SEARCH FOR THE STANDARD MODEL HIGGS BOSON PRODUCED IN ASSOCIATION WITH TOP QUARK PAIRS IN MULTI-LEPTONIC FINAL STATES WITH THE ATLAS DETECTOR AT THE LHC

MARCO MILESI

orcid.org/0000-0002-8805-1886

School of Physics

The University of Melbourne

Dipartimento di Fisica

Università degli Studi di Milano

Prof. Elisabetta Barberio
Prof. Attilio Andreazza

June, 2018

Submitted in total fulfilment of the requirements
of the degree of Doctor of Philosophy

Produced on archival quality paper

ABSTRACT

The search for the Standard Model (SM) Higgs boson produced in association with top quarks - known as $t\bar{t}H$ production - plays a crucial role in the Large Hadron Collider (LHC) physics programme, as it allows a direct measurement of the Higgs field Yukawa coupling to the heaviest fermion and can constrain effects of new physics beyond the Standard Model in the top coupling sector.

This thesis presents a search for the $t\bar{t}H$ production in an inclusive multileptonic final state, with a proton-proton collision dataset corresponding to an integrated luminosity of $\int L\ dt = 36.1\ \text{fb}^{-1}$, collected by the ATLAS experiment at the LHC in 2015 and 2016 at a centre-of-mass energy of $\sqrt{s} = 13$ TeV.

The final state is characterised by high jet multiplicity, and the presence of several electrons and muons, as well as hadronically decaying tau leptons. The multiplicity of these physics objects allows the definition of several categories to enhance the sensitivity of the analysis.

The particular focus of my work lies on the final state where exactly two light leptons with the same electric charge and no hadronic taus are found - indicated as 2ℓ SS $0\tau_{had}$ - for which I developed a novel technique to estimate the reducible background of non-prompt (fake) electrons and muons.

Boosted decision tree algorithms are trained to discriminate the $t\bar{t}H$ signal events from the two major background processes in this channel: $t\bar{t}V$ (V=W,Z) and events with fake leptons. A fit of our model to the observed data is performed, and the results are interpreted using a frequentist approach. A best-fit value for the strength of the $t\bar{t}H$ production cross section with respect to the Standard Model expectation of $\mu_{t\bar{t}H} = 1.5^{+0.7}_{-0.6}$ is observed.

The observed sensitivity of this search corresponds to a 2.7σ excess of events above the SM background-only hypothesis, with an expected median sensitivity of 1.9σ for a model where the SM $t\bar{t}H$ production is assumed.

Combination with the other categories of the $t\bar{t}H$ to multi-leptons analysis eventually leads to a signal strength of $\mu_{t\bar{t}H} = 1.6^{+0.5}_{-0.4}$, with an observed (expected) sensitivity of 4.1σ (2.8σ) above the SM background-only hypothesis. This indicates the strongest evidence to date for the $t\bar{t}H$ production mode.

Furthermore, I present a study on improvements to the ATLAS track reconstruction algorithm to enhance its performance in environments with

high density of tracks, such as the core of boosted hadronic jets and hadronically decaying tau leptons.

DECLARATION

This is to certify that:

1. this thesis comprises only my original work towards the PhD except where indicated in the Preface;
2. due acknowledgement has been made in the text to all other material used;
3. the thesis is less than 100,000 words in length, exclusive of tables, bibliographies and appendices.

Parkville, June, 2018

Marco Milesi

PREFACE

In recent years, accomplishments in experimental particle physics have been the result of increasingly large international scientific collaborative efforts. This is especially the case for experiments at the Large Hadron Collider, like ATLAS. As a result of these collaborative undertakings, it is incredibly rare that any substantial body of work is done in isolation. The standard procedure is to work in groups, each carrying out a specific purpose within the wider collaboration such as data acquisition, detector monitoring and development, particle reconstruction and event simulation. All these activities represent essential inputs to any scientific achievement in ATLAS. This is the reason why all the members of the collaboration are listed as authors on each ATLAS publication.

