *The Dangerous* [1911]

References

[1] Carlo Alberto Gottardo, *Search for charged lepton-flavour violation in top-quark decays at the LHC with the ATLAS detector*, PhD thesis: Universität Bonn, 2019, URL: https://cds.cern.ch/record/2676841.

[2] ATLAS Collaboration, *Search for charged lepton-flavour violation in top-quark decays at the LHC with the ATLAS detector*, ATLAS-CONF-2018-044, 2018, URL: https://cds.cern.ch/record/2638305.

[3] David Griffiths, *Introduction to Elementary Particles*, 2nd, Wiley-VCH, 2008, ISBN: 978-3-527-40601-2.

[4] Francis Halzen and Alan D. Martin, *Quarks and Leptons*, Wiley & Sons, 1984, ISBN: 0-471-88741-2.

[5] Ta-Pei Cheng and Ling-Fong Li, *Gauge Theory of Elementary Particle Physics*, Oxford science publications, 1984, ISBN: 0-19-851961-3.

[6] Donald H. Perkins, *Introduction to High Energy Physics*, 4th, Cambridge University Press, 2000, ISBN: 0-521-62196-8.

[7] Ian J.R. Aitchison and Anthony J.G. Hey, *Gauge Theories in Particle Physics*, 4th, CRC Press, Taylor & Francis Group, 2013, ISBN: 9781315275253, URL: https://doi.org/10.1201/9781315275253.

[8] Mark Thomson, *Modern Particle Physics*, Cambridge University Press, 2013, ISBN: 978-1-107-03426-6.

[9] Wikipedia, *Standard Model*, accessed 21 October 2023, URL: https://en.wikipedia.org/wiki/Standard_Model.

[10] M. Tanabashi et al. (Particle Data Group), *Review of Particle Physics*, Phys. Rev. D **98**.030001 (2018), URL: https://pdg.lbl.gov/2018/.

[11] The Nobel Prize, *The Nobel Prize in Physics 2004*, accessed 21 October 2023, URL: https://www.nobelprize.org/prizes/physics/2004/popular-information/.

[12] Lorenzo Calibbi and Giovanni Signorelli, *Charged Lepton Flavour Violation: An Experimental and Theoretical Introduction*, Riv. Nuovo Cim. **41**.2 (2018) 1, arXiv: 1709.00294 [hep-ph].

[13] CDF Collaboration, *Observation of top quark production in $p\bar{p}$ collisions with the Collider-Detector at Fermilab*, Physical Review Letters **74**.14 (1995) 2626, arXiv: 9503002 [hep-ex].

[14] DØ Collaboration, *Observation of the Top Quark*, Physical Review Letters **74**.14 (1995) 2632, arXiv: 9503003 [hep-ex].

[15] R.L. Workman et al. (Particle Data Group), *Review of Particle Physics*, PTEP **2022** (2022) 083C01.

[16] The ATLAS Collaboration, *Top working group cross-section summary plots - June 2023*, accessed 22 October 2023, URL: https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PUBNOTES/ATL-PHYS-PUB-2023-014/.

[17] Geoffrey Gilles, *Top quark production at the LHC*, ATL-PHYS-SLIDE-2018-190, 2018, URL: https://cds.cern.ch/record/2314658/files/ATL-PHYS-SLIDE-2018-190.pdf.

[18] The CDF and DØ collaborations, *Combination of the top-quark mass measurements from the Tevatron collider*, Physical Review D **86** (2012).

[19] J. N. Bahcall, R. Davis, P. Parker, A. Smirnov, and R. Ulrich, *Solar neutrinos: The first thirty years*, CRC Press, Taylor & Francis Group, 2018, ISBN: 978-0-8133-4037-1.

[20] S.M.Bilenky, *Early years of neutrino oscillations*, Proceedings, International Meeting, Turin, Italy, March 25-27, 1999 (1999), arXiv: hep-ph/9908335 [hep-ph].

[21] S.M.Bilenky, *Neutrino in Standard Model and beyond*, (2015), arXiv: 1501.00232 [hep-ph], URL: https://doi.org/10.48550/arXiv.1501.00232.

[22] H1 Collaboration, *Search for Single Top Quark Production at HERA*, Phys.Lett.B **678**.450-458 (2009), arXiv: 0904.3876 [hep-ex], URL: https://doi.org/10.48550/arXiv.0904.3876.

[23] ATLAS Collaboration, *Probing lepton flavour violation via neutrinoless $\tau \to 3\mu$ decays with the ATLAS detector*, Eur. Phys. J. C **76** (2016) 232, arXiv: 1601.03567 [hep-ex].

[24] ATLAS Collaboration, *Search for the lepton flavor violating decay $Z \to e\mu$ in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector*, Phys. Rev. D **90** (2014) 072010, arXiv: 1408.5774 [hep-ex].

[25] ATLAS Collaboration, *Search for lepton-flavor-violating decays of the Z boson into a τ lepton and a light lepton with the ATLAS detector*, Phys. Rev. D **98** (2018) 092010, arXiv: 1804.09568 [hep-ex].

[26] ATLAS Collaboration, *Search for lepton-flavour-violating decays of the Higgs and Z bosons with the ATLAS detector*, Eur. Phys. J. C **77** (2017) 70, arXiv: 1604.07730 [hep-ex].

[27] ATLAS Collaboration, *Search for charged-lepton-flavour violating $\mu\tau qt$ interactions in top-quark production and decay with the ATLAS detector at the LHC*, ATLAS-CONF-2023-001, 2023, URL: https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/CONFNOTES/ATLAS-CONF-2023-001/.

[28] ATLAS Collaboration, *Search for lepton-flavor violation in different-flavor, high-mass final states in pp collisions at* $\sqrt{s} = 13$ *TeV with the ATLAS detector*, Phys. Rev. D **98** (2018) 092008, arXiv: 1807.06573 [hep-ex].

[29] ATLAS Collaboration, *Search for lepton flavour violation in the eμ continuum with the ATLAS detector in* $\sqrt{s} = 7$ *TeV pp collisions at the LHC*, Eur. Phys. J. C **72** (2012) 2040, arXiv: 1205.0725 [hep-ex].

[30] Ana M. Teixeira, *Theoretical aspects of charged Lepton Flavour Violation*, J. Phys.: Conf. Ser. **888**.012029 (2017), URL: https://iopscience.iop.org/article/10.1088/1742-6596/888/1/012029.

[31] J. Hisano et al., *Lepton flavor violation via right-handed neutrino Yukawa couplings in supersymmetric standard model*, Phys. Rev. **D53**.2442 (1996), arXiv: hep-ph/9510309 [hep-ph], URL: https://doi.org/10.1103/PhysRevD.53.2442.

[32] A. Pilaftsis A. Ilakovac and L. Popov, *Charged lepton flavor violation in supersymmetric low-scale seesaw models*, Phys. Rev. **D87**.053014 (2013), arXiv: 1212.5939 [hep-ph], URL: https://doi.org/10.1103/PhysRevD.87.053014.

[33] G. C. Branco et al., *Theory and phenomenology of two-Higgs-doublet models*, Phys. Rept. **516** (2012), arXiv: 1106.0034 [hep-ph], URL: https://doi.org/10.1016/j.physrep.2012.02.002.

[34] A. Kokulu A. Crivellin and C. Greub, *Flavor-phenomenology of two-Higgs-doublet models with generic Yukawa structure*, Phys. Rev. **D87**.094031 (2013), arXiv: 1303.5877 [hep-hp], URL: https://doi.org/10.1103/PhysRevD.87.094031.

[35] A. Vicente, *Lepton flavor violation beyond the MSSM*, Adv. High Energy Phys. 686572 (2015), arXiv: 1503.08622 [hep-hp], URL: http://dx.doi.org/10.1155/2015/686572.

[36] J. C. Pati and A. Salam, *Unified Lepton-Hadron Symmetry and a Gauge Theory of the Basic Interactions*, Phys. Rev. **D8**.1240 (1973), URL: https://doi.org/10.1103/PhysRevD.8.1240.

[37] H. Georgi and S. L. Glashow, *Unity of All Elementary Particle Forces*, Phys. Rev. Lett. **32**.438 (1974), URL: https://doi.org/10.1103/PhysRevLett.32.438.

[38] H. Fritzsch and P. Minkowski, *Unified Interactions of Leptons and Hadrons*, Annals Phys. **93**.193 (1975), URL: https://doi.org/10.1016/0003-4916(75)90211-0.

[39] A. De Rujula, H. Georgi, and S. L. Glashow, *Flavor goniometry by proton decay*, Phys. Rev. Lett. **45**.413 (1980), URL: https://doi.org/10.1103/PhysRevLett.45.413.

[40] L. J. Hall and M. Suzuki, *Explicit R-Parity Breaking in Supersymmetric Models*, Nucl. Phys. **B231**.419 (1984), URL: https://doi.org/10.1016/0550-3213(84)90513-3.

[41] S. Dimopoulos, *Technicolored Signatures*, Nucl. Phys. **B168**.69 (1980), URL: https://doi.org/10.1016/0550-3213(80)90277-1.

[42] J. Wudka, *Composite leptoquarks*, Phys. Lett. **B167**.337 (1986), URL: https://doi.org/10.1016/0370-2693(86)90356-4.

[43] J. K. Elwood and A. E. Faraggi, *Exotic leptoquarks from superstring derived models*, Nucl. Phys. **B512**.42 (1998), arXiv: hep-ph/9704363 [hep-hp].

[44] I. Doršner et al., *Physics of leptoquarks in precision experiments and at particle colliders*, Phys. Rept. **641** (2016), arXiv: 1603.04993 [hep-hp], URL: https://doi.org/10.1016/j.physrep.2016.06.001.

[45] Ilja Doršner, Svjetlana Fajfer, and Nejc Košnik, *Heavy and light scalar leptoquarks in proton decay*, Phys. Rev. **D86**.015013 (2012), arXiv: 1204.0674 [hep-ph], URL: https://doi.org/10.1103/PhysRevD.86.015013.

[46] Heavy Flavor Averaging Group (HFLAV), *Average of $R(D)$ and $R(D^*)$ for Summer 2018*, URL: https://hflav-eos.web.cern.ch/hflav-eos/semi/summer18/RDRDs.html.

[47] LHCb Collaboration, *Test of lepton universality using $B^+ \to K^+\ell^+\ell^-$ decays*, Phys. Rev. Lett. **113**.151601 (2014), arXiv: 1406.6482 [hep-hp].

[48] LHCb Collaboration, *Test of lepton universality with $B^0 \to K^{*0}\ell^+\ell^-$ decays*, JHEP **08**.055 (2017), arXiv: 1705.05802 [hep-hp].

[49] LHCb Collaboration, *Angular analysis of the $B^0 \to K^{*0}\mu^+\mu^-$ decay using 3 fb^{-1} of integrated luminosity*, JHEP **02**.104 (2016), arXiv: 1512.04442 [hep-hp].

[50] ATLAS Collaboration, *Angular analysis of $B_d^0 \to K^*\mu^+\mu^-$ decays in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, ATLAS-CONF-2017-023, 2017, URL: https://cds.cern.ch/record/2258146.

[51] LHCb collaboration, *Test of lepton universality in beauty-quark decays*, Nature Physics **18**.277-282 (2022), arXiv: 2103.11769 [hep-ex], URL: https://doi.org/10.48550/arXiv.2103.11769.

[52] The LHCb collaboration, *Test of lepton universality with $\Lambda_b^0 \to pK^-\ell^+\ell^-$ decays*, JHEP **05**.040 (2020), URL: https://doi.org/10.1007/JHEP05(2020)040.

[53] LHCb collaboration, *Measurement of lepton universality parameters in $B^+ \to K^+\ell^+\ell^-$ and $B^0 \to K^{*0}\ell^+\ell^-$ decays*, Phys. Rev. D **108**.3 (2023) 032002, arXiv: 2212.09153 [hep-ex].

[54] LHCb collaboration, *Test of lepton universality in $b \to s\ell^+\ell^-$ decays*, Phys. Rev. Lett. **131**.5 (2023) 051803, arXiv: 2212.09152 [hep-ex].

[55] T.J. Kim, P. Ko, J. Li, J. Park, and P. Wu, *Correlation between $R_{D^{(*)}}$ and top quark FCNC decays in leptoquark models*, JHEP **07**.25 (2019), URL: https://doi.org/10.1007/JHEP07(2019)025.

[56] B. Dumont, K. Nishiwaki, and R. Watanabe,
LHC constraints and prospects for S_1 scalar leptoquark explaining the $\bar{B} \to D^{(*)}\tau\bar{\nu}$ anomaly,
Phys. Rev. D **94**.034001 (2016), arXiv: 1603.05248 [hep-ph],
URL: https://doi.org/10.48550/arXiv.1603.05248.

[57] B. Abi et al. (Muon $g-2$ Collaboration),
Measurement of the Positive Muon Anomalous Magnetic Moment to 0.46 ppm,
Phys.Rev.Lett. **126**.141801 (2021),
URL: https://doi.org/10.1103/PhysRevLett.126.141801.

[58] David W. Hertzog and William M. Morse,
The Brookhaven Muon Anomalous Magnetic Moment Experiment,
Annual Review of Nuclear and Particle Science **54**.1 (2004) 141.

[59] A. A. Petrov and A. E. Blechman, *Effective Field Theories*, World Scientific, 2016,
ISBN: 978-9814434928.

[60] B. Henning, X. Lu, and H. Murayama, *How to use the Standard Model effective field theory*,
JHEP **01**.023 (2016), arXiv: 1412.1837 [hep-hp].

[61] B. Grzadkowski, M. Iskrzynski, M. Misiak, and J. Rosiek,
Dimension-Six Terms in the Standard Model Lagrangian, JHEP **10** (2010) 85,
arXiv: 1008.4884 [hep-ph].

[62] Michael Carpentier and Sacha Davidson, *Constraints on two-lepton two-quark operators*,
Eur. Phys. J. **C70**.1071 (2010), arXiv: 1008.0280 [hep-ph],
URL: https://doi.org/10.1140/epjc/s10052-010-1482-4.

[63] Sacha Davidson, Michelangelo L. Mangano, Stephane Perries, and Viola Sordini,
Lepton Flavour Violating top decays at the LHC, Eur. Phys. J. **C75**.9 (2015) 450,
arXiv: 1507.07163 [hep-ph].

[64] Mikael Chala, Jose Santiago, and Michael Spannowsky,
Constraining four-fermion operators using rare top decays, JHEP **04** (2019) 014,
arXiv: 1809.09624 [hep-ph].

[65] ATLAS Collaboration, *Search for flavour-changing neutral current top-quark decays $t \to qZ$ in proton–proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, JHEP **07** (2018) 176,
arXiv: 1803.09923 [hep-ex].

[66] CMS Collaboration, *Search for charged-lepton flavor violation in top quark production and decay in pp collisions at $\sqrt{s}$ = 13 TeV*, JHEP **06**.082 (2022), arXiv: 2201.07859 [hep-ex].

[67] CMS Collaboration, *Search for charged lepton flavor violation in the top quark sector in trilepton final states with the CMS detector at $\sqrt{s}$=13 TeV*, CMS-PAS-TOP-22-005, 2023,
URL: https://cms-results.web.cern.ch/cms-results/public-results/preliminary-results/TOP-22-005/index.html.

[68] L. Evans and P. Bryant, *LHC machine*, JINST **3** (2008) S08001,
URL: http://cds.cern.ch/record/1129806.

[69] The ALICE Collaboration, *The ALICE experiment at the CERN LHC*, JINST **3** (2008) S08002, URL: https://www.researchgate.net/publication/230981232_The_alice_experiment_at_the_CERN_LHC.

[70] The CMS Collaboration, *The CMS experiment at the CERN LHC*, JINST **3** (2008) S08004, URL: https://inspirehep.net/literature/796887.

[71] The LHCb Collaboration, *The LHCb Detector at the LHC*, JINST **3** (2008) S08005, URL: https://inspirehep.net/literature/796248.

[72] The ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, JINST **3** (2008) S08003, URL: https://cds.cern.ch/record/1129811.

[73] CERN website, *The Large Hadron Collider*, accessed 4 July 2023, URL: https://home.web.cern.ch/science/accelerators/large-hadron-collider.

[74] *Public ATLAS Luminosity Results for Run-2 of the LHC*, URL: https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2.

[75] *Public ATLAS Luminosity Results for Run-3 of the LHC*, URL: https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun3.

[76] The ATLAS Collaboration, *ATLAS Detector and Physics Performance – Technical Design Report, Vol.1*, CERN/LHCC/99-14, ATLAS-TDR-14, 1999, URL: https://inspirehep.net/literature/511648.

[77] The ATLAS Collaboration – ATLAS IBL Community, *ATLAS Insertable B-Layer Technical Design Report*, CERN/LHCC/2010-013, ATLAS-TDR-019, 2010, URL: https://inspirehep.net/literature/1609508.

[78] Yosuke Takubo on behalf of the ATLAS collaboration, *ATLAS IBL operational experience*, ATL-INDET-PROC-2016-012, PoS **Vertex2016** (2016) 004.

[79] ATLAS Experiment website, *Detector and technology*, accessed 4 July 2023, URL: https://atlas.cern/Discover/Detector.

[80] The ATLAS Collaboration – ATLAS Level-1 Trigger Group, *ATLAS Level-1: Trigger Technical Design Report*, CERN/LHCC/98-14, ATLAS-TDR-12, 1998, URL: https://cds.cern.ch/record/381429.

[81] The ATLAS Collaboration, *Performance of the ATLAS Trigger System in 2010*, CERN-PH-EP-2011-078, Eur. Phys. J. **C72**.1849 (2012), arXiv: 1110.1530 [hep-ex].

[82] The ATLAS Collaboration, *Performance of the ATLAS Trigger System in 2015*, Eur. Phys. J. C **77**.317 (2017), arXiv: 1611.09661 [hep-ex], URL: https://doi.org/10.48550/arXiv.1611.09661.

[83] Arantxa Ruiz Martínez on behalf of the ATLAS Collaboration,
The Run-2 ATLAS Trigger System, ATL-DAQ-PROC-2016-003,
J. Phys.: Conf. Ser. **762**.012003 (2016), URL: https://cds.cern.ch/record/2133909.

[84] ATLAS Collaboration, *Operation of the ATLAS trigger system in Run 2*, CERN-EP-2020-109,
JINST **15** (2020) P10004, arXiv: 2007.12539v2 [physics.ins-det],
URL: https://inspirehep.net/literature/1808700.

[85] The ATLAS Collaboration,
ATLAS High-Level Trigger, Data Acquisition and Controls Technical Design Report,
CERN/LHCC/2003-022, ATLAS-TDR-016, 2003,
URL: https://cds.cern.ch/record/616089.

[86] The ATLAS Collaboration, *The ATLAS Fast Tracker system*, CERN-EP-2020-232,
JINST **16**.P07006 (2021), arXiv: 2101.05078v2 [physics.ins-det].

[87] The ATLAS Collaboration,
Technical Design Report for the Phase-I Upgrade of the ATLAS TDAQ System,
CERN-LHCC-2013-018, ATLAS-TDR-023, 2013,
URL: https://cds.cern.ch/record/1602235.

[88] ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider: A Description of the Detector Configuration for Run 3*, CERN-EP-2022-259, (2023),
arXiv: 2305.16623v1 [physics.ins-det],
URL: https://inspirehep.net/literature/2663519.