The results presented in this thesis have been developed in concert with different working groups and includes some external contributions, each detailed here.

Chapter 1 includes information sourced from several publications - mostly by the ATLAS and CMS collaborations - and is meant to provide a motivation to the personal work of the author. Chapter 2 is an original summary of the ATLAS experiment, with extracts from a number of public results about the ATLAS detector design and performance.

The author's original work is described in Chapter 3, 4, 5. The work presented in Chapter 3 was undertaken by the author as part of the ATLAS Tracking CP working group, in collaboration with Dr. G. Facini and Dr. A. Morley and supervised by Prof. A. Andreazza. Chapter 4 and 5 illustrate the work of the author conducted within the ATLAS HTop (Higgs-top coupling) working group, under the supervision of Prof. E. Barberio and Dr. D. Zanzi. Unless specifically stated, all the information provided in these two chapters represents original work of the author, although results outlined in section 4.6 were carried out in close collaboration with Dr. P. Pralavorio, Dr. F. Hubaut and K. De Vasconcelos. Result presented in subsection 4.4.5 owe to studies performed by R. Roberts, K. Mochizuki and G. Tarna. Studies described in subsection 4.8.1 were peformed by A. Chomont.

ACKNOWLEDGEMENTS

Ah, you know, lad, that love business is a powerful thing. — Greater than gravity? — Well, yes, boy. In its way, I'd, uh... Yes, I'd say it's the greatest force on earth. (Merlin)

At the conclusion of some intense, yet immensely rewarding years, I wish to express my heartfelt thanks to all those who made some special contribution to achieving this great goal.

First and foremost, I am extremely grateful to my supervisor Elisabetta Barberio for giving me the chance to join the outstanding particle physics community at the University of Melbourne in the CoEPP group, and for opening me the doors to the wonders of Australia. Her strong leadership and motivation as a scientist have been highly inspirational to me. I also wish to greatly thank Elisabetta for her trust, and for her care through all the hardships I had to face in this quest. I owe the fulfilment of my PhD also to the experienced mentorship of my second supervisor Attilio Andreazza, and I thank him for having encouraged the continuation of my research after my master studies in Milan.

My deepest gratitude goes out to Daniele Zanzi. I am very much grateful to him for constantly pushing me to reach the best, never settling for shallow explanations. In spite of his young age, he surely is a great scientist already, and I wish him a brilliant career ahead.

Special thanks go to Geoff Taylor and Ray Volkas for guiding CoEPP and the School of Physics to great success during the past years. I feel blessed for having shared so much knowledge and fun with such an amazing team of people. I wish also to thank Lucien Boland and Sean Crosby for all their invaluable help on the numerous technical and software-related issues I had to deal with. A honourable mention to Lucien for being a top-notch basketball coach (not to mention an ultra-competitive player).

Working in a large collaboration as ATLAS definitely is a challenge, mostly due to the sometimes overwhelming pressure. Luckily, I had the chance to meet several exceptional people along the way who patiently helped me out gaining deeper insight into physics problems, and taught me a lot to improve as a scientist. To this end, I wish to mention in particular Gabriel Facini, Pascal Pralavorio, Fabrice Hubaut and Ximo Poveda. I am also genuinely thankful to all the people with whom I enjoyed my time at CERN in these

years, most notably Andrea Favareto, Simone Mazza, Lidia Dell'Asta and Marco Rimoldi.

Certainly, I cannot forget about all my fellow PhD candidates. It has been a true pleasure to share my work days and trips with some awesome people out there. I am particularly grateful to Amelia Brennan, Giacomo Caria, John Gargalionis, Chia-Ling Hsu, David Jennens, Brian Le, Pere Rados, Laurence Spiller (a huge thanks for proofreading the whole manuscript!) and Eiasha Waheed for their advice and consideration. If then I had to single someone out, that would be Francesco Nuti. His knowledgeable help has been precious to me, especially when ramping up on the $t\bar{t}H$ business. Too bad he is a Juventus supporter, but I will turn a blind eye on that for this time.

During these years, I have also been fortunate to come across some greatly inspirational, affable and fun-loving colleagues here in Melbourne (yes, they are all physicist. Well, not quite *all* of them...): the "Baroncelli Bros" Tommaso and Tiziano, Noel (Nello) Dawe, Sara Diglio, Federico Scutti, Francesco Tenchini, Francesca Ungaro, Matteo Volpi. Thanks guys for all the good times and shenanigans, you are truly friends of mine.

Accomplishing this milestone would never have been possible without the affection and concrete aid of my family: Maurizio, Matilde, Giulia and Mattia, uncle Arnaldo and nonna Edda. No matter the distance that separates us, my heart and thoughts are always with you. If I have become the man I am, I owe it above all to you. May the Lord bless you, always. And thank you, Martina, for your courage, strenuous support, and deep love. My journey towards the PhD coincided with our life journey together, made of hard work and sacrifice, but ultimately rewarded by immense joy and happiness. My achievement is yours, too.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

INTRODUCTION

A new particle consistent with the Standard Model (SM) Higgs boson was discovered in 2012 by the ATLAS and CMS collaborations at the Large Hadron Collider (LHC). Ever since then, precise measurement of its properties has been of paramount importance in fundamental physics.

Given the large mass of the top quark of $m_t \approx 173$ GeV when compared to the other fermions, the Higgs boson Yukawa coupling to the top is by far the strongest among all fermion couplings. This feature makes it of particular interest, as it can shed a light on the scale of new physics beyond the Standard Model.

Despite the challenge of a very low production cross section compared to other Higgs boson production modes accessible at the LHC, the $t\bar{t}H$ mechanism provides the cleanest probe for measuring the Higgs-top coupling, for it does not require loop diagrams to describe the interaction at the leading order in perturbation theory. This significantly reduces the dependence of the coupling measurement procedure on specific model hypotheses, and disentangles it from any potential effect due to non-SM particles.

At the ATLAS experiment, searches for $t\bar{t}H$ production target several Higgs boson decay signatures. This thesis describes the search for $t\bar{t}H$ to multi-leptons - which is sensitive to the $H \to WW^*$, $H \to \tau^+\tau^-$, $H \to ZZ^*$ decays - with a proton-proton collision dataset corresponding to an integrated luminosity of $\int L\, dt = 36.1$ fb^{-1}, collected by the ATLAS experiment at the LHC in 2015 and 2016 at a centre-of-mass energy of $\sqrt{s} = 13$ TeV (LHC Run 2). My work specifically concentrates on the final state with exactly two light leptons of same electric charge, and no hadronically decaying tau leptons (2ℓ SS $0\tau_{had}$). An outline of the thesis follows:

CHAPTER 1 contains a concise theoretical overview of the Spontaneous Symmetry Breaking (SSB) mechanism in the Standard Model, which justifies the mass degree of freedom of the W, Z weak interaction mediators and of fermions through Yukawa couplings. It provides a summary of the Higgs boson searches at collider experiments and a review of the most recent measurements of its properties, focusing on the top quark coupling measurements.

CHAPTER 2 gives a broad description of the LHC and of the ATLAS detector, providing a description of the particle reconstruction and identification algorithms employed in ATLAS, as well as their performance.

CHAPTER 3 presents an overhaul to the ATLAS tracking software algorithm for the LHC Run 2 data taking to improve the reconstruction efficiency of close-by tracks in the core of boosted jets and hadronic taus, and its implications on the performance of b-tagging algorithms to identify jets from heavy flavour quark decays.