[89] Y. Ermoline on behalf of The ATLAS Collaboration,
The Phase-I upgrade of the ATLAS Level-1 calorimeter trigger, JINST **15**.C06040 (2020),
URL: https://inspirehep.net/files/192893ddb1e96fbcb39823d1f89e85a4.

[90] V. Andrei et al., *Tile Rear Extension module for the Phase-I upgrade of the ATLAS L1Calo PreProcessor system*, JINST **12**.C03034 (2017),
URL: https://inspirehep.net/literature/1517067.

[91] *Achieving Backplane Redundancy in AdvancedTCA Systems, An Approach with Commonly Understood Standard VLAN Protocols on the Radisys Promentum® Platform*,
Radisys® White Paper (2009),
URL: http://go.radisys.com/rs/radisys/images/paper-atca-achieving.pdf.

[92] Xilinx, *Xilinx 7 Series FPGAs Data Sheet: Overview: Product Specification*, 2018, URL: https://www.xilinx.com/support/documentation/data_sheets/ds180_7Series_Overview.pdf.

[93] Richard Staley, *FTM Programming Model*, 2019, URL: http://epweb2.ph.bham.ac.uk/user/staley/FTM/Documentation/FTM_Programming_Model_v1_2_0.pdf.

[94] Richard Staley,
Technical Specification: A Multi-Gigabit/s Data Source/Sink for commissioning the eFEX module of the ATLAS L1 Calo. Upgrade Programme, eFEX / jFEX Test Module (FTM), 2019,

URL: http://epweb2.ph.bham.ac.uk/user/staley/FTM/Documentation/FTM_Spec_Production_1_0.pdf.

[95] Thomas L. Floyd, *Digital Fundamentals*, 11th, Pearson Education Limited, 2015, ISBN: 978-1-292-07598-3.

[96] John Wakerly, *Digital Design: Principles and Practices*, 5th, Pearson, 2018, ISBN: 978-0134460093.

[97] *I2C Bus website*, URL: https://www.i2c-bus.org/.

[98] *I2C info – I2C Bus, Interface and Protocol*, URL: https://i2c.info/.

[99] *Understanding the I^2C Bus*, 2015, URL: http://www.ti.com/lit/an/slva704/slva704.pdf.

[100] C. Ghabrous Larrea et al., *IPbus: a flexible Ethernet-based control system for xTCA hardware*, JINST **10**.C02019 (2015), URL: https://doi.org/10.1088/1748-0221/10/02/C02019.

[101] *Xilinx: System Monitor and XADC*, URL: https://www.xilinx.com/products/technology/analog-mixed-signal.htm.

[102] Cathal Murphy, *XADC Layout Guidelines, Application Note: Kintex-7 FPGAs*, Xilinx, 2013, URL: https://www.xilinx.com/support/documentation/application_notes/xapp554-xadc-layout-guidelines.pdf.

[103] Cathal Murphy, *Driving the Xilinx Analog-to-Digital Converter, Application Note: 7 Series FPGAs*, Xilinx, 2015, URL: https://www.xilinx.com/support/documentation/application_notes/xapp795-driving-xadc.pdf.

[104] *LM82 Remote Diode and Local Digital Temperature Sensor with Two-Wire Interface*, 2013, URL: http://www.ti.com/lit/ds/symlink/lm82.pdf.

[105] Weiming Quian, *Design and test performance of the ATLAS Feature Extractor trigger boards for the Phase-1 Upgrade*, JINST **12**.C01079 (2017), URL: https://inspirehep.net/literature/1513689.

[106] Weiming Quian, *Design and test of the electron Feature EXtractor (eFEX) preproduction module for the ATLAS Phase-I Upgrade*, ATL-DAQ-SLIDE-2019-642, (2019), URL: https://cds.cern.ch/record/2689075/.

[107] R. Oishi, *LHC-ATLAS Phase-1 upgrade: firmware validation for real time digital processing for new trigger readout system of the Liquid Argon calorimeter*, JINST **15**.C05013 (2020), URL: https://inis.iaea.org/search/search.aspx?orig_q=RN:52079003.

[108] Luka Selem on behalf of ATLAS Liquid Argon Calorimeter Group, *ATLAS LAr Calorimeter commissioning for LHC Run-3: Energy computation in LATOME boards*, ATL-LARG-PROC-2021-003, 2021, URL: https://cds.cern.ch/record/2777241/.

[109] Xilinx, *Vivado Design Suite User Guide, Programming and Debugging*, 2022, URL: https://www.xilinx.com/support/documents/sw_manuals/xilinx2022_1/ug908-vivado-programming-debugging.pdf.

[110] Mentor Graphics Corporation, ModelSim *Reference Manual*, (2012), URL: https://www.microsemi.com/document-portal/doc_view/131617-modelsim-reference-manual.

[111] Siemens, ModelSim *website*, accessed 1 January 2024, URL: https://eda.sw.siemens.com/en-US/ic/modelsim/.

[112] The ATLAS Collaboration, *Technical Design Report for the Phase-II Upgrade of the ATLAS TDAQ System*, CERN-LHCC-2017-020, 2017, rev.2022, URL: https://cds.cern.ch/record/2285584.

[113] Torbjörn Sjöstrand, *Monte Carlo Generators*, (2006), URL: https://cds.cern.ch/record/999717.

[114] G.Altarelli and G.Parisi, *Asymptotic freedom in parton language*, Nuclear Physics B **126** (1977), URL: https://doi.org/10.1016/0550-3213(77)90384-4.

[115] O. Behnke, K. Kröninger, G. Schott, and T. Schörner-Sadenius, eds., *Data Analysis in High Energy Physics, A Practical Guide to Statistical Methods*, WILEY-VCH Verlag GmbH & Co. KGaA, 2013, ISBN: 978-3-527-41058-3.

[116] Daniel de Florian and Massimiliano Grazzini, *Next-to-next-to-leading logarithmic corrections at small transverse momentum in hadronic collisions*, Phys.Rev.Lett. **85** (2000), arXiv: 0008152 [hep-ph], URL: https://doi.org/10.48550/arXiv.hep-ph/0008152.

[117] Zoltan Nagy and Davison E. Soper, *What is a parton shower?*, Phys. Rev. D **98**.014034 (2018), arXiv: 1705.08093v3 [hep-ph], URL: https://doi.org/10.48550/arXiv.1705.08093.

[118] S. Agostinelli et al., *GEANT4: A Simulation toolkit*, Nucl. Instrum. Meth. **A506** (2003) 250.

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

[120] *Overview of the MC samples released in the 13 TeV ATLAS Open Data*, URL: http://opendata.atlas.cern/release/2020/documentation/datasets/mc.html.

[121] Enrico Bothmann et al., *Event generation with Sherpa 2.2*, SciPost Phys. **7**.3 (2019) 034, arXiv: 1905.09127 [hep-ph].

[122] C Ay et al., *Monte Carlo generators in ATLAS software*, Journal of Physics: Conference Series **219**.3 (2010) 032001, URL: https://dx.doi.org/10.1088/1742-6596/219/3/032001.

[123] Torbjorn Sjostrand, Stephen Mrenna, and Peter Z. Skands, *PYTHIA 6.4 Physics and Manual*, JHEP **05** (2006) 026, arXiv: hep-ph/0603175 [hep-ph].

[124] Torbjörn Sjöstrand et al., *An introduction to PYTHIA 8.2*, Comput. Phys. Commun. **191** (2015) 159, arXiv: 1410.3012 [hep-ph].

[125] Johannes Bellm et al., *Herwig 7.0/Herwig++ 3.0 release note*, Eur. Phys. J. **C76**.4 (2016) 196, arXiv: 1512.01178 [hep-ph].

[126] M. Bähr et al., *Herwig++ physics and manual*, Eur. Phys. J. C **58** (2008) 639, arXiv: 0803.0883 [hep-ph].

[127] Johannes Bellm et al., *Herwig 7.0/Herwig++ 3.0 release note*, Eur. Phys. J. C **76**.4 (2016) 196, arXiv: 1512.01178 [hep-ph].

[128] T. Gleisberg et al., *Event generation with SHERPA 1.1*, JHEP **02** (2009) 007, arXiv: 0811.4622 [hep-ph].

[129] Emanuele Re, *Single-top Wt-channel production matched with parton showers using the POWHEG method*, Eur. Phys. J. C **71** (2011) 1547, arXiv: 1009.2450 [hep-ph].

[130] Rikkert Frederix, Emanuele Re, and Paolo Torrielli, *Single-top t-channel hadroproduction in the four-flavour scheme with POWHEG and aMC@NLO*, JHEP **09** (2012) 130, arXiv: 1207.5391 [hep-ph].

[131] Paolo Nason, *A new method for combining NLO QCD with shower Monte Carlo algorithms*, JHEP **11** (2004) 040, arXiv: hep-ph/0409146.

[132] Stefano Frixione, Paolo Nason, and Carlo Oleari, *Matching NLO QCD computations with parton shower simulations: the POWHEG method*, JHEP **11** (2007) 070, arXiv: 0709.2092 [hep-ph].

[133] Simone Alioli, Paolo Nason, Carlo Oleari, and Emanuele 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].

[134] W. F. Long T. Stelzer, *Automatic Generation of Tree Level Helicity Amplitudes*, Comput.Phys.Commun. **81** (1994), arXiv: 9401258v1 [hep-ph], URL: https://doi.org/10.48550/arXiv.hep-ph/9401258.

[135] T. Stelzer F. Maltoni, *MadEvent: Automatic Event Generation with MadGraph*, JHEP02 **027** (2003), arXiv: 0208156v1 [hep-ph], URL: https://doi.org/10.48550/arXiv.hep-ph/0208156.

[136] J. Alwall at al., *MadGraph/MadEvent v4: The New Web Generation*, JHEP09 **028** (2007), arXiv: 0706.2334v1 [hep-ph], URL: https://doi.org/10.48550/arXiv.0706.2334.

[137] J. Alwall at al., *New Developments in MadGraph/MadEvent*, AIP Conf.Proc **1078** (2009), arXiv: 0809.2410v1 [hep-ph], URL: https://doi.org/10.48550/arXiv.0809.2410.

[138] Johan Alwall, Michel Herquet, Fabio Maltoni, Olivier Mattelaer, and Tim Stelzer, *MadGraph 5 : Going Beyond*, JHEP **06** (2011) 128, arXiv: 1106.0522 [hep-ph].

[139] Stefano Frixione and Bryan R. Webber, *Matching NLO QCD computations and parton shower simulations*, JHEP **06** (2002) 029, arXiv: hep-ph/0204244.

[140] Stefano Frixione, Paolo Nason, and Bryan R. Webber, *Matching NLO QCD and parton showers in heavy flavour production*, JHEP **08** (2003) 007, arXiv: hep-ph/0305252.

[141] 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].

[142] M. Jezabek, Z. Was, S. Jadach, and J.H. Kuhn, *The tau decay library TAUOLA, update with exact O(alpha) QED corrections in tau → mu (e) neutrino anti-neutrino decay modes*, Comput.Phys.Commun. **70** (1992), URL: https://doi.org/10.1016/0010-4655(92)90092-D.

[143] P. Golonka and Z. Was,
PHOTOS Monte Carlo: a precision tool for QED corrections in Z and W decays,
The European Physical Journal C (2006),
URL: https://doi.org/10.1140/epjc/s2005-02396-4.

[144] D. J. Lange, *The EvtGen particle decay simulation package*,
Nucl. Instrum. Meth. A **462** (2001) 152.

[145] Céline Degrande et al., *UFO - The Universal FeynRules Output*,
Comput.Phys.Commun. **183** (2012), arXiv: 1108.2040v2 [hep-ph],
URL: https://doi.org/10.48550/arXiv.1108.2040.

[146] Adam Alloul, Neil D. Christensen, Céline Degrande, Claude Duhr, and Benjamin Fuks,
FeynRules 2.0 - A complete toolbox for tree-level phenomenology,
Comput. Phys. Commun. **185** (2014) 2250, arXiv: 1310.1921 [hep-ph].

[147] D. Barducci et al.,
Interpreting top-quark LHC measurements in the standard-model effective field theory,
(2018), ed. by Juan Antonio Aguilar-Saavedra et al., arXiv: 1802.07237 [hep-ph].

[148] Richard D. Ball et al., *Parton distributions for the LHC run II*, JHEP **04** (2015) 040,
arXiv: 1410.8849 [hep-ph].

[149] *ATLAS Run 1 Pythia8 tunes*, tech. rep. ATL-PHYS-PUB-2014-021, CERN, 2014,
URL: https://cds.cern.ch/record/1966419.

[150] David Griffiths, *Introduction to Quantum Mechanics*, 2nd, Pearson – Prentice Hall, 2004,
ISBN: 978-0131118928.

[151] Giulia Negro, *Top Quark Modelling and Tuning in ATLAS and CMS*,
14th International Workshop on Top Quark Physics (2022), arXiv: 2201.03517v2 [hep-ex],
URL: https://doi.org/10.48550/arXiv.2201.03517.

[152] ATLAS Collaboration, *Studies on top-quark Monte Carlo modelling for Top2016*,
ATL-PHYS-PUB-2016-020, 2016, URL: https://cds.cern.ch/record/2216168.

[153] ATLAS Collaboration, *ATLAS Pythia 8 tunes to 7 TeV data*, ATL-PHYS-PUB-2014-021,
2014, URL: https://cds.cern.ch/record/1966419.

[154] M. Beneke, P. Falgari, S. Klein, and C. Schwinn,
Hadronic top-quark pair production with NNLL threshold resummation,
Nucl. Phys. B **855** (2012) 695, arXiv: 1109.1536 [hep-ph].

[155] Matteo Cacciari, Michal Czakon, Michelangelo Mangano, Alexander Mitov, and Paolo Nason, *Top-pair production at hadron colliders with next-to-next-to-leading logarithmic soft-gluon resummation*, Phys. Lett. B **710** (2012) 612, arXiv: 1111.5869 [hep-ph].

[156] Peter Bärnreuther, Michal Czakon, and Alexander Mitov, *Percent-Level-Precision Physics at the Tevatron: Next-to-Next-to-Leading Order QCD Corrections to $q\bar{q} \to t\bar{t} + X$*, Phys. Rev. Lett. **109** (2012) 132001, arXiv: 1204.5201 [hep-ph].

[157] Michal Czakon and Alexander Mitov, *NNLO corrections to top-pair production at hadron colliders: the all-fermionic scattering channels*, JHEP **12** (2012) 054, arXiv: 1207.0236 [hep-ph].

[158] Michal Czakon and Alexander Mitov, *NNLO corrections to top pair production at hadron colliders: the quark-gluon reaction*, JHEP **01** (2013) 080, arXiv: 1210.6832 [hep-ph].

[159] Michal Czakon, Paul Fiedler, and Alexander Mitov, *Total Top-Quark Pair-Production Cross Section at Hadron Colliders Through $O(\alpha_S^4)$*, Phys. Rev. Lett. **110** (2013) 252004, arXiv: 1303.6254 [hep-ph].

[160] Michal Czakon and Alexander 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].

[161] Stefano Frixione, Eric Laenen, Patrick Motylinski, Chris White, and Bryan R. Webber, *Single-top hadroproduction in association with a W boson*, JHEP **07** (2008) 029, arXiv: 0805.3067 [hep-ph].

[162] Nikolaos 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].

[163] Nikolaos Kidonakis, *Top Quark Production*, KSU-HEP-110113, Proceedings, Helmholtz International Summer School on Physics of Heavy Quarks and Hadrons (HQ 2013) (2013) 139, arXiv: 1311.0283 [hep-ph].

[164] Andrea Giammanco, *Single top quark production at the LHC*, Reviews in Physics **1** (2016) 1, ISSN: 2405-4283, URL: https://www.sciencedirect.com/science/article/pii/S2405428315000027.

[165] Chris D. White, Stefano Frixione, Eric Laenen, and Fabio Maltoni, *Isolating Wt production at the LHC*, JHEP **11**.074 (2009).

[166] Stefano Frixione, Eric Laenen, Patrick Motylinski, and Bryan 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.

[167] Pierre Artoisenet, Rikkert Frederix, Olivier Mattelaer, and Robbert Rietkerk, *Automatic spin-entangled decays of heavy resonances in Monte Carlo simulations*, JHEP **03** (2013) 015, arXiv: 1212.3460 [hep-ph].

[168] 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].

[169] Tanju Gleisberg and Stefan Höche, *Comix, a new matrix element generator*,
JHEP **12** (2008) 039, arXiv: 0808.3674 [hep-ph].

[170] Steffen Schumann and Frank Krauss,
A parton shower algorithm based on Catani–Seymour dipole factorisation,
JHEP **03** (2008) 038, arXiv: 0709.1027 [hep-ph].

[171] Stefan Höche, Frank Krauss, Marek Schönherr, and Frank Siegert,
A critical appraisal of NLO+PS matching methods, JHEP **09** (2012) 049,
arXiv: 1111.1220 [hep-ph].

[172] Stefan Höche, Frank Krauss, Marek Schönherr, and Frank Siegert,
QCD matrix elements + parton showers. The NLO case, JHEP **04** (2013) 027,
arXiv: 1207.5030 [hep-ph].

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

[174] Stefan Höche, Frank Krauss, Steffen Schumann, and Frank Siegert,
QCD matrix elements and truncated showers, JHEP **05** (2009) 053,
arXiv: 0903.1219 [hep-ph].

[175] Federico Buccioni et al., *OpenLoops 2*, Eur. Phys. J. C **79**.10 (2019) 866,
arXiv: 1907.13071 [hep-ph].

[176] Fabio Cascioli, Philipp Maierhöfer, and Stefano Pozzorini,
Scattering Amplitudes with Open Loops, Phys. Rev. Lett. **108** (2012) 111601,
arXiv: 1111.5206 [hep-ph].

[177] Ansgar Denner, Stefan Dittmaier, and Lars Hofer,
COLLIER*: A fortran-based complex one-loop library in extended regularizations*,
Comput. Phys. Commun. **212** (2017) 220, arXiv: 1604.06792 [hep-ph].

[178] Jan-Christopher Winter, Frank Krauss, and Gerhard Soff,
A modified cluster-hadronization model, Eur. Phys. J. C **36** (2004) 381,
arXiv: hep-ph/0311085.

[179] Charalampos Anastasiou, Lance Dixon, Kirill Melnikov, and Frank Petriello,
High-precision QCD at hadron colliders: Electroweak gauge boson rapidity distributions at next-to-next-to leading order, Phys. Rev. D **69** (2004) 094008, arXiv: hep-ph/0312266.

[180] T. Cornelissen et al. on behalf of the ATLAS Inner Detector software group,
The new ATLAS Track Reconstruction (NEWT),
Journal of Physics: Conference Series **119**.032014 (2008),
URL: https://s3.cern.ch/inspire-prod-files-0/050f4cd4d3945f3847859bd6c1c8ebc9.

[181] V. Kartvelishvili, *Electron bremsstrahlung recovery in ATLAS*, Nuclear Physics B - Proceedings Supplements **172** (2007) 208, ISSN: 0920-5632, URL: https://www.sciencedirect.com/science/article/pii/S0920563207005944.

[182] ATLAS Collaboration, *ATLAS Track Reconstruction Tutorial*, URL: https://atlassoftwaredocs.web.cern.ch/trackingTutorial/idoverview/.

[183] ATLAS Collaboration, *Characterization of Interaction-Point Beam Parameters Using the pp Event-Vertex Distribution Reconstructed in the ATLAS Detector at the LHC*, ATLAS-CONF-2010-027, 2010, URL: https://cds.cern.ch/record/1277659.