CHAPTER 4 illustrates the $t\bar{t}H$ to multi-leptons analysis strategy, concentrating on the 2ℓ SS $0\tau_{had}$ final state. It details the trigger, event and physics object selection, and describes the set of boosted decision tree (BDT) discriminants developed to separate the $t\bar{t}H$ process from the major backgrounds. The chapter then describes in detail how the irreducible and reducible fake lepton backgrounds of this analaysis are estimated and validated. Focus is made on the data-driven technique developed to predict the fake leptons background known as the *matrix method*. Finally, it ends with a full review of the sources of systematic uncertainty for the analysis.

CHAPTER 5 introduces the framework used for the statistical analysis. Then, it moves on to the description of the modelling of the input variables to the BDT discriminants and of their output. Finally, it presents a summary of the results.

1 THE HIGGS BOSON IN THE STANDARD MODEL

1.1 THE STANDARD MODEL OF PARTICLE INTERACTIONS

The Standard Model (SM) of particle interactions describes the structure of ordinary matter and the fundamental interactions of nature, with the exception of gravity. Despite being a result of an almost one century long effort carried out by several theoretical and experimental physicists, the current formulation of the Standard Model as a renormalisable quantum field theory has been set only in the early 1970's, most notably from the work of Glashow, Weinberg, Salam, Gross, Wilczek, t'Hooft and Veltman [1, 2, 3, 4, 5].

A very important feature of the Standard Model is the large number of observable properties and processes that can be calculated and compared with experimental measurements. The Standard Model has been established through the years as an outstandingly well-tested physics theory [6]. Over the last decades, numerous experiments have checked the predictions of the Standard Model, finding an impressive agreement with the experimental measurements [7, 8]. In spite of increasingly accurate predictions and more precise experimental results, no unambiguous deviation between the Standard Model and measurements has been observed thus far. Yet experimental evidence suggests the existence of new physics beyond the SM, in particular for what concerns the neutrino sector [9, 10, 11], the nature of dark matter and dark energy [12], and the asymmetry between ordinary matter and antimatter [13]. All this motivates the ongoing efforts in searching for new physics, and the improvement in the precision of measurements and theoretical predictions.

1.1.1 Elementary particles and their interactions

In the SM [14, 15], all the known matter is ultimately made up of twelve spin-$^1/_2$ elementary particles known as *fermions*, divided into *quarks* and *leptons*, and grouped in three families, or *generations*, which exhibit similar physical behaviour. Masses of the fermions span over a broad range from near zero mass of neutrinos, to about 173 GeV[1] of the heaviest fermion, the top quark. Apart from neutrinos, all fermions are electrically charged; leptons have unit charge, whereas quarks are characterised by fractional charges. Each fermion has an associated antiparticle, whose spin and electric charge quantum numbers are reversed in sign with respect to the particle state. The electromagnetic, strong and weak interactions between particles are driven via exchange of fundamental *gauge bosons*, arising from the requirement of the underlying local gauge symmetries of the Lagrangian which describes the theory. A schematic representation of the SM particles and their quantum numbers is given in Figure 1.

In the group theory formalism, the gauge symmetries of the Standard Model are described by the transformations of the group $SU(3)_C \otimes SU(2)_L \otimes U(1)_Y$.

Quantum chromodynamics

Quarks in the $SU(3)_C$ symmetry group have a fundamental representation as *triplets* of colour - hence the "C" in the subscript - meaning that the members of each triplet differ only by their colour charge (r, g, b). The eight generators of $SU(3)_c$ - required by gauge invariance of the Lagrangian - correspond to *gluons*, spin-1 fields which mediate the interaction between coloured particles by carrying colour charge. From the gauge invariance principle, it follows that gluons have zero mass. Experimental results currently limit the rest mass of the gluon to less than few MeV [16], which support the assumption of $SU(3)_C$ being an unbroken symmetry of nature. Leptons are singlets of $SU(3)_C$: they do not carry a colour charge, hence do not interact via the strong force.

Within the SM, the theory of the strong interaction between quarks and gluons is known as quantum chromodynamics (QCD). A complete, formal discussion of QCD, and in particular of the concepts of asymptotic freedom and confinement, is outside the scope of this thesis, for which the reader may be referred to standard textbooks [14, 17]. One of its most striking successes is the possibility to interpret interactions of hadrons observed at high energy

1 In the rest of the discussion we will always consider natural units, where $c = \hbar = 1$.

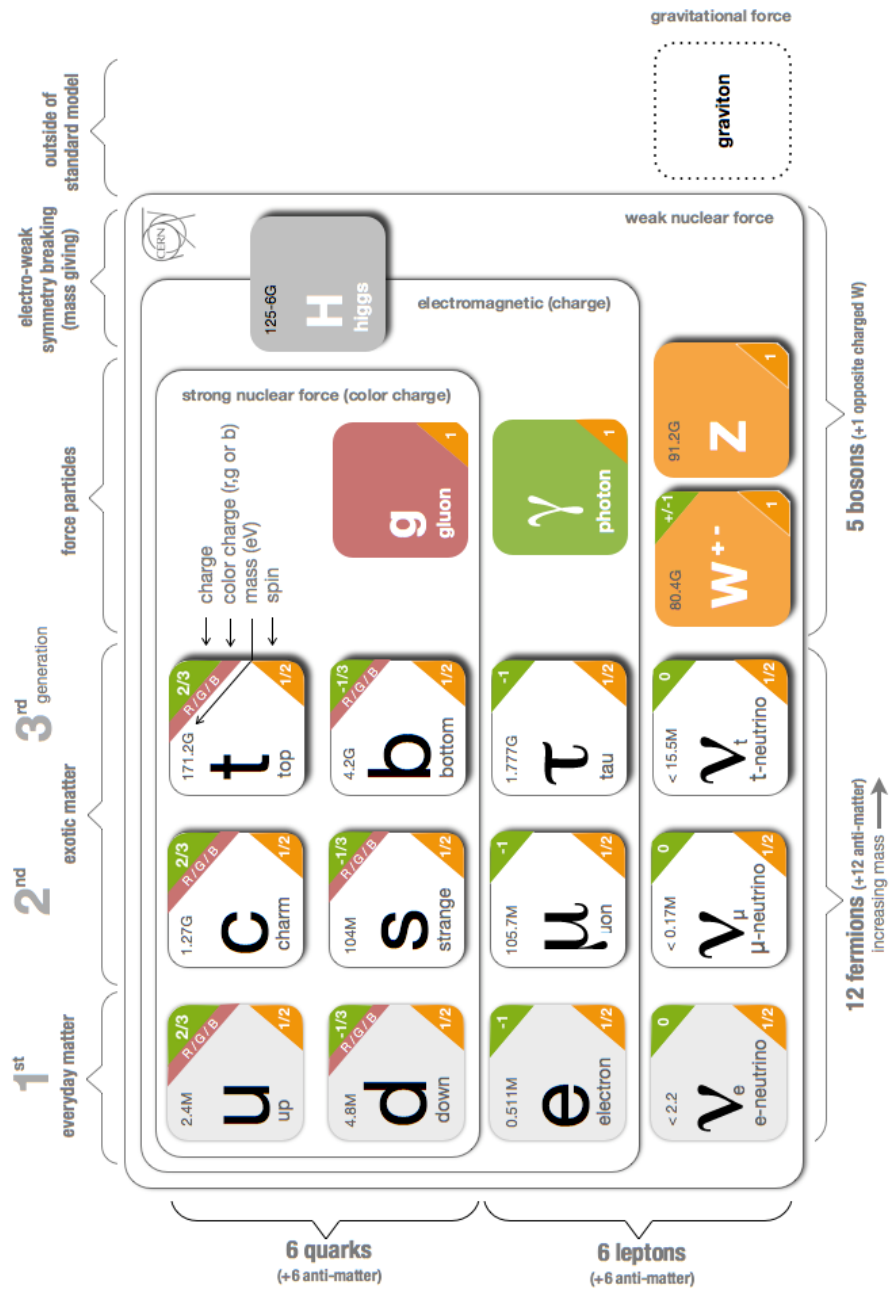


Figure 1: Elementary particles in the Standard Model. Image courtesy of CERN.

particle colliders - such as the Large Hadron Collider (LHC) [18] - in terms of elementary processes involving quarks and gluons, collectively indicated as *partons*.