[184] ATLAS Collaboration, *Vertex Reconstruction Performance of the ATLAS Detector at* $\sqrt{s} = 13$ *TeV*, ATL-PHYS-PUB-2015-026, 2015, URL: https://cds.cern.ch/record/2037717.

[185] URL: https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PLOTS/IDTR-2017-006.

[186] URL: https://twiki.cern.ch/twiki/bin/view/Main/VertexingGroup.

[187] Chiara Magliocca, *The Transverse Impact Parameter as a Relevant Variable to Study the Adjusted-MC Method in the ATLAS Experiment*, DESY Summer Student Programme, 2018, URL: https://www.zeuthen.desy.de/students/2018/Summerstudents2018/Reports/Magliocca.pdf.

[188] ATLAS Collaboration, *Electron efficiency measurements with the ATLAS detector using the 2012 LHC proton–proton collision data*, CERN-EP-2016-262, Eur.Phys.J.C. **77**.195 (2017), arXiv: 1612.01456v2 [hep-ex].

[189] ATLAS Collaboration, *Electron and photon performance measurements with the ATLAS detector using the 2015–2017 LHC proton–proton collision data*, JINST **14** (2019) P12006, arXiv: 1908.00005 [hep-ex].

[190] F. He for the ATLAS Collaboration, *High Precision Electron and Muon Reconstruction Performance with ATLAS at LHC Run-2*, ATL-PHYS-PROC-2022-111, PoS ICHEP (2022) 675, URL: https://cds.cern.ch/record/2841383/.

[191] ATLAS Collaboration, *Electron reconstruction and identification in the ATLAS experiment using the 2015 and 2016 LHC proton–proton collision data at* $\sqrt{s} = 13$ *TeV*, Eur. Phys. J. C **79** (2019) 639, arXiv: 1902.04655 [hep-ex].

[192] ATLAS Collaboration, *Tools for estimating fake/non-prompt lepton backgrounds with the ATLAS detector at the LHC*, CERN-EP-2022-214, (2022), arXiv: 2211.16178v1 [hep-ex], URL: https://doi.org/10.48550/arXiv.2211.16178.

[193] ATLAS Collaboration, *Electron and photon efficiencies in LHC Run 2 with the ATLAS experiment*, CERN-EP-2023-182, JHEP (2023), arXiv: 2308.13362v1 [hep-ex], URL: https://cds.cern.ch/record/2868520/.

[194] ATLAS Collaboration, *Electron efficiency with full Run2*, 2022, URL: https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PLOTS/EGAM-2022-02/.

[195] Fudong He, Rustem Ospanov, and Rhys Roberts,
Prompt LeptonIso BDT in rel21, Muon Combined Performance meeting, ATLAS, 2017,
URL: https://indico.cern.ch/event/673290/.

[196] R. Ospanov, F. He, R. Roberts, and T. Shroeder,
Status and plans for prompt lepton tagging with release 21, ATLAS physics workshop,
ATLAS, 2017, URL: https://indico.cern.ch/event/675559/.

[197] S. Akatsuka and S. Chen, *Isolation WPs summary: PLV + LowPtPLV*, IFF: summary of selected lepton and photon ISO WPs for the precision recommendations, ATLAS, 2019,
URL: https://indico.cern.ch/event/854783/.

[198] F. Cardillo and O. Ducu, *PromptLeptonTaggerIFF*, 2021,
URL: https://twiki.cern.ch/twiki/bin/view/AtlasProtected/PromptLeptonTaggerIFF.

[199] C. Magliocca, *Measurement of the track impact parameters resolution with the ATLAS experiment at LHC using 2016–2018 data*, IL NUOVO CIMENTO **44 C**.55 (2021),
URL: https://cds.cern.ch/record/2791559/files/ncc12181.pdf.

[200] ATLAS Collaboration, *Electron efficiency measurements with the ATLAS detector using the 2015 LHC proton–proton collision data*, ATLAS-CONF-2016-024, 2016,
URL: https://cds.cern.ch/record/2157687.

[201] ATLAS Collaboration, *Muon reconstruction performance of the ATLAS detector in proton–proton collision data at* $\sqrt{s} = 13$ *TeV*, Eur. Phys. J. C **76** (2016) 292,
arXiv: 1603.05598 [hep-ex].

[202] ATLAS Collaboration, *Muon reconstruction and identification efficiency in ATLAS using the full Run 2 pp collision data set at* $\sqrt{s} = 13$ *TeV*, Eur. Phys. J. C **81**.578 (2021),
arXiv: 2012.00578v2 [hep-ex], URL: https://doi.org/10.48550/arXiv.2012.00578.

[203] ATLAS Collaboration,
Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1,
Eur. Phys. J. C **77** (2017) 490, arXiv: 1603.02934 [hep-ex].

[204] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez, *The anti-*k_t *jet clustering algorithm*,
JHEP **04** (2008) 063, arXiv: 0802.1189 [hep-ph].

[205] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez, *FastJet user manual*,
Eur. Phys. J. C **72** (2012) 1896, arXiv: 1111.6097 [hep-ph].

[206] 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].

[207] ATLAS Collaboration, *Performance of pile-up mitigation techniques for jets in pp collisions at* $\sqrt{s} = 8$ *TeV using the ATLAS detector*, Eur. Phys. J. C **76** (2016) 581,
arXiv: 1510.03823 [hep-ex].

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

[209] ATLAS Collaboration, *Optimisation and performance studies of the ATLAS b-tagging algorithms for the 2017-18 LHC run*, ATL-PHYS-PUB-2017-013, 2017, URL: https://cds.cern.ch/record/2273281.

[210] ATLAS Collaboration, *TopRecoObjTwikiModel*, URL: https://twiki.cern.ch/twiki/bin/view/AtlasProtected/TopRecoObjTwikiModel#b_Tagging.

[211] ATLAS Collaboration, *Tagger Calibration Recommendations for Release 21*, URL: https://twiki.cern.ch/twiki/bin/view/AtlasProtected/BTagCalibrationRecommendationsRelease21#Recommendation_March_2020.

[212] Martino Centonze on behalf of the ATLAS collaboration, *Jet flavour tagging for the ATLAS Experiment*, ATL-PHYS-PROC-2021-107, (2021), URL: https://cds.cern.ch/record/2792526/.

[213] Michela Paganini and on behalf of the ATLAS Collaboration, *Machine Learning Algorithms for b-Jet Tagging at the ATLAS Experiment*, Journal of Physics: Conference Series **1085**.042031 (2018), URL: https://iopscience.iop.org/article/10.1088/1742-6596/1085/4/042031/pdf.

[214] DØ collaboration, *Observation of single top quark production*, 2009, URL: https://www-d0.fnal.gov/Run2Physics/top/singletop_observation/singletop_observation_updated.html.

[215] D Adams et al., *Recommendations of the Physics Objects and Analysis Harmonisation Study Groups 2014*, tech. rep. ATL-PHYS-INT-2014-018, CERN, 2014, URL: https://cds.cern.ch/record/1743654.

[216] URL: https://indico.cern.ch/event/539619/contributions/2191033/attachments/1284239/1909241/Farrell_ort_asg.pdf.

[217] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez, *The Catchment Area of Jets*, JHEP **04** (2008) 005, arXiv: 0802.1188 [hep-ph].

[218] ATLAS Collaboration, *Performance of missing transverse momentum reconstruction with the ATLAS detector using proton–proton collisions at $\sqrt{s} = 13$ TeV*, Eur. Phys. J. C **78** (2018) 903, arXiv: 1802.08168 [hep-ex].

[219] ATLAS Collaboration, *E_T^{miss} performance in the ATLAS detector using 2015–2016 LHC pp collisions*, ATLAS-CONF-2018-023, 2018, URL: https://cds.cern.ch/record/2625233.

[220] ATLAS Collaboration, *Tagging and suppression of pileup jets with the ATLAS detector*, ATLAS-CONF-2014-018, URL: https://cds.cern.ch/record/1700870.

[221] ATLAS Collaboration,
Performance of electron and photon triggers in ATLAS during LHC Run 2,
Eur. Phys. J. C **80** (2020) 47, arXiv: 1909.00761 [hep-ex].

[222] S. D. Jones and on behalf of the ATLAS Collaboration,
The ATLAS Electron and Photon Triggers,
Journal of Physics: Conference Series **1085**.4 (2018) 042001,
URL: https://dx.doi.org/10.1088/1742-6596/1085/4/042001.

[223] ATLAS Collaboration, *Performance of the ATLAS muon triggers in Run 2*,
JINST **15** (2020) P09015, arXiv: 2004.13447 [hep-ex].

[224] ATLAS Collaboration, *Public Egamma Trigger Plots for Collision Data*, URL:
https://twiki.cern.ch/twiki/bin/view/AtlasPublic/EgammaTriggerPublicResults.

[225] URL: https://twiki.cern.ch/twiki/bin/view/Atlas/TrigEgammaRecommendations2015.

[226] URL: https://twiki.cern.ch/twiki/bin/view/Atlas/
MuonTriggerPhysicsRecommendationsRel212017.

[227] ATLAS Collaboration, *Measurement of $W^{\pm}Z$ production cross sections and gauge boson polarisation in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*,
Eur. Phys. J. C **79** (2019) 535, arXiv: 1902.05759 [hep-ex].

[228] URL: https://twiki.cern.ch/twiki/bin/view/AtlasProtected/
EGammaIdentificationRun2#Using_the_Ambiguity_tool_in_anal.

[229] Michael Wichmann,
Electron Charge Misidentification in the ATLAS Detector, University of Munich thesis, (2008),
URL: https://www-static.etp.physik.uni-
muenchen.de/dokumente/thesis/diplom_wichmann.pdf.

[230] URL: https:
//twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/ElectronChargeFlipTaggerTool.

[231] W. Andrews et al.,
Studies of the Electron charge misidentification rate, DPG/PH egamma meeting, (2009),
URL: https://slideplayer.com/slide/8816342/.

[232] ATLAS Collaboration, *Electron reconstruction and identification efficiency measurements with the ATLAS detector using the 2011 LHC proton-proton collision data*,
Eur. Phys. J. C **74**.2941 (2014), arXiv: 1404.2240v3 [hep-ex],
URL: https://doi.org/10.48550/arXiv.1404.2240.

[233] ATLAS Physics Modelling Group, *Weak Boson Processes*, accessed 20 October 2023,
URL: https://twiki.cern.ch/twiki/bin/view/AtlasProtected/PmgWeakBosonProcesses#
Treating_pathological_high_weigh.

[234] Julia Bauer, *Prospects for the Observation of Electroweak Top Quark Production with the CMS Experiment. Perspektiven zur Beobachtung der elektroschwachen Produktion einzelner Top-Quarks mit dem CMS Experiment*, PhD thesis: Karlsruhe, Inst. Technol., 2010 214, URL: http://cds.cern.ch/record/1308713.

[235] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning*, 2nd, Springer, 2017, ISBN: 978-0387848570, URL: https://hastie.su.domains/ElemStatLearn/.

[236] A. Hoecker et al., *TMVA - Toolkit for Multivariate Data Analysis*, PoS ACAT 040 (2007), arXiv: physics/0703039 [physics.data-an].

[237] Yoav Freund and Robert E. Schapire, *A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting*, Journal of Computer and System Sciences **55**.1 (1997) 119, ISSN: 0022-0000, URL: http://www.sciencedirect.com/science/article/pii/S002200009791504X.

[238] R. Brun and F. Rademakers, *ROOT: An object oriented data analysis framework*, Nucl. Instrum. Meth. **A389** (1997) 81.

[239] 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.

[240] ATLAS Collaboration, *Electron and photon energy calibration with the ATLAS detector using 2015–2016 LHC proton–proton collision data*, JINST **14** (2019) P03017, arXiv: 1812.03848 [hep-ex].

[241] ATLAS Collaboration, *Tagging and suppression of pileup jets with the ATLAS detector*, ATLAS-CONF-2014-018, 2014, URL: https://cds.cern.ch/record/1700870.

[242] 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].

[243] ATLAS Collaboration, *Jet energy resolution in proton–proton collisions at $\sqrt{s} = 7$ TeV recorded in 2010 with the ATLAS detector*, Eur. Phys. J. C **73** (2013) 2306, arXiv: 1210.6210 [hep-ex].

[244] ATLAS Collaboration, *Measurement of b-tagging efficiency of c-jets in $t\bar{t}$ events using a likelihood approach with the ATLAS detector*, ATLAS-CONF-2018-001, 2018, URL: https://cds.cern.ch/record/2306649.

[245] ATLAS Collaboration, *Calibration of light-flavour b-jet mistagging rates using ATLAS proton–proton collision data at $\sqrt{s} = 13$ TeV*, ATLAS-CONF-2018-006, 2018, URL: https://cds.cern.ch/record/2314418.

[246] Jon Butterworth et al., *PDF4LHC recommendations for LHC Run II*, J. Phys. G **43** (2016) 023001, arXiv: 1510.03865 [hep-ph].

[247] Rikkert Frederix and Ioannis Tsinikos, *On improving NLO merging for ttW production*, Journal of High Energy Physics **2021**.11 (2021), URL: https://doi.org/10.1007.

[248] Rikkert Frederix and Ioannis Tsinikos,
Subleading EW corrections and spin-correlation effects in $t\bar{t}W$ multi-lepton signatures, 2020, arXiv: 2004.09552 [hep-ph].

[249] Kulesza, Anna, Motyka, Leszek, Schwartländer, Daniel, Stebel, Tomasz, and Theeuwes, Vincent, *Associated top quark pair production with a heavy boson: differential cross sections at NLO + NNLL accuracy*, Eur. Phys. J. C **80**.5 (2020) 428, URL: https://doi.org/10.1140/epjc/s10052-020-7987-6.

[250] D. de Florian et al.,
Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector, CERN Yellow Reports (2016), arXiv: 1610.07922 [hep-ph].

[251] ATLAS Collaboration, *Measurement of the cross-section of the production of a W boson in association with a single top quark with ATLAS at $\sqrt{s} = 13$ TeV*, ATLAS-CONF-2016-065, 2016, URL: https://cds.cern.ch/record/2206207.

[252] ATLAS Collaboration, *Observation of the associated production of a top quark and a Z boson in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, JHEP **07** (2020) 124, arXiv: 2002.07546 [hep-ex].

[253] ATLAS Collaboration,
Search for flavor-changing neutral-current couplings between the top quark and the Z boson with LHC Run 2 proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector, Phys.Rev.D (2023), arXiv: 2301.11605 [hep-ex].

[254] ATLAS Collaboration, *Measurements of $Z\gamma$ + jets differential cross sections in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, JHEP (2022), arXiv: 2212.07184 [hep-ex].

[255] 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.Journal C **79** (2019), arXiv: 1812.01697 [hep-ex].

[256] ATLAS Collaboration, *Search for leptonic charge assymetry in $t\bar{t}W$ production in final states with three leptons at $\sqrt{s} = 13$ TeV*, JHEP07 (2023), arXiv: 2301.04245 [hep-ex].

[257] ATLAS Collaboration,
Search for flavor-changing neutral-current couplings between the top quark and the Z boson with LHC Run 2 proton–proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector, ATLAS-CONF-2021-049, 2021, URL: https://cds.cern.ch/record/2781174.

[258] ATLAS Collaboration, *Measurement of the $t\bar{t}Z$ and $t\bar{t}W$ cross sections in proton–proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, Phys. Rev. D **99** (2019) 072009, arXiv: 1901.03584 [hep-ex].

[259] Glen Cowan, Kyle Cranmer, Eilam Gross, and Ofer Vitells,
Asymptotic formulae for likelihood-based tests of new physics, Eur. Phys. J. C **71** (2011) 1554, arXiv: 1007.1727 [physics.data-an], Erratum: Eur. Phys. J. C **73** (2013) 2501.

[260] J.S. Conway,
Incorporating Nuisance Parameters in Likelihoods for Multisource Spectra, PHYSTAT2011, University of California, 2011, URL: https://indico.cern.ch/event/107747/contributions/32677/attachments/24368/35057/conway.pdf.

[261] Li-Gang Xia,
Study of constraint and impact of a nuisance parameter in maximum likelihood method, J. Phys. G: Nucl. Part. Phys. **46**.085004 (2019), arXiv: 1805.03961v4 [physics.data-an], URL: https://doi.org/10.48550/arXiv.1805.03961.

[262] Kyle Cranmer, *Practical Statistics for the LHC*, CERN-2014-003 (2015), arXiv: 1503.07622v1 [physics.data-an], URL: https://doi.org/10.48550/arXiv.1503.07622.

[263] *TRExFitter documentation*, 2021, URL: https://trexfitter-docs.web.cern.ch/trexfitter-docs/.

[264] *Statistical Methods for the LHC*, URL: https://atlas-stats-doc-dev.web.cern.ch/atlas-stats-doc-dev/statisticaltests/.

[265] S. S. Wilks,
The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses, The Annals of Mathematical Statistics **9**.1 (1938) 60, URL: https://doi.org/10.1214/aoms/1177732360.

[266] Abraham Wald, *On Cumulative Sums of Random Variables*, The Annals of Mathematical Statistics **15**.3 (1944) 283, URL: https://doi.org/10.1214/aoms/1177731235.

[267] Alexander L. Read, *Presentation of search results: the CL_S technique*, J. Phys. G **28** (2002) 2693.

[268] Isaac Asimov, *The Complete Stories*, vol. 1, Doubleday Foundation, 1990, ISBN: 0-385-41606-7.

[269] Kyle Cranmer, George Lewis, Lorenzo Moneta, Akira Shibata, and Wouter Verkerke,
HistFactory: A tool for creating statistical models for use with RooFit and RooStats, CERN-OPEN-2012-016, 2012, URL: https://cds.cern.ch/record/1456844.

[270] Wouter Verkerke and David P. Kirkby, *The RooFit toolkit for data modelling*, CHEP-2003-MOLT007, 2003 MOLT007, arXiv: physics/0306116.

[271] Lorenzo Moneta et al., *The RooStats Project*, PoS **ACAT2010** (2010) 057, ed. by T. Speer et al., arXiv: 1009.1003 [physics.data-an].

[272] R. Brun and F. Rademakers, *ROOT: An object oriented data analysis framework*, Nucl. Instrum. Meth. A **389** (1997) 81, ed. by M. Werlen and D. Perret-Gallix.

[273] I. Antcheva et al.,
ROOT - A C++ Framework for Petabyte Data Storage, Statistical Analysis and Visualization, Computer Physics Communications **180**.12 (2009), arXiv: 1508.07749v1 [physics.data-an], URL: https://doi.org/10.48550/arXiv.1508.07749.

[274] F. James and M. Roos, *Minuit: A System for Function Minimization and Analysis of the Parameter Errors and Correlations*, Comput. Phys. Commun. **10** (1975) 343.

[275] F. James,
MINUIT Function Minimization and Error Analysis: Reference Manual Version 94.1,
CERN-D-506 (1994), URL: https://inspirehep.net/literature/1258343.

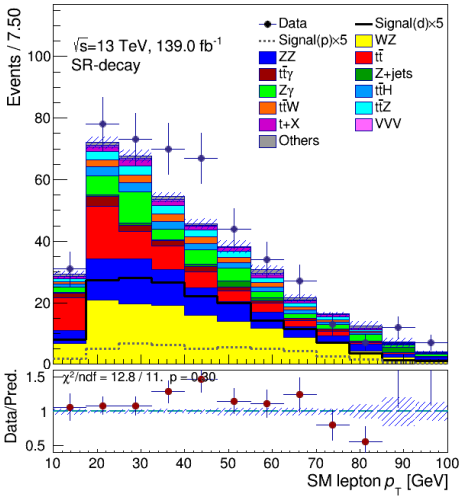
[276] Fred James and Matthias Winkler, *MINUIT User's Guide*, 2004,
URL: https://inspirehep.net/literature/1258345.

Appendices

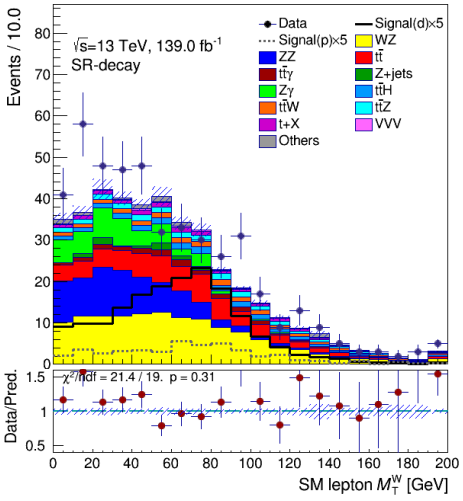
Appendix A

Kinematic distributions in the analysis regions

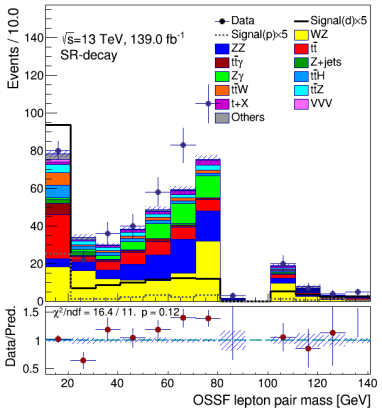
Distributions of kinematic variables in the analysis regions, defined in section 5.4, are given from fig. A.1 to fig. A.14, starting from reconstructed variables in the SR-decay region and dilepton-mass distributions in all SRs, then proceeding to p_{T} distributions of leptons and jets, as well as jet multiplicities, in all regions. The error bars and hatched areas represent statistical uncertainties only.



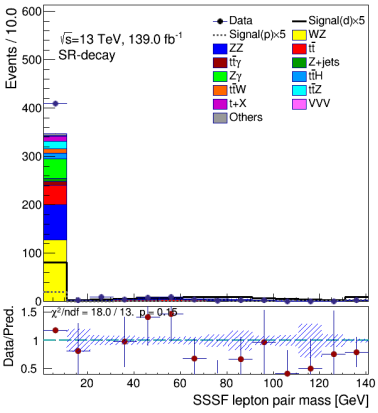
(a) SM lepton p_{T}



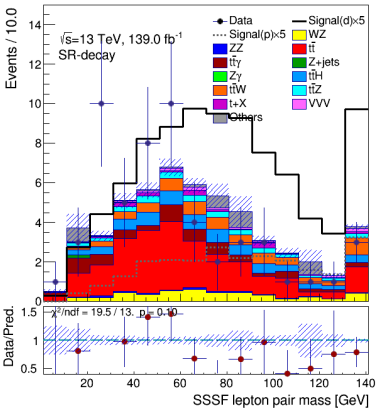
(b) W transverse mass of SM lepton



(c) OSSF dilepton inv. mass

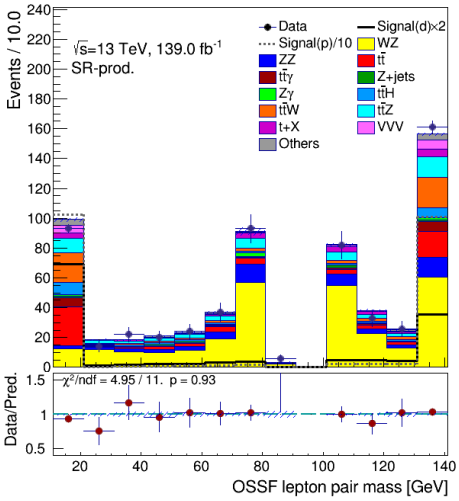


(d) SSSF dilepton inv. mass

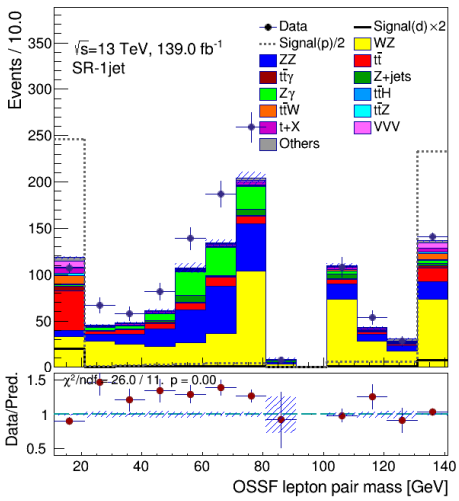


(e) SSSF dilepton inv. mass, no underflow

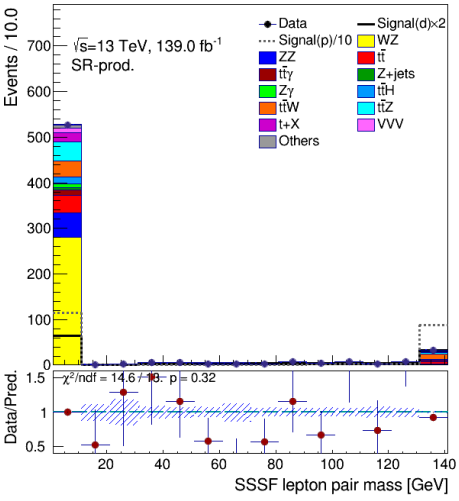
Figure A.1: Distributions of kinematic variables related to the top quarks' reconstruction (a,b) and dilepton invariant masses (c-e) in the SR-decay region. The underflow/overflow are added in the first/last bin, apart from plot (e), where the underflow is not added (provided for better illustration). When no OSSF or SSSF dilepton is found, then the underflow is filled in the respective plots (default value zero). It is obvious that for dibosons and $Z+\mathrm{jets}(+\gamma)$, we mostly do not have a SSSF lepton pair, while for $t\bar{t}$ this is "balanced", i.e. the Z gives an OSSF lepton pair, while the SRs demand an additional (3^{rd}) lepton of different flavour.



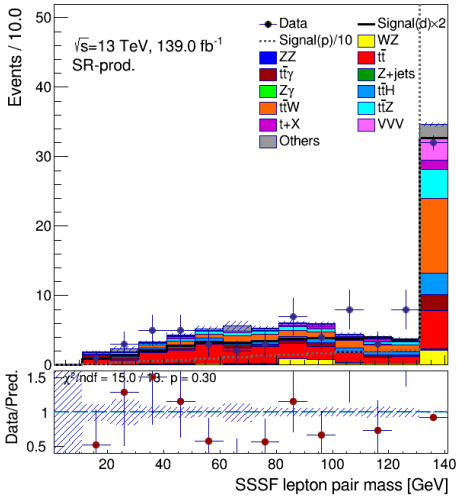
(a) OSSF dilepton inv. mass



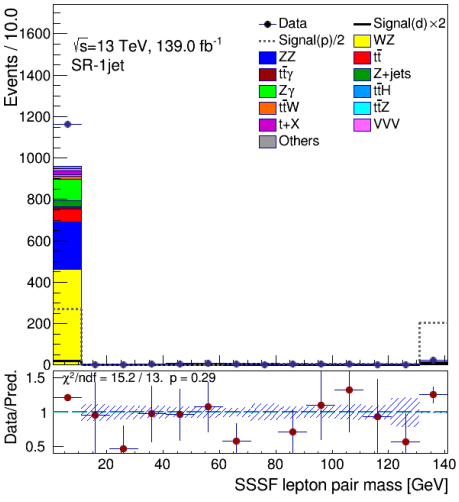
(b) OSSF dilepton inv. mass



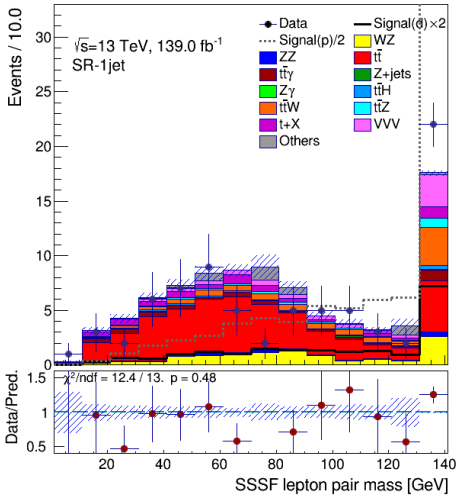
(c) SSSF dilepton inv. mass



(d) SSSF dilepton inv. mass, no underflow

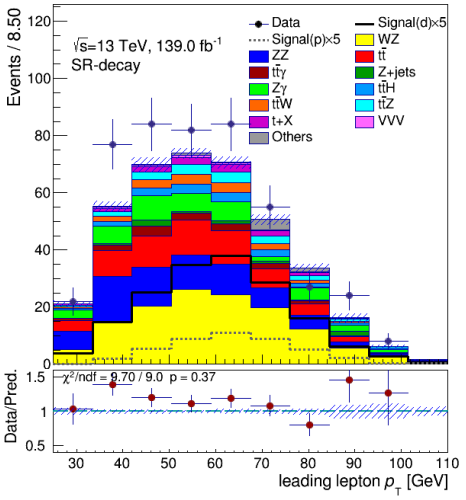


(e) SSSF dilepton inv. mass

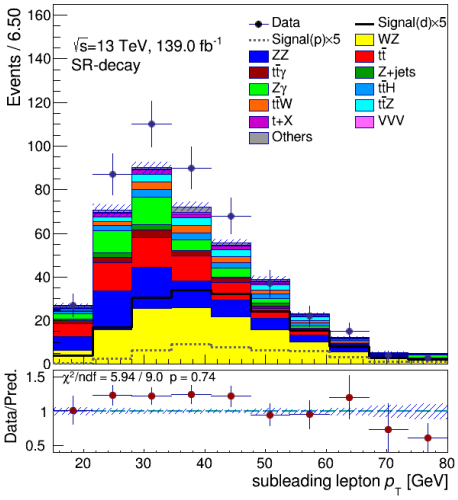


(f) SSSF dilepton inv. mass, no underflow

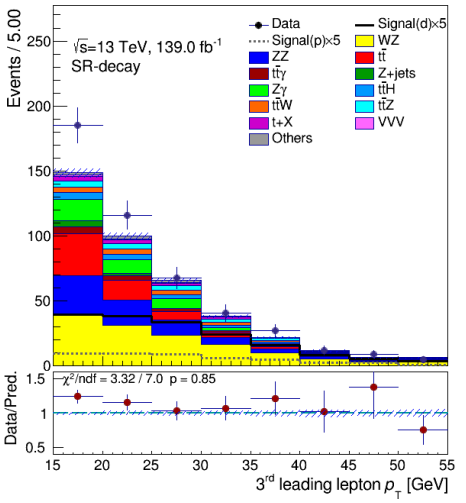
Figure A.2: Distributions of dilepton invariant masses in SR-production and SR-1jet regions. The underflow/overflow are added in the first/last bin, apart from (d,f), where the underflow is not added (provided for better illustration). When no OSSF/SSSF dilepton is found, then the underflow is filled in the respective plots (default value zero).



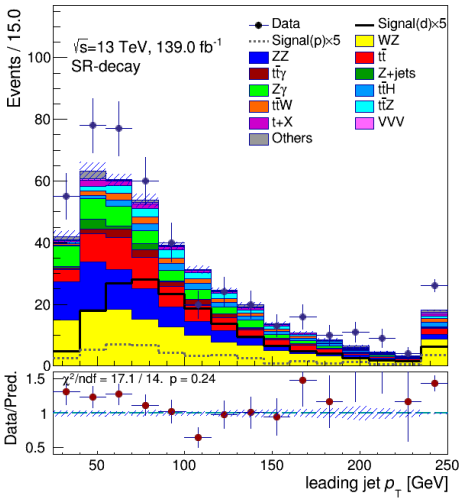
(a) Leading lepton p_{T}



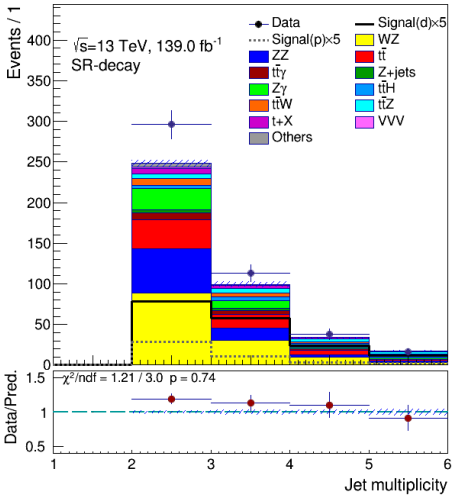
(b) Sub-leading lepton p_{T}



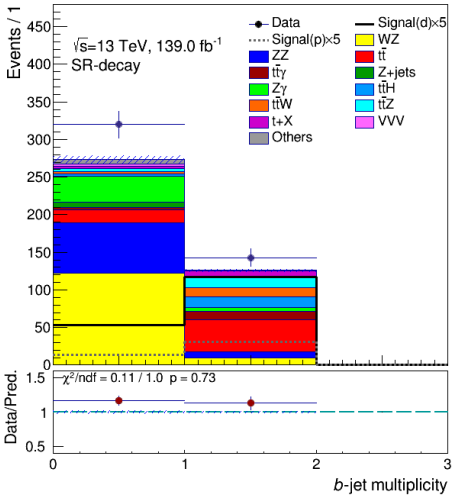
(c) Third leading lepton p_{T}



(d) Leading jet p_{T}

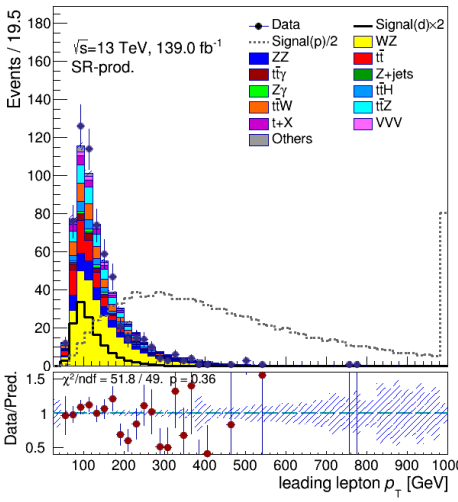


(e) Jet multiplicity

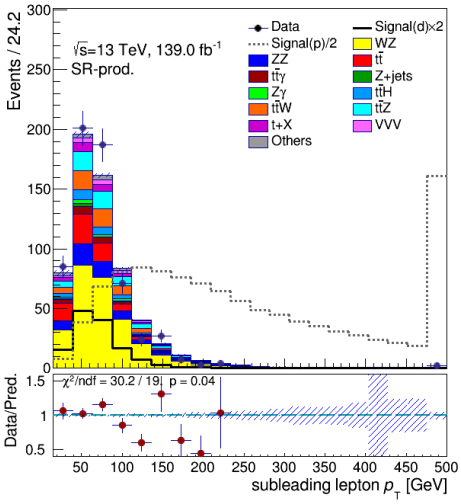


(f) b-jet multiplicity

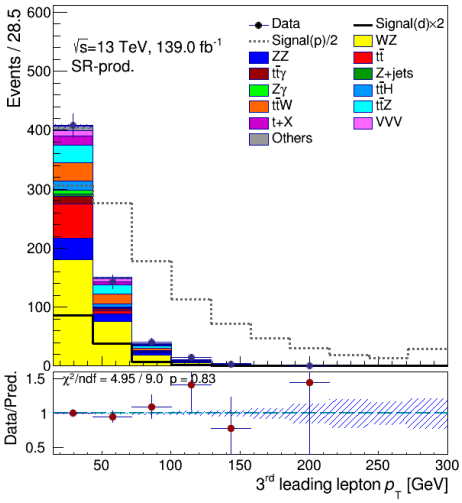
Figure A.3: Distributions of p_{T} variables of leptons and jets, as well as jet multiplicities, in SR-decay region.



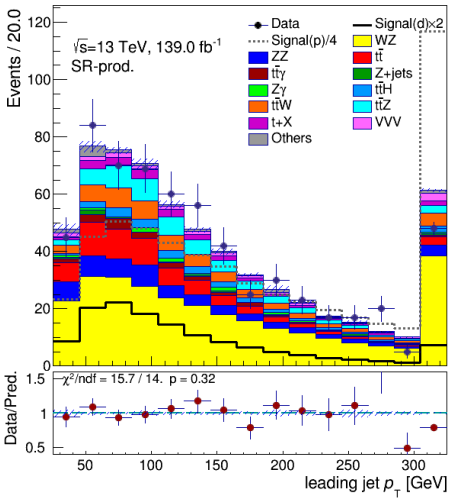
(a) Leading lepton p_{T}



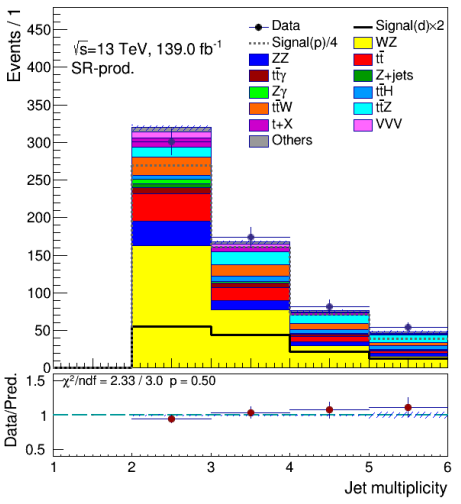
(b) Sub-leading lepton p_{T}



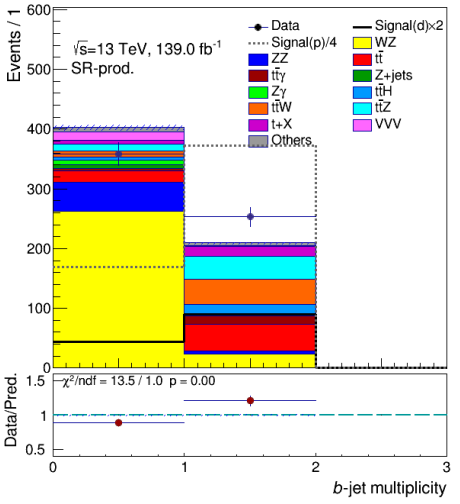
(c) Third leading lepton p_{T}



(d) Leading jet p_{T}

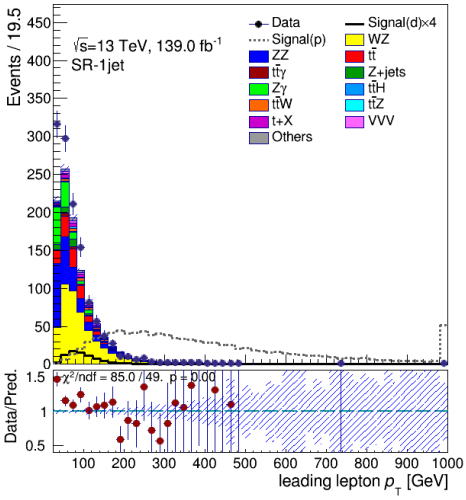


(e) Jet multiplicity

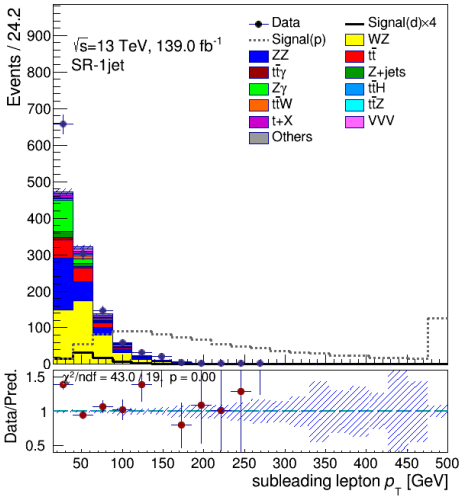


(f) b-jet multiplicity

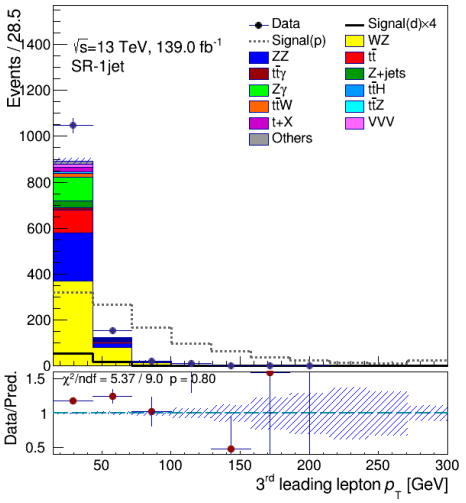
Figure A.4: Distributions of p_{T} variables of leptons and jets, as well as jet multiplicities, in SR-production region.