Partons represent the elementary constituents of the protons. At the LHC, elementary particles are created as a result of the head-on hard scattering of pairs of partons in proton-proton collisions. The fraction of proton momentum carried by any of these constituents is unknown *a priori*, which prevents one from exploiting momentum conservation of the hard scattering system along the longitudinal direction. However, the probability densities of finding any parton with a certain longitudinal hadron momentum fraction x at some energy scale - referred to as *parton distribution functions* (PDFs) $f(x)$ - are known from measurements of deep-inelastic-scattering [19], inclusive single jet production [20] or electroweak W, Z production at hadron colliders [21, 22]. Eventually PDFs can be extrapolated to the LHC hard scattering energy scales by means of perturbative QCD evolution equations [17]. A set of PDFs as obtained by the NNPDF collaboration [23] is illustrated in Figure 2.

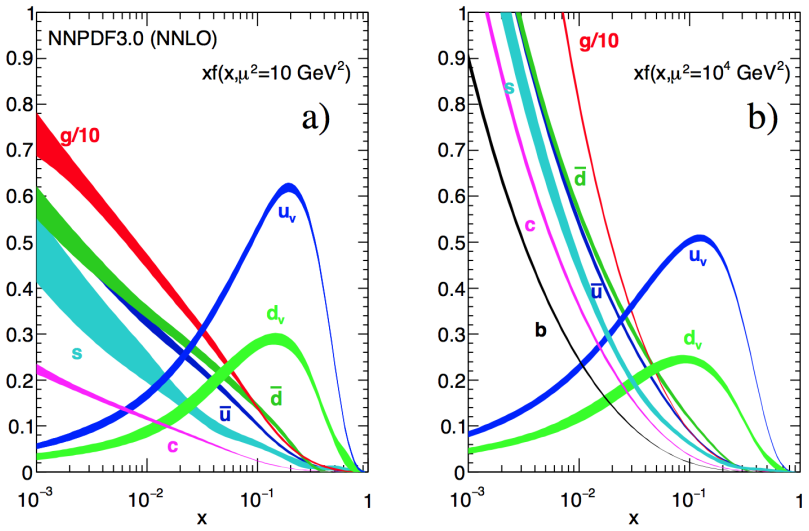


Figure 2: Proton parton distribution functions $xf(x)$ at energy scales of $\mu^2 = 10$ GeV^2 and $\mu^2 = 10^4$ GeV^2 from the NNPDF3.0 global analysis [24]. Widths of the curves represent PDF uncertainties.

The PDFs are a fundamental input to the algorithms simulating hard scattering processes for incoming hadrons a, b with centre of mass energy $\sqrt{s}$, since they enter in the definition of the *parton-parton luminosity*:

$$\frac{1}{s}\frac{d\mathcal{L}_{ij}}{d\tau} = \frac{1}{1+\delta_{ij}}\frac{1}{s}\int_\tau^1 \frac{dx}{x}\left(f_i^{(a)}(x,\mu_F^2)f_j^{(b)}(\tau/x,\mu_F^2) + f_j^{(a)}(x,\mu_F^2)f_i^{(b)}(\tau/x,\mu_F^2)\right),$$
$$\tau = \hat{s}/s = x_i x_j, \tag{1}$$