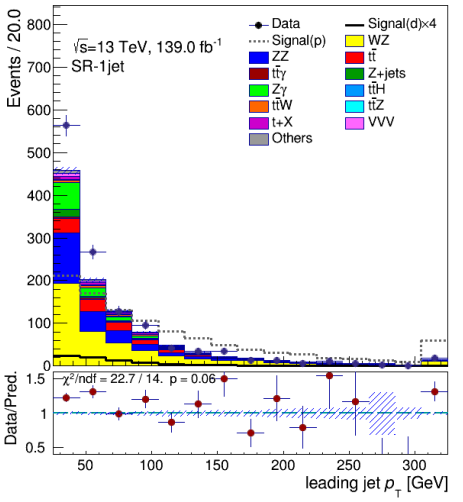
(a) Leading lepton p_{T}



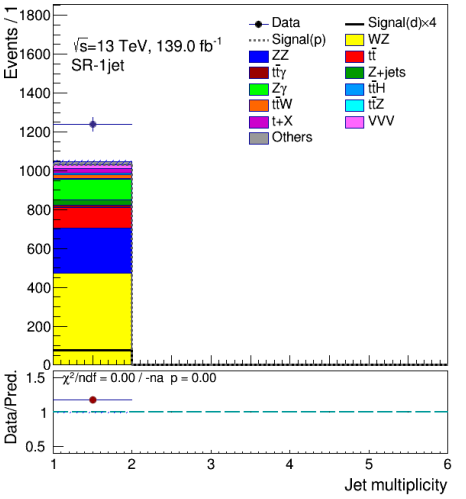
(b) Sub-leading lepton p_{T}



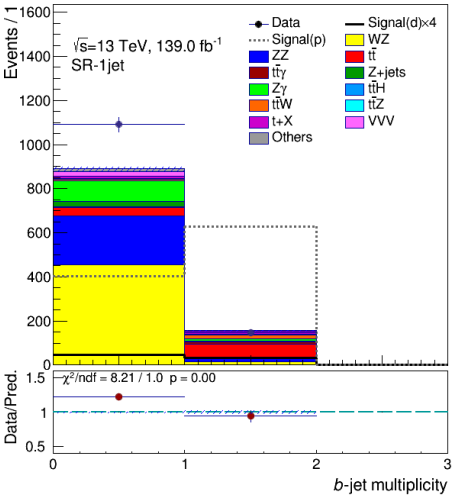
(c) Third leading lepton p_{T}



(d) Leading jet p_{T}

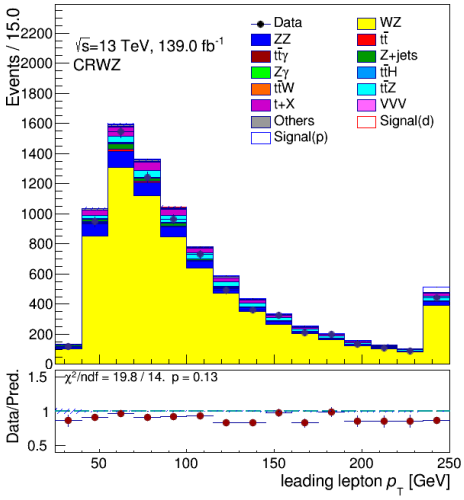


(e) Jet multiplicity

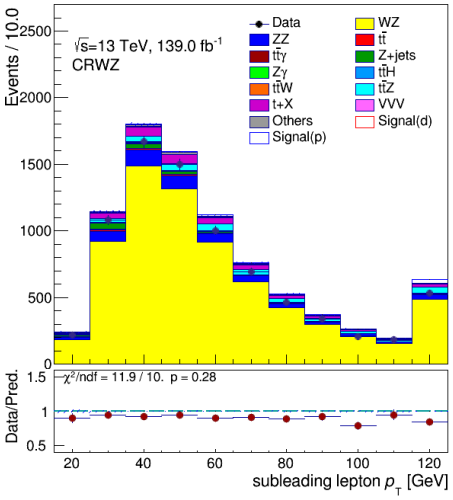


(f) b-jet multiplicity

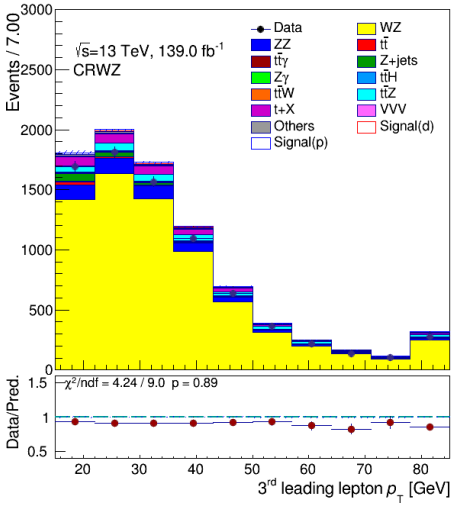
Figure A.5: Distributions of p_{T} variables of leptons and jets, as well as jet multiplicities, in SR-1jet region.



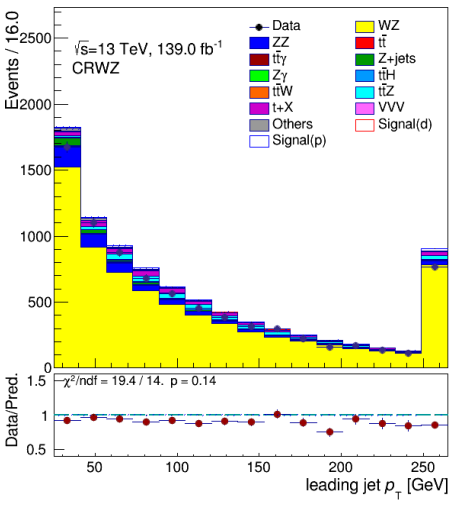
(a) Leading lepton p_{T}



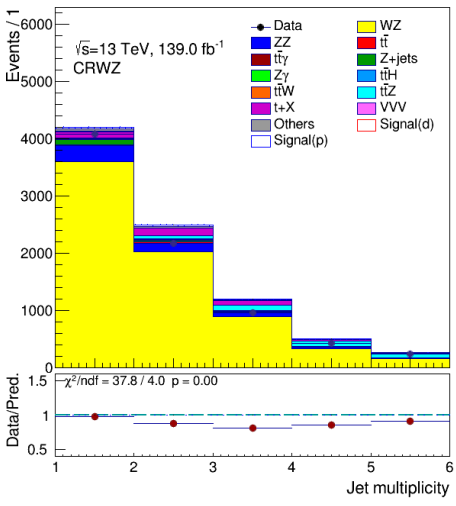
(b) Sub-leading lepton p_{T}



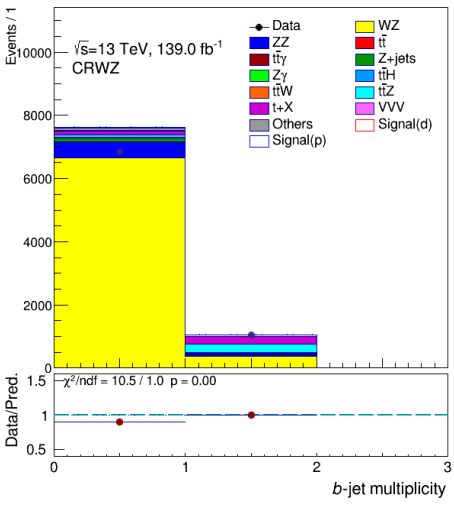
(c) Third leading lepton p_{T}



(d) Leading jet p_{T}

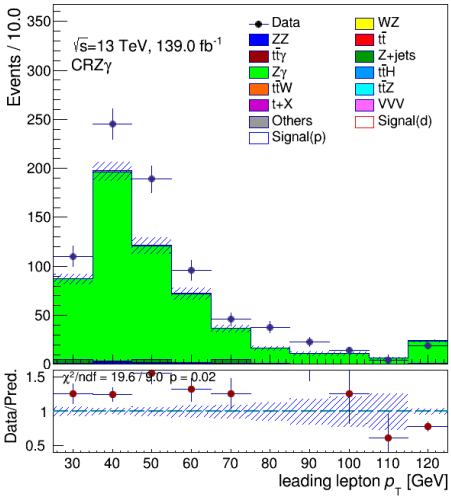


(e) Jet multiplicity

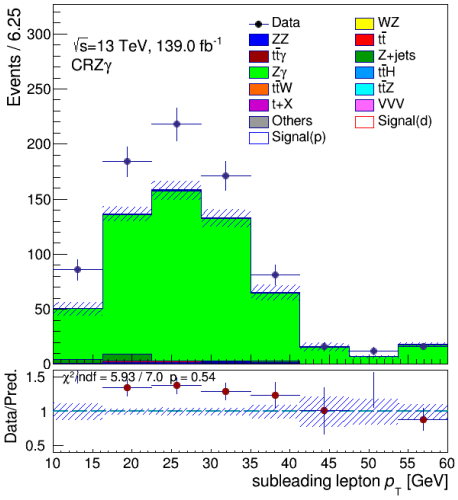


(f) b-jet multiplicity

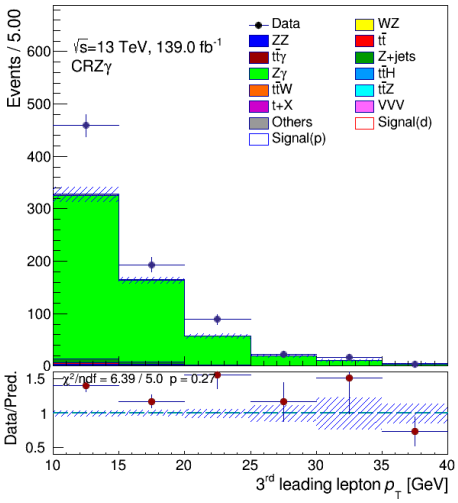
Figure A.6: Distributions of p_{T} variables of leptons and jets, as well as jet multiplicities, in CRWZ region.



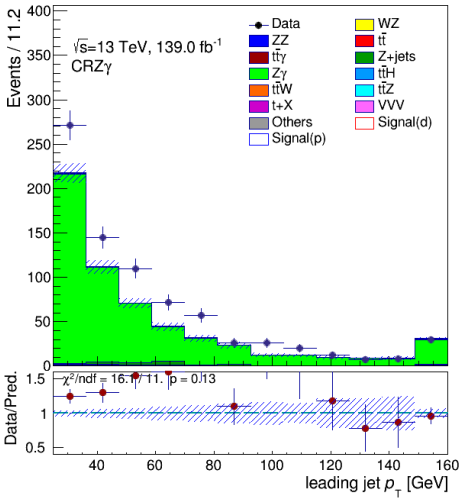
(a) Leading lepton p_{T}



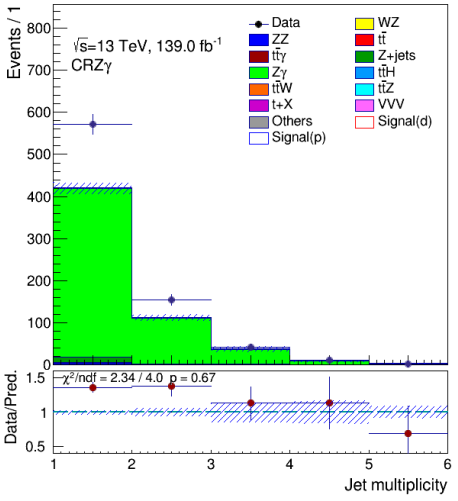
(b) Sub-leading lepton p_{T}



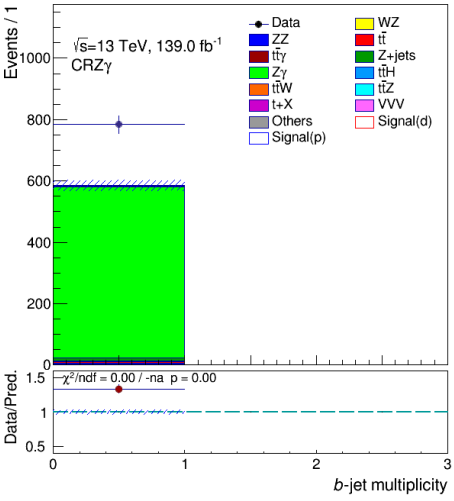
(c) Third leading lepton p_{T}



(d) Leading jet p_{T}

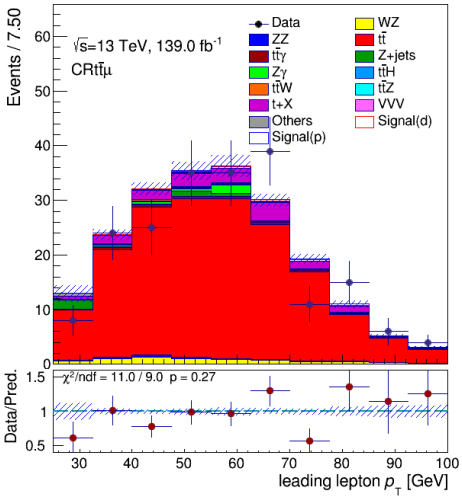


(e) Jet multiplicity

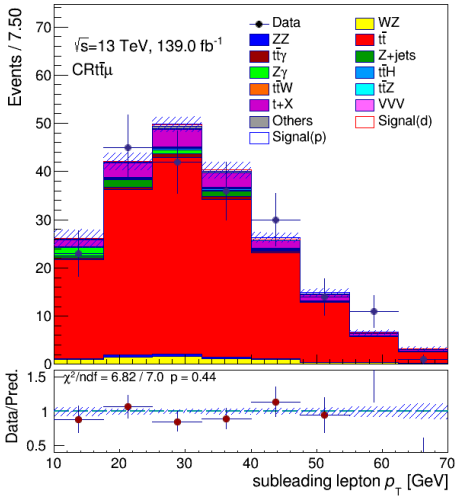


(f) b-jet multiplicity

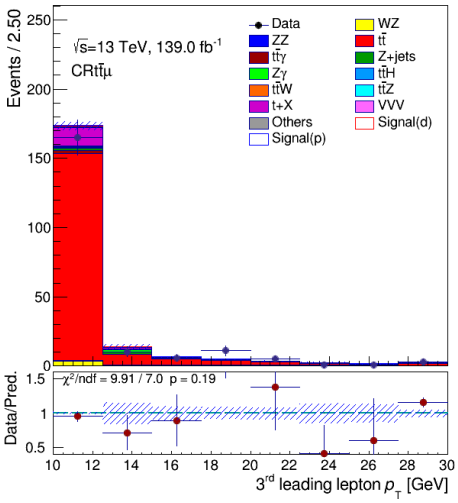
Figure A.7: Distributions of p_{T} variables of leptons and jets, as well as jet multiplicities, in CR$Z\gamma$ region.



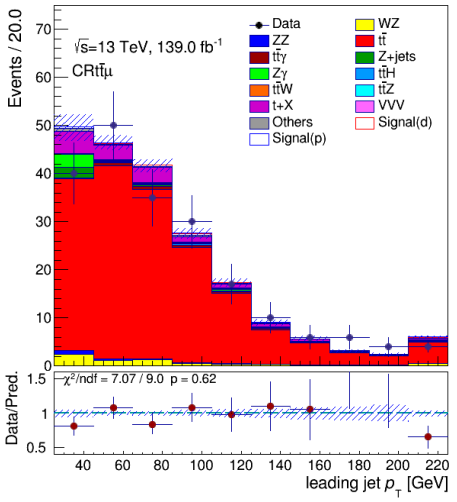
(a) Leading lepton p_{T}



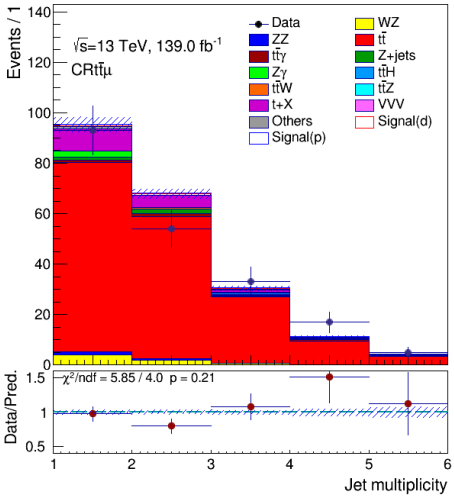
(b) Sub-leading lepton p_{T}



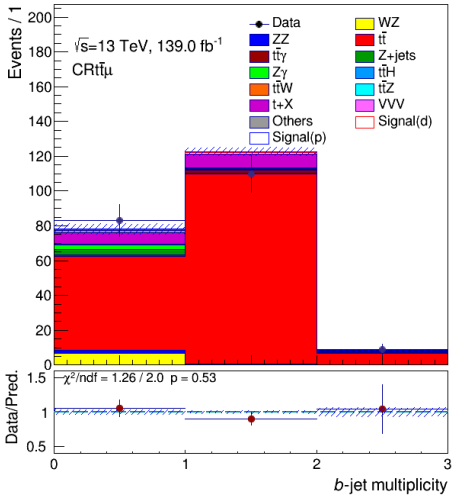
(c) Third leading lepton p_{T}



(d) Leading jet p_{T}

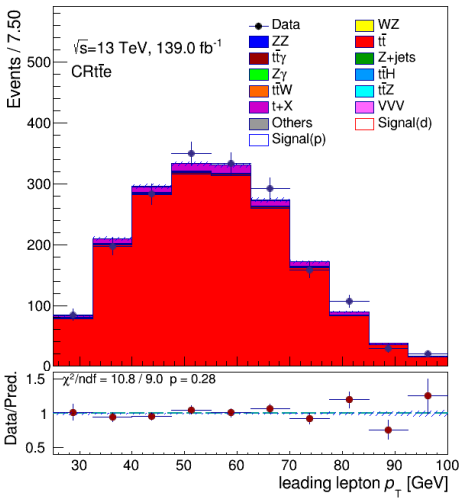


(e) Jet multiplicity

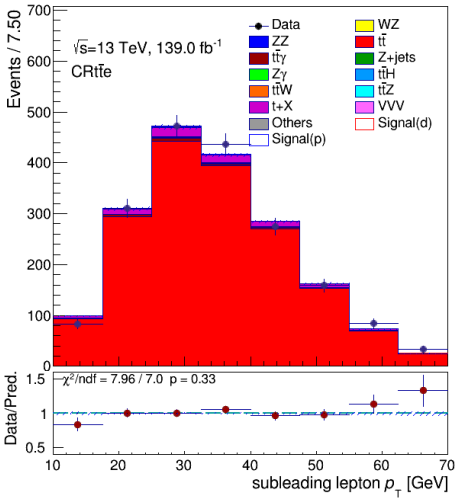


(f) b-jet multiplicity

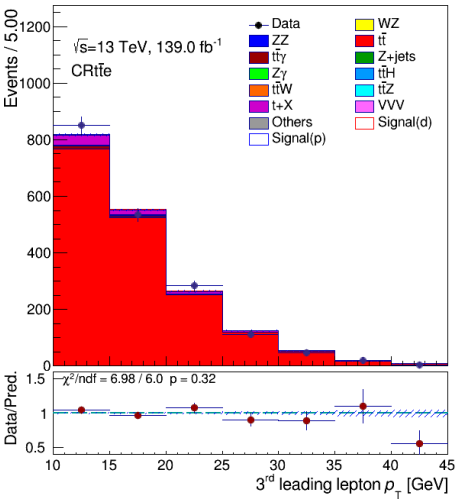
Figure A.8: Distributions of p_{T} variables of leptons and jets, as well as jet multiplicities, in CR$t\bar{t}\mu$ region.



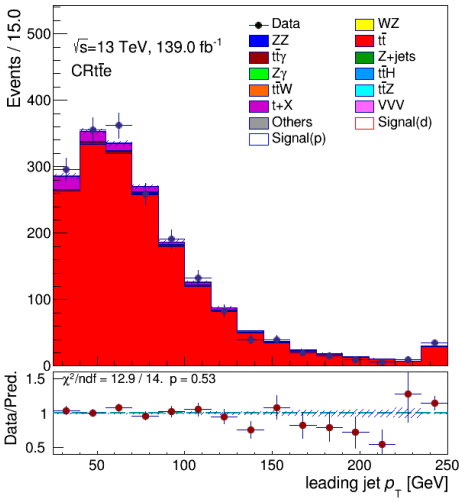
(a) Leading lepton p_{T}



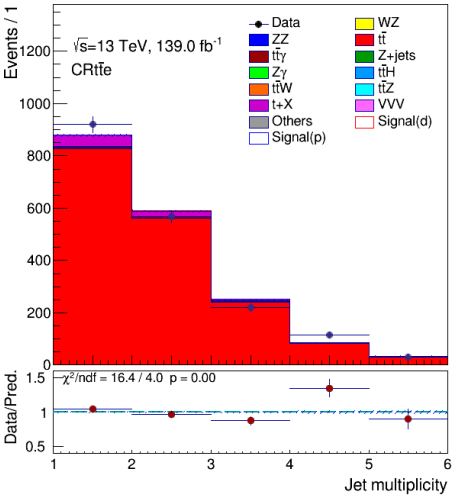
(b) Sub-leading lepton p_{T}



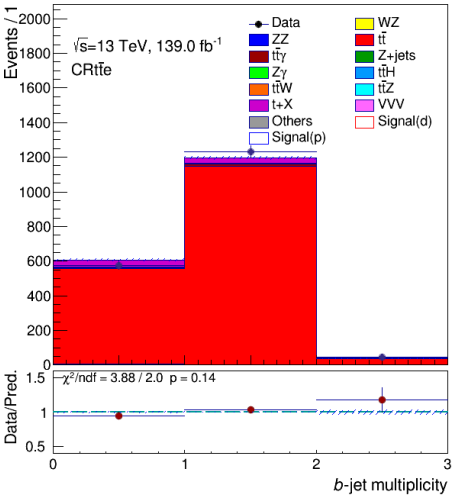
(c) Third leading lepton p_{T}



(d) Leading jet p_{T}

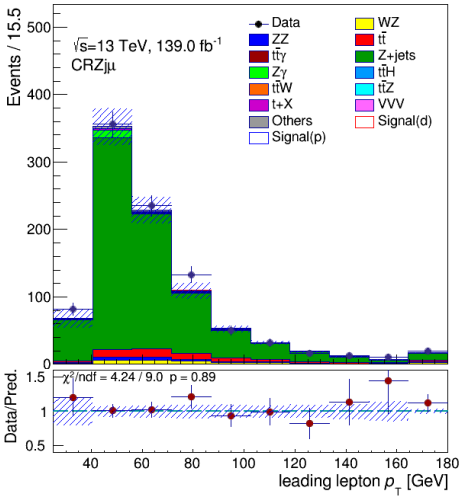


(e) Jet multiplicity

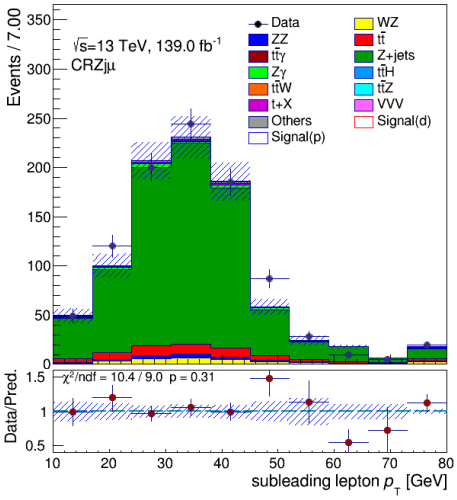


(f) b-jet multiplicity

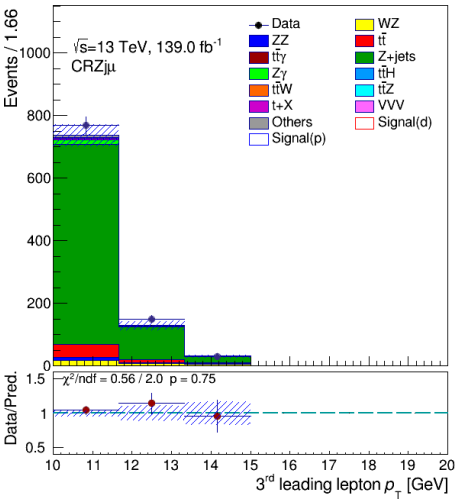
Figure A.9: Distributions of p_{T} variables of leptons and jets, as well as jet multiplicities, in CR$t\bar{t}e$ region.



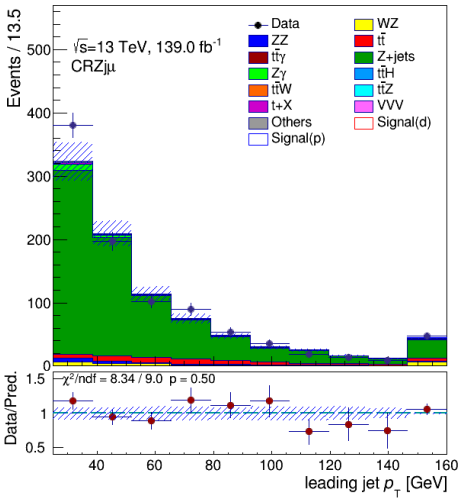
(a) Leading lepton p_{T}



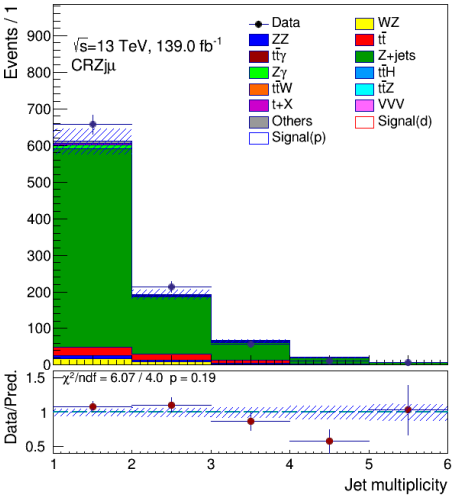
(b) Sub-leading lepton p_{T}



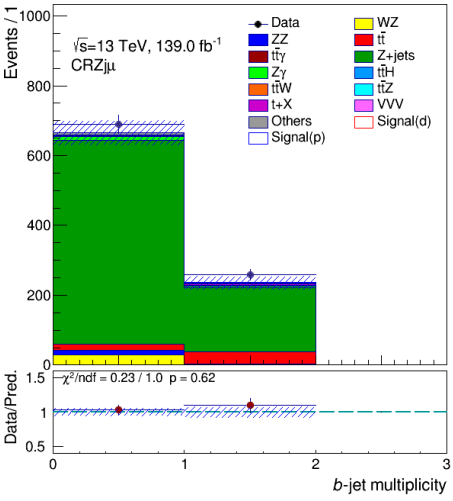
(c) Third leading lepton p_{T}



(d) Leading jet p_{T}

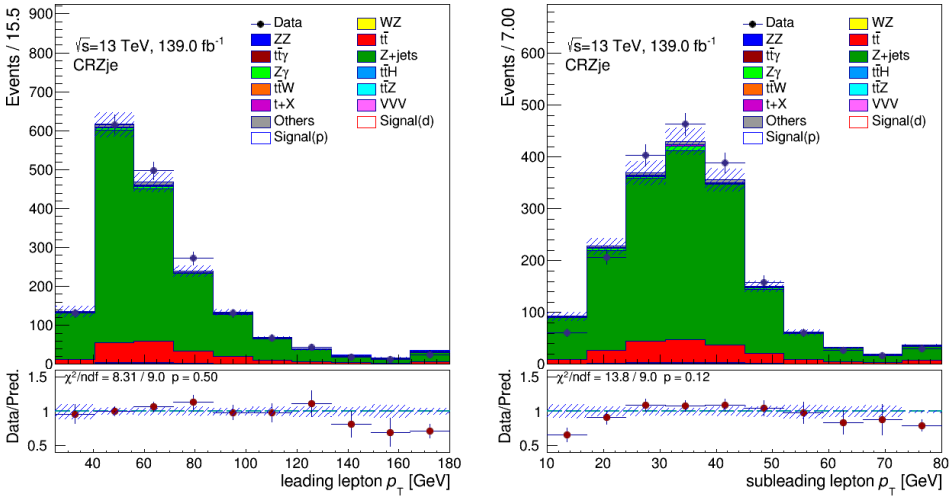


(e) Jet multiplicity

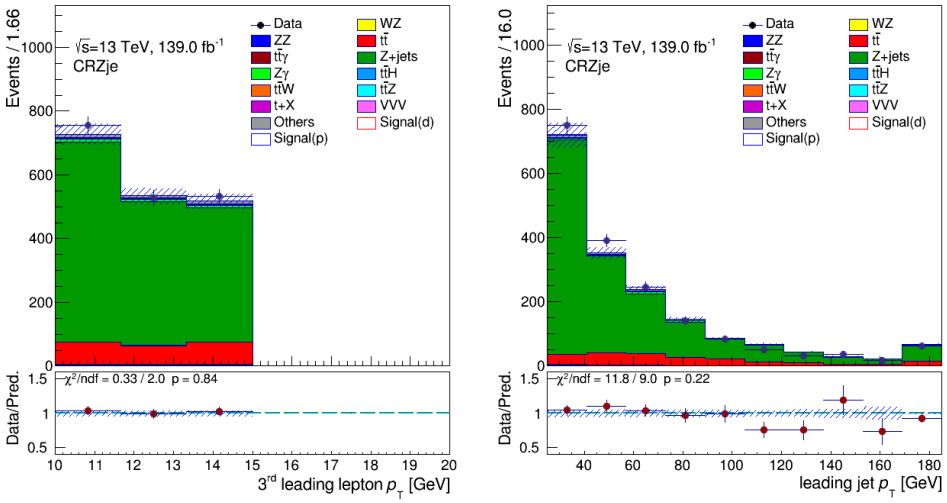


(f) b-jet multiplicity

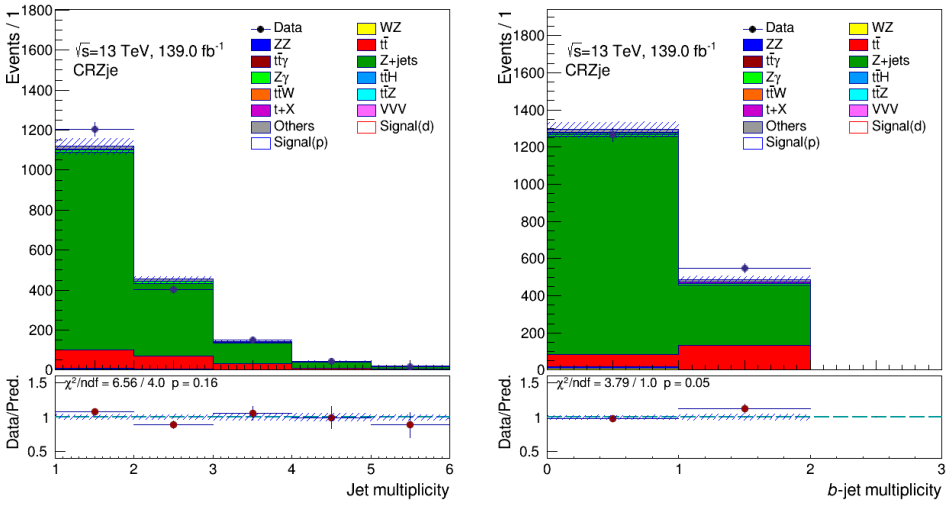
Figure A.10: Distributions of p_{T} variables of leptons and jets, as well as jet multiplicities, in CR$Zj\mu$ region.



(a) Leading lepton p_{T} (b) Sub-leading lepton p_{T}

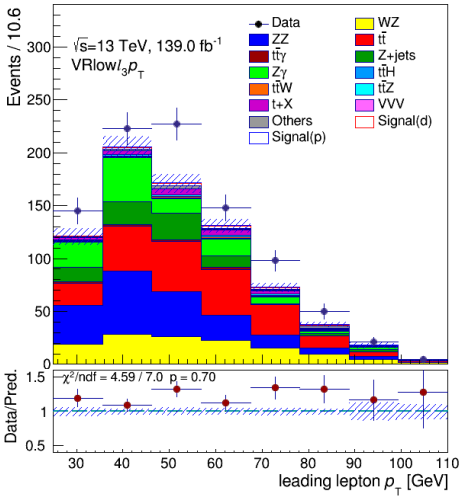


(c) Third leading lepton p_{T} (d) Leading jet p_{T}

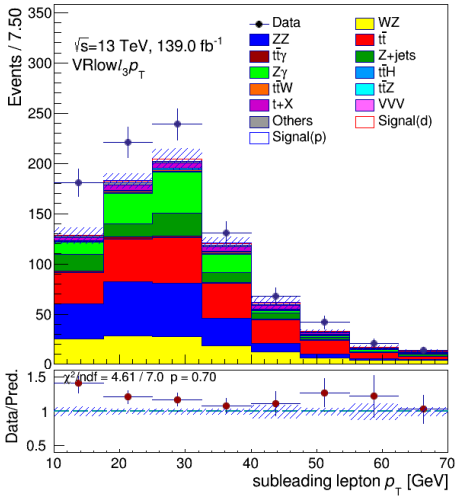


(e) Jet multiplicity (f) b-jet multiplicity

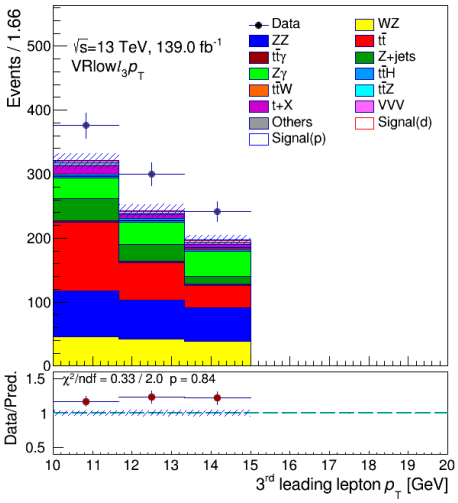
Figure A.11: Distributions of p_{T} variables of leptons and jets, as well as jet multiplicities, in CRZje region.



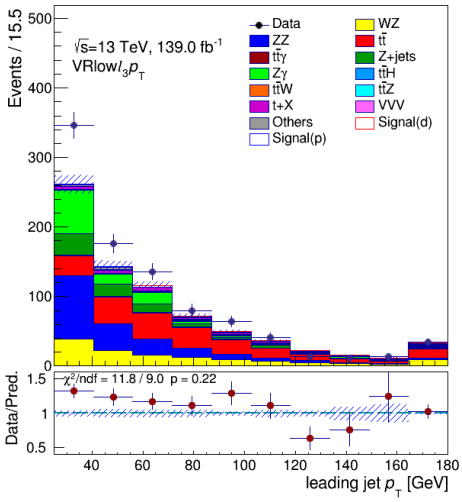
(a) Leading lepton p_{T}



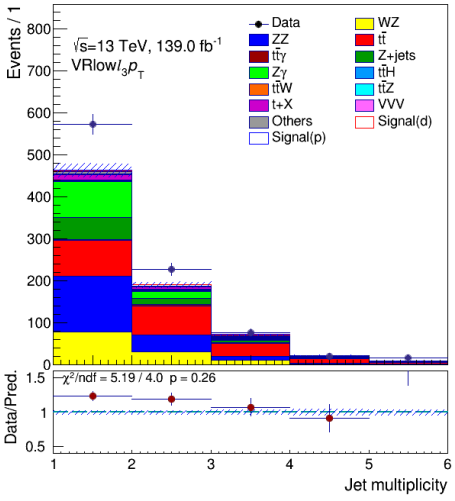
(b) Sub-leading lepton p_{T}



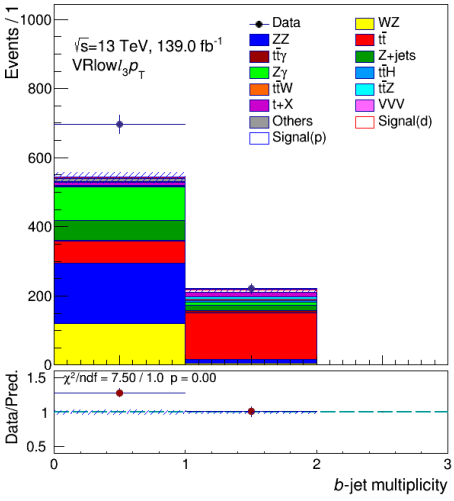
(c) Third leading lepton p_{T}



(d) Leading jet p_{T}

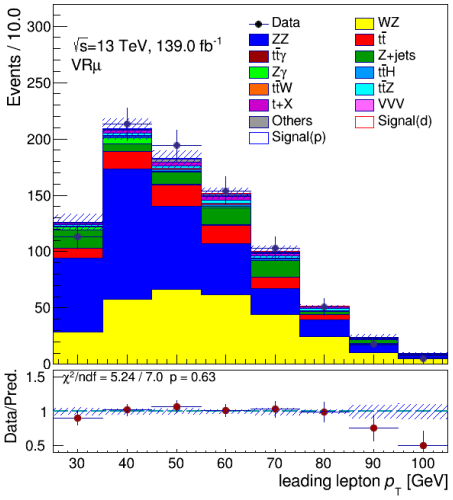


(e) Jet multiplicity

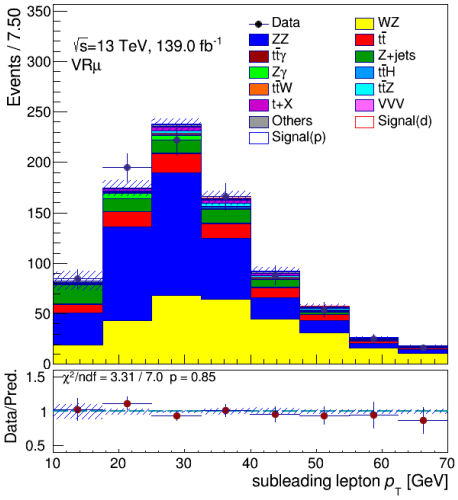


(f) b-jet multiplicity

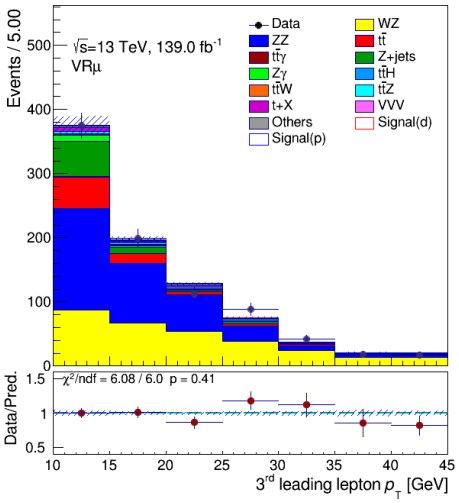
Figure A.12: Distributions of p_{T} variables of leptons and jets, as well as jet multiplicities, in VRlow$\ell_3 p_{\mathrm{T}}$ region.



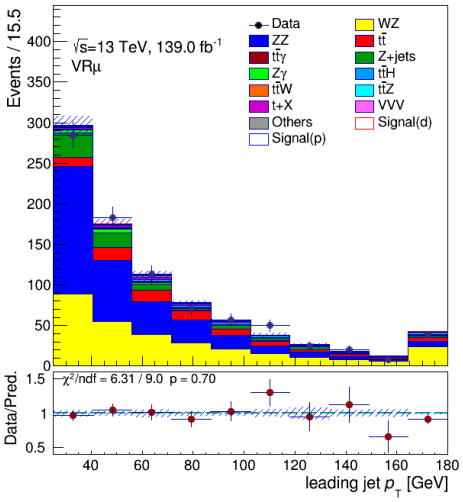
(a) Leading lepton p_{T}



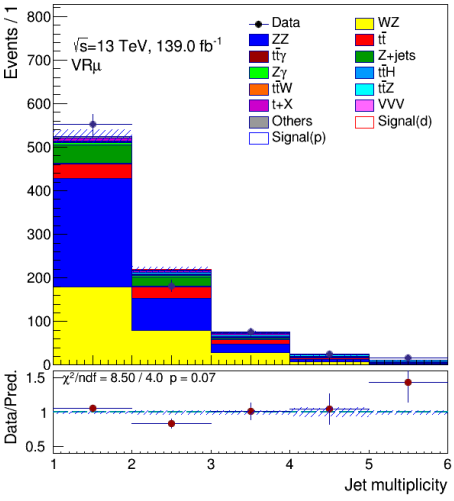
(b) Sub-leading lepton p_{T}



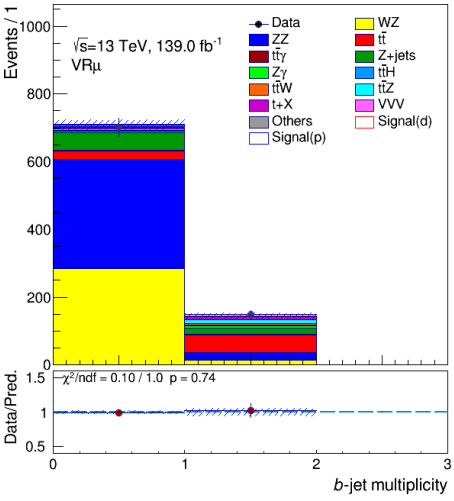
(c) Third leading lepton p_{T}



(d) Leading jet p_{T}

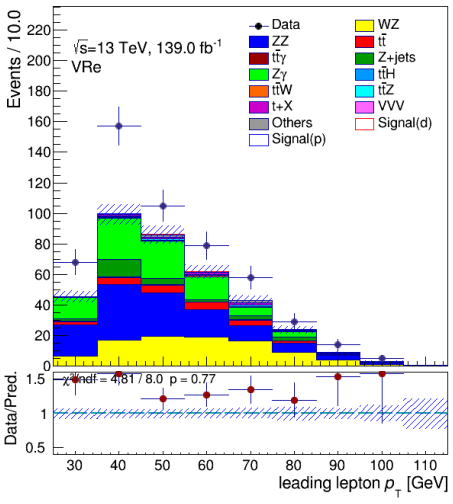


(e) Jet multiplicity

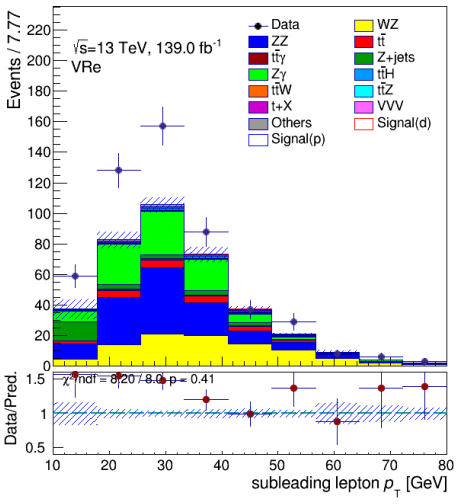


(f) b-jet multiplicity

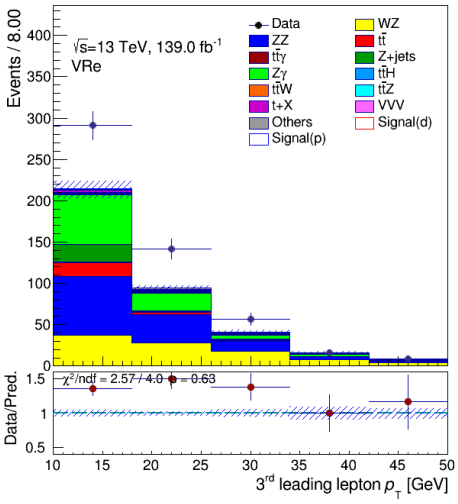
Figure A.13: Distributions of p_{T} variables of leptons and jets, as well as jet multiplicities, in VRμ region.



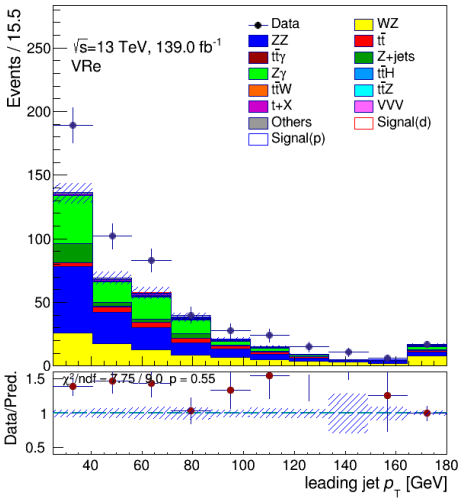
(a) Leading lepton p_{T}



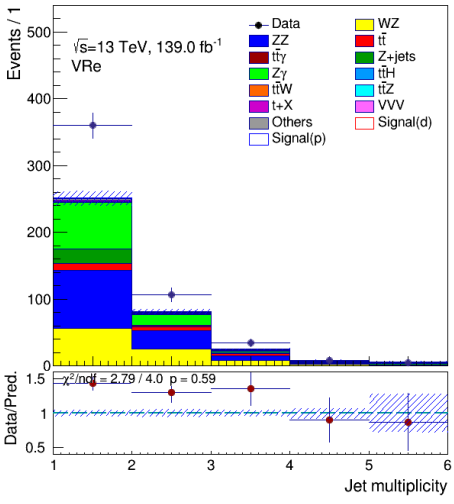
(b) Sub-leading lepton p_{T}



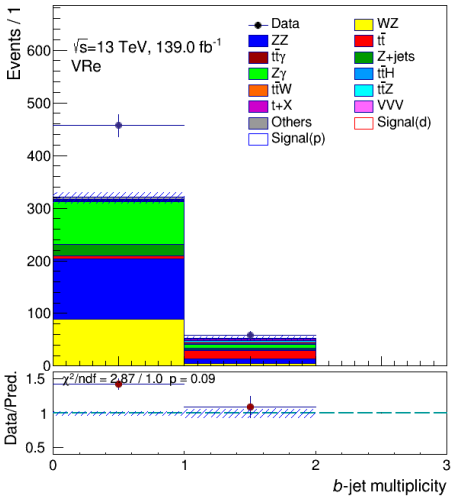
(c) Third leading lepton p_{T}



(d) Leading jet p_{T}



(e) Jet multiplicity



(f) b-jet multiplicity

Figure A.14: Distributions of p_{T} variables of leptons and jets, as well as jet multiplicities, in VRe region.

Appendix B

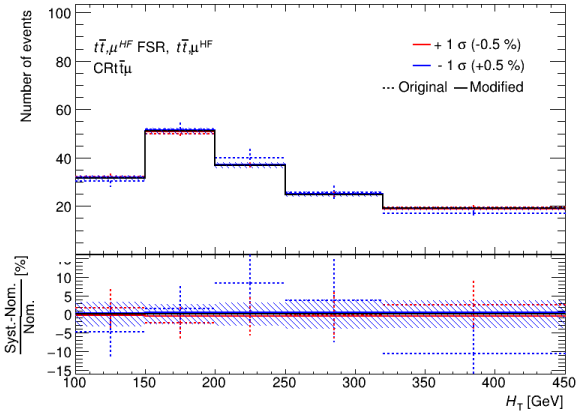
Statistical Analysis

B.1 Theory systematic uncertainties before and after smoothing and symmetrisation

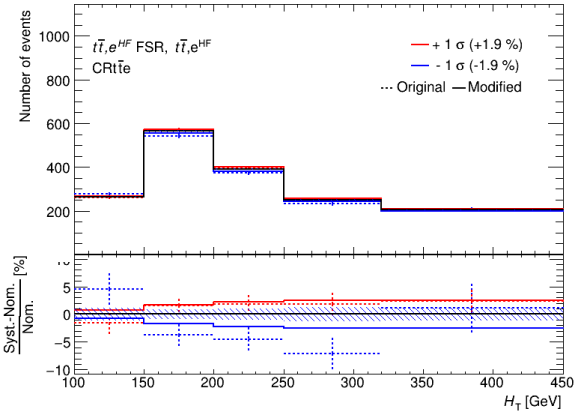
Additional plots of important modelling systematic uncertainties on $t\bar{t}$, $t\bar{t}Z$ and ZZ background sources, shown in fig. B.1, fig. B.2 and figs. B.3a to B.3d respectively, are presented. Finally, some important modelling systematic uncertainties on the production-signal process (figs. B.3e and B.3f) are also provided.

B.2 Fully Asimov Fit

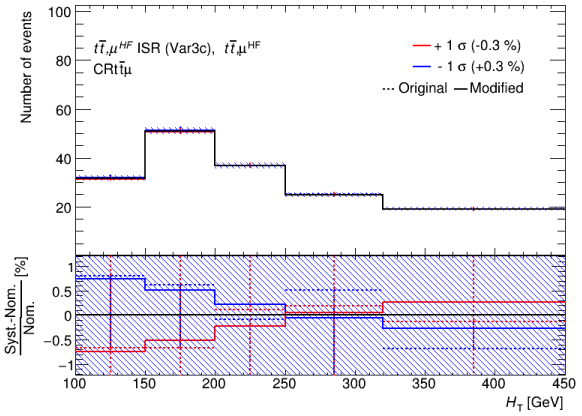
The NP correlation matrix, fig. B.4, and the NP pull plots, fig. B.5, are presented for the fit which uses Asimov data in both CRs and SRs. The plots are provided for comparison with the respective plots obtained from the "mixed" data (in CRs) and Asimov data (in SRs) fit (fig. 5.31 and fig. 5.32). A compatibility between the two cases means that the constraints and correlations observed are expected from the fully Asimov fit, and the pulls in the "mixed" setup are due to the fit trying to compensate for data/MC disagreements in the CRs. The NP ranking is also provided, fig. B.6, and is compatible with the "mixed" fit setup (fig. 5.42(a)).



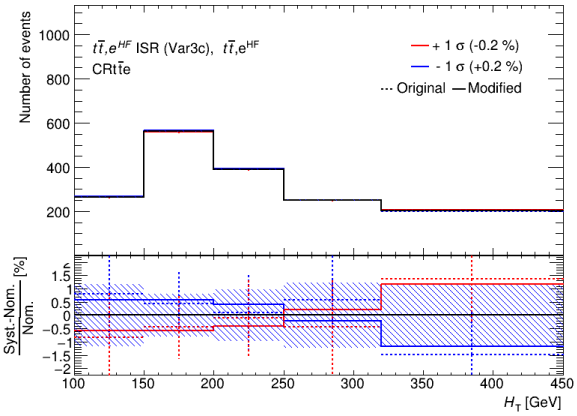
(a) $t\bar{t}$,HFμ FSR uncertainty.



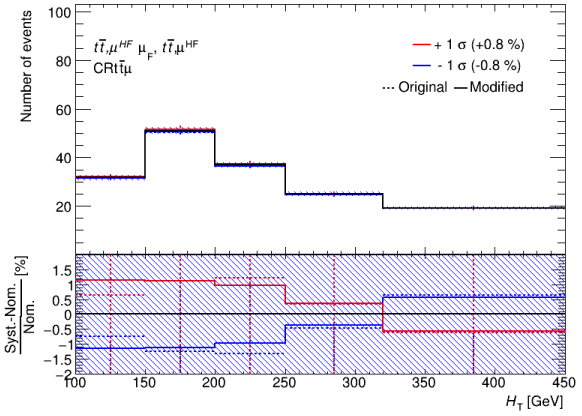
(b) $t\bar{t}$,HFe FSR uncertainty.



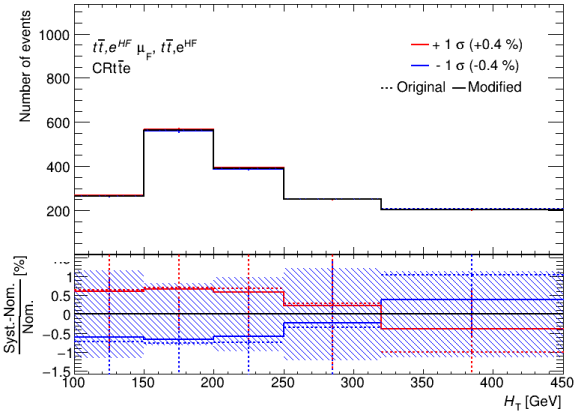
(c) $t\bar{t}$,HFμ ISR uncertainty.



(d) $t\bar{t}$,HFe ISR uncertainty.

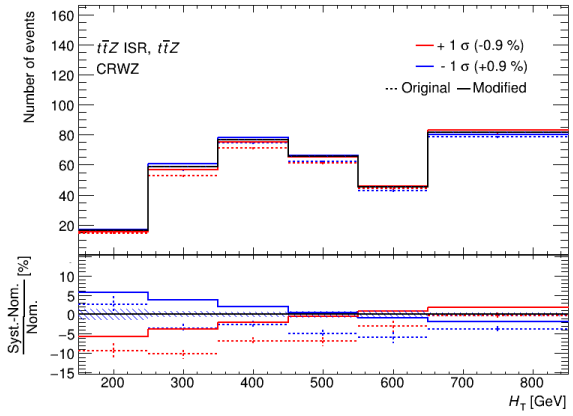


(e) $t\bar{t}$,HFμ μ_F uncertainty.

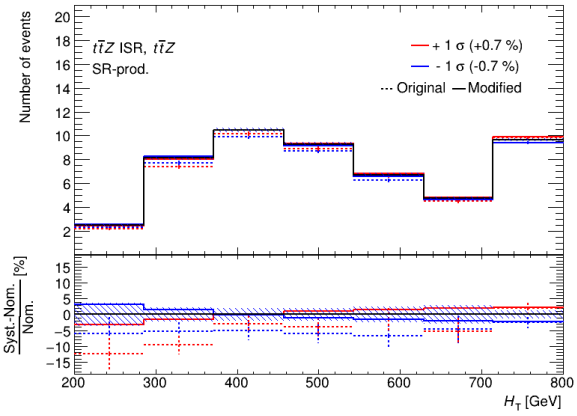


(f) $t\bar{t}$,HFe μ_F uncertainty.

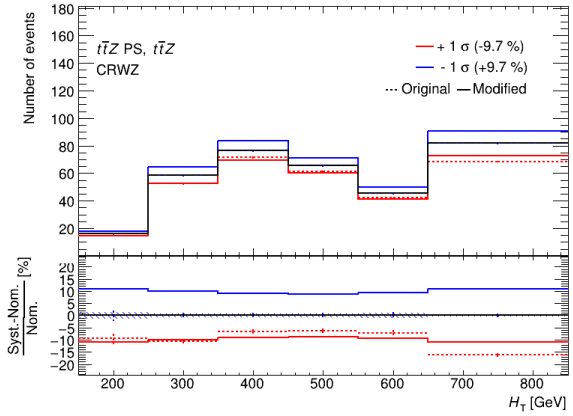
Figure B.1: Plots in the control regions targeting $t\bar{t}$, showing the effect of $\pm 1\sigma$ variations of FSR (a, b), ISR (c, d) and μ_F (e, f) modelling nuisance parameters before and after the application of smoothing and symmetrisation by TRExFitter. The continuous (original) and dashed (modified) lines refer to the up/down (red/blue) variations. The black histogram line, in the upper panel, corresponds to the nominal sample. `TWOSIDED` symmetrisation has been used for the fit, while the `MAXIMUM` has been tested and resulted in similar results on the signal strength limit setting. The hatched area represents the statistical uncertainty (divided by the MC prediction in the ratio pad).



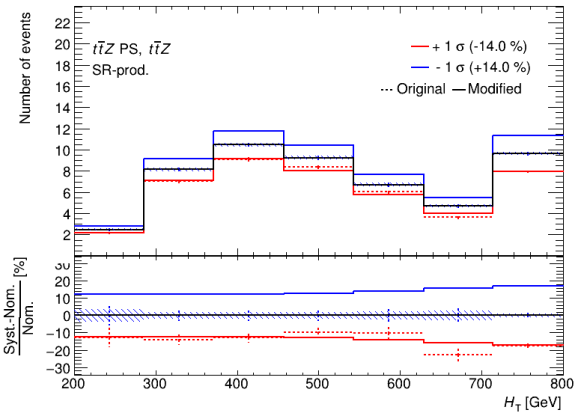
(a) $t\bar{t}Z$ ISR uncertainty in CRWZ.



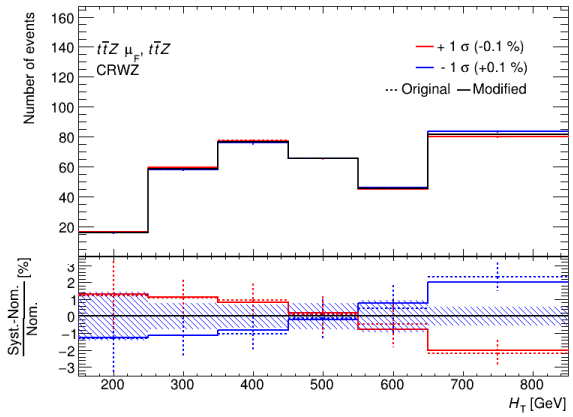
(b) $t\bar{t}Z$ ISR uncertainty in SR-prod.



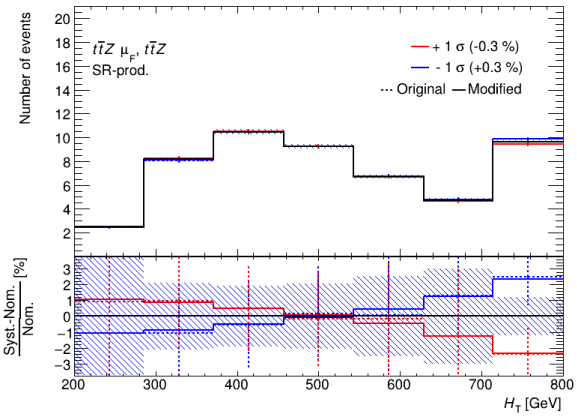
(c) $t\bar{t}Z$ PS uncertainty in CRWZ.



(d) $t\bar{t}Z$ PS uncertainty in SR-prod.

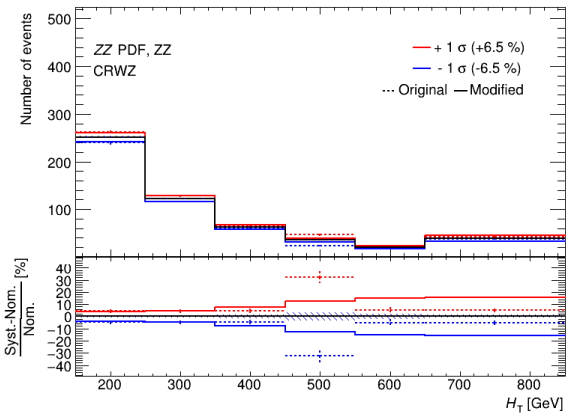


(e) $t\bar{t}Z$ μ_F uncertainty in CRWZ.

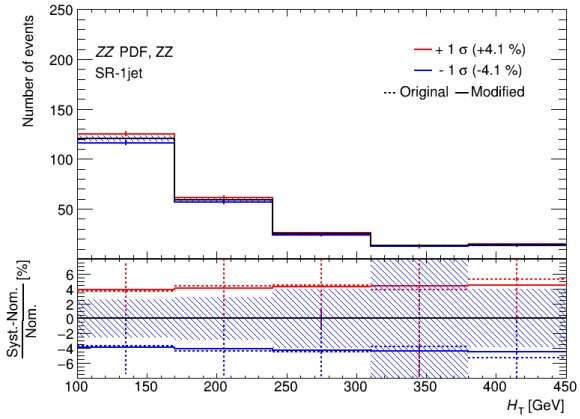


(f) $t\bar{t}Z$ μ_F uncertainty in SR-prod.

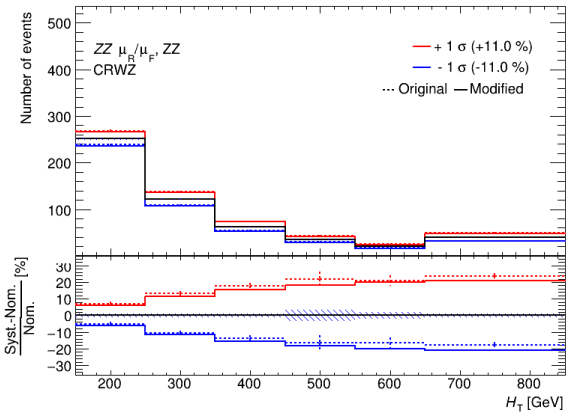
Figure B.2: Plots, in CRWZ (left) and SR-production (right) regions, where $t\bar{t}Z$ is mostly concentrated, showing the effect of $\pm 1\sigma$ variations of (a,b) ISR, (c,d) PS, (e,f) factorization scale modelling nuisance parameters on $t\bar{t}Z$, before and after the application of smoothing and symmetrisation by TRExFitter. The continuous (original) and dashed (modified) lines refer to the up/down (red/blue) variations. The black histogram line, in the upper panel, corresponds to the nominal sample. The symmetrisation used for the PS systematic is `ONESIDED`, i.e. the "down" variation is added as a mirror image of the "up" variation. Hence the modified (blue) line is not accompanied by a dashed line. The hatched area represents the statistical uncertainty (divided by the MC prediction in the ratio pad).



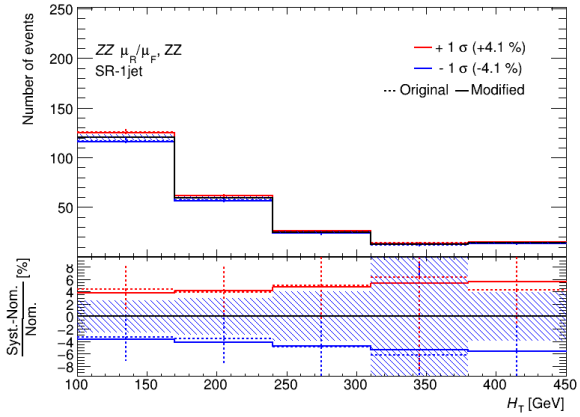
(a) ZZ PDF set uncertainty in CRWZ.



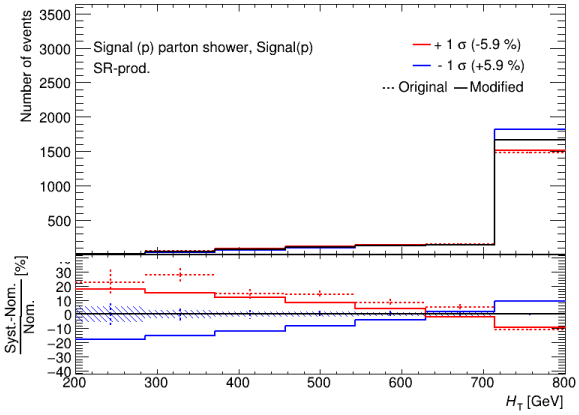
(b) ZZ PDF set uncertainty in SR-1jet.



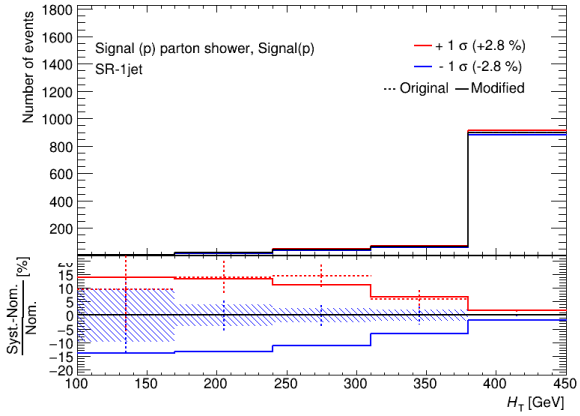
(c) ZZ scale uncertainty in CRWZ.



(d) ZZ scale uncertainty in SR-1jet.



(e) Production-signal PS uncertainty in SR-prod.



(f) Production-signal PS uncertainty in SR-1jet.

Figure B.3: ZZ PDF set and μ_R/μ_F (scale) modelling systematics, and production-signal PS uncertainty, in the regions most populated by each process, before and after the application of smoothing and symmetrisation by TRExFitter. The continuous (original) and dashed (modified) lines refer to the up/down (red/blue) variations. The black histogram line, in the upper panel, corresponds to the nominal sample. The symmetrisation used for the signal PS systematic is `ONESIDED`, i.e. the "down" variation is added as a mirror image of the "up" variation. Hence the modified (blue) line is not accompanied by a dashed line. The hatched area represents the statistical uncertainty (divided by the MC prediction in the ratio pad).

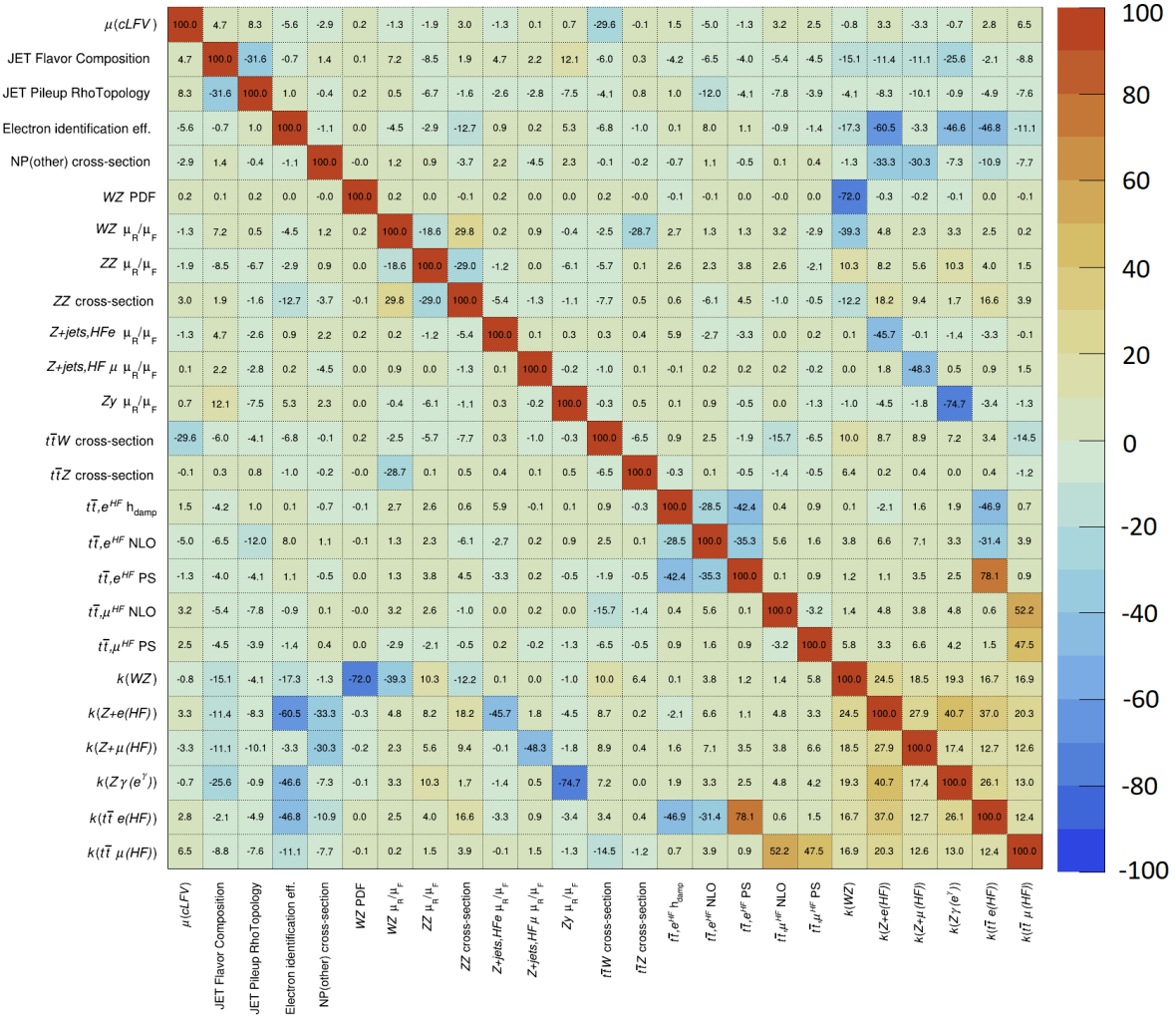
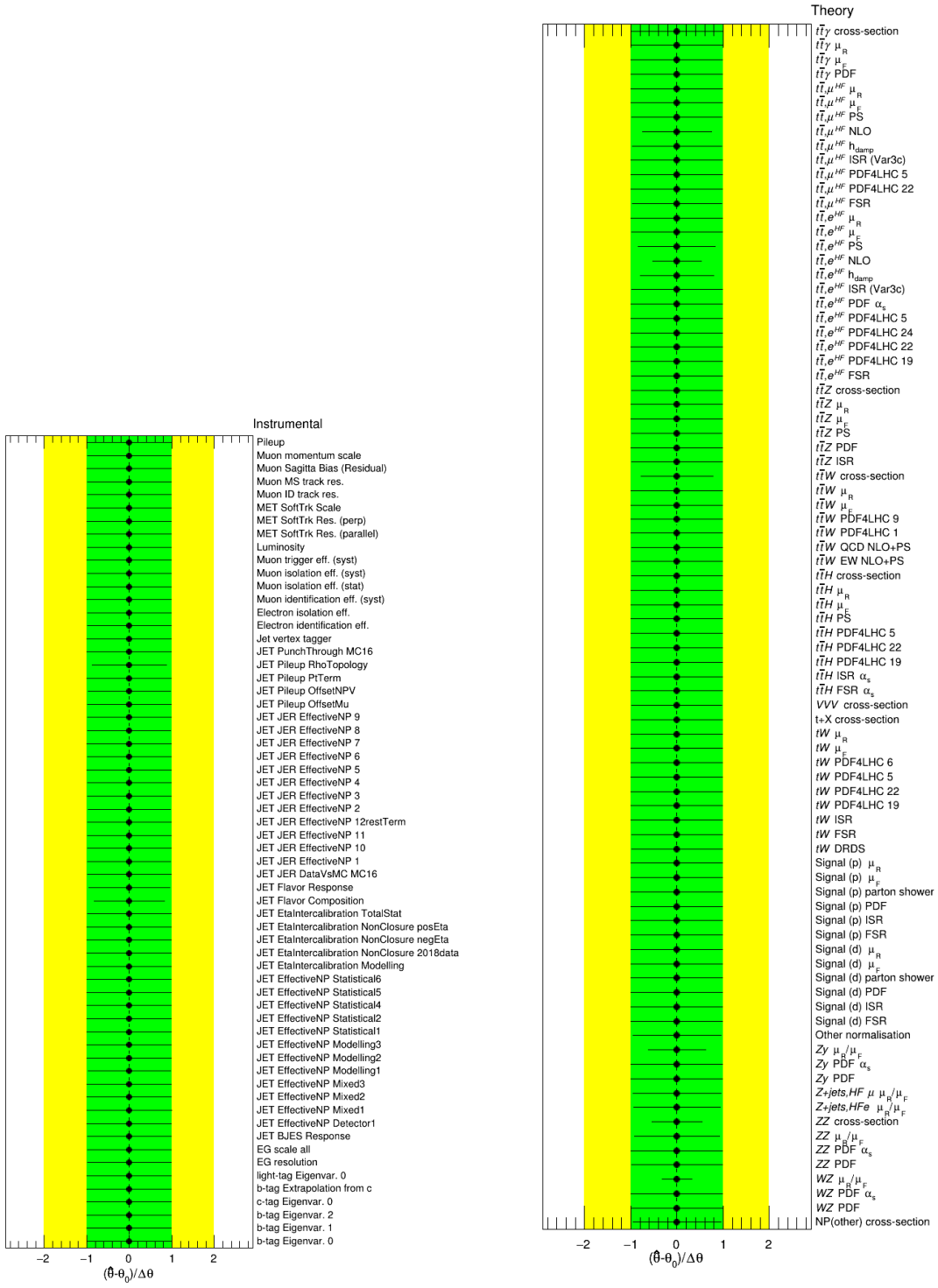


Figure B.4: Correlation matrix (in %) of the NPs obtained from the fit using Asimov data in all regions. A threshold of 25% has been set, so that NPs with lower correlation values don't appear, in order to make the plot more visible.



(a) (b)

Figure B.5: Nuisance parameter pull plot obtained from the fit using Asimov data in all regions, showing how the values of the nuisance parameters and their uncertainties changed relative to those before the fit. To improve visibility, the $t\bar{t}W$ PDF set NPs, for which no constraint or pull is observed, are omitted.

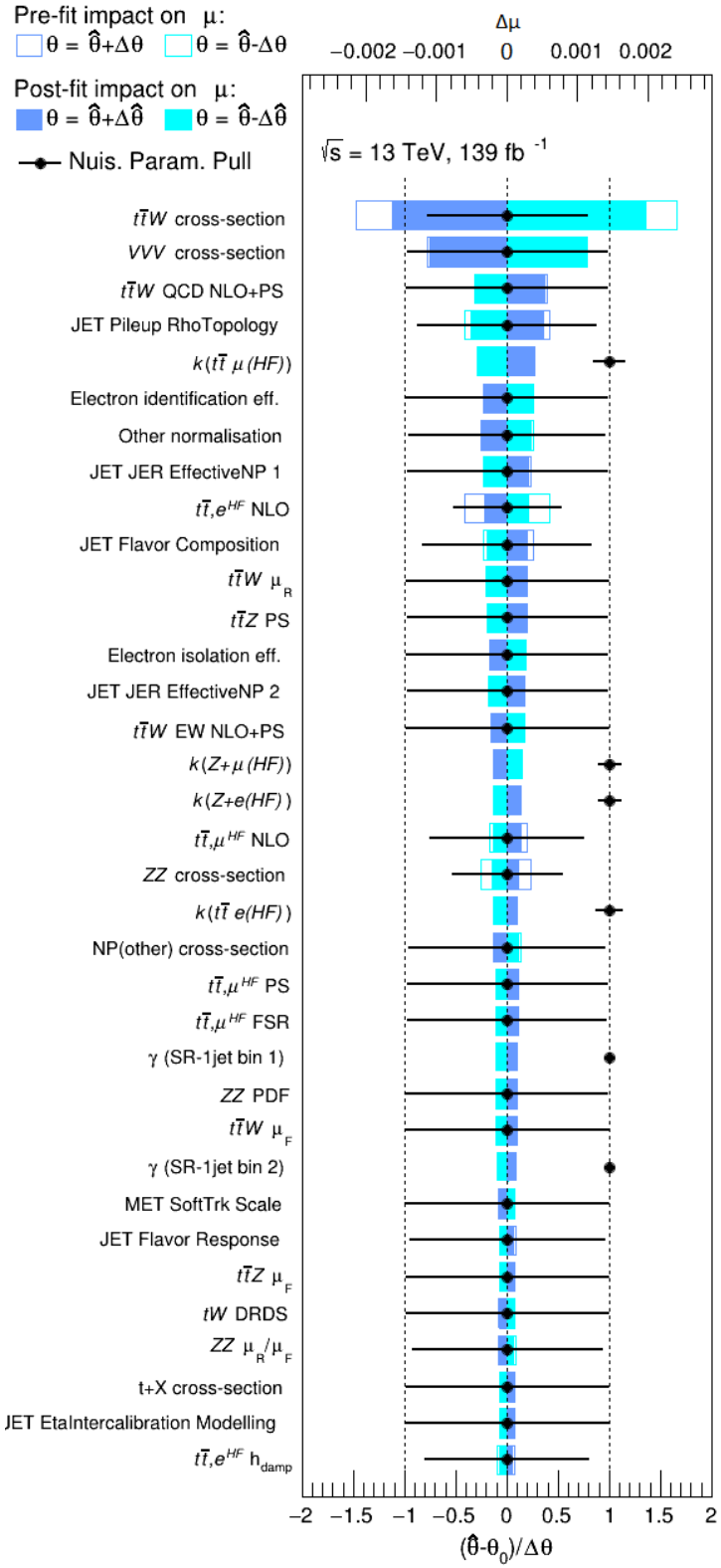


Figure B.6: Nuisance parameter ranking plot for the fully Asimov fit. The information in the caption of fig. 5.42 applies.