where $f_i^{(a)}$ is the PDF for the parton of species i carrying a momentum fraction x of hadron a, and $\sqrt{\hat{s}}$ is the centre-of-mass energy of the (i,j) parton system. The parton-parton luminosity in turn determines the inclusive cross section for $a + b \to \ \alpha + X$ by means of the QCD factorisation theorem:

$$\sigma(s) = \sum_{ij}\int_\tau^1 d\tau \frac{d\mathcal{L}_{ij}}{d\tau}\int \sum_{n=0}^{\infty} \alpha_s^{(n)}(\mu_R^2)\ d\hat{\sigma}_{ij\to\alpha}^{(n)}(\hat{s},\mu_F^2,\mu_R^2), \tag{2}$$

where $d\hat{\sigma}_{ij\to\alpha}^{(n)}(\hat{s},\mu_F^2,\mu_R^2)$ is the differential parton-level cross section in an element of phase space of the final state α, calculated at the n-th order in perturbative QCD, and the external sum is over the types of partonic constituents pairs. The quantity α_s represents the strong coupling constant.

In a nutshell, in the above formula the PDFs absorb the unknown information about the remaining parts of the two protons which interact "softly", that is, in the non perturbative range of QCD. The PDFs depend on a particular scale μ_F^2 - known as the *factorisation scale* - which qualitatively represents the resolution with which the hadron is being probed. This is an unphysical parameter, which appears also in the parton-level cross section. In the latter, there is an additional dependence on an unphysical energy scale - the *renormalisation scale* μ_R^2 - which also affects α_s and arises from the regularisation of ultraviolet divergences.

Even so, the physical cross section result calculated to all orders in perturbation theory is formally independent of the choices of μ_F^2 and μ_R^2. In fact, the dependence of the coefficients of the perturbative expansion of $\hat{\sigma}_{ij\to\alpha}(\hat{s},\mu_F^2,\mu_R^2)$ on these parameters is compensated exactly by the scale dependence of the parton distribution functions and of the strong coupling constant. The compensation becomes more exact as more terms are included in the perturbation series. In case of a limited set of higher-order corrections, it is necessary to make a specific choice for μ_F^2 and μ_R^2 in the calculation of the cross section prediction. A common one is to assign values of the order of the typical four-momentum transfer of the hard-scattering process under examination. Making this unphysical scales fluctuate in the calculations, it is possible to estimate the impact of the theoretical uncertainty. Therefore, the lower the order in perturbation theory to which the parton-level cross sec-

tion calculation is carried out, the larger the theoretical uncertainties arising from the fixed value for μ_F^2, μ_R^2 will be.

Electroweak unification

The weak and electromagnetic interactions are considered as part of a unified $SU(2)_L \otimes U(1)_Y$ electroweak group. The generators of the transformations of this group correspond to four spin-1 fields, two of which carry electric charge, and two are neutral. The "L" subscript indicates that the $SU(2)$ symmetry group transforms only the left-handed (L) chiral components of fermions - which are represented as weak force isospin *doublets* - leaving the right-handed (R) components unchanged. The latter are thus represented as $SU(2)_L$ *singlets*. This property expresses parity violation in the weak sector of the SM. In the SM, neutrinos are originally assumed to be massless, and cannot exist in a right-handed chiral state. Electromagnetic interaction arises from the exchange of a neutral gauge boson, which results from the mixing of the neutral gauge bosons of the $SU(2)_L$ group and the $U(1)$ group. The hypercharge Y quantum number associated to the $U(1)$ group is effectively a linear combination of the electrical charge and the longitudinal component of the weak isospin.

Gauge invariance under $SU(2)_L \otimes U(1)_Y$ transformation implies electroweak force carriers to be massless. Also, fermion fields are not allowed to have a mass, as the presence of a mass term in the Lagrangian would imply mixing of left-handed and right-handed fermion states, which live in different representations of the fundamental gauge group.

However, experimental evidence of the short range of the weak interaction, and of the mass differences between fermions within weak isospin doublets indicate that the $SU(2)_L$ symmetry has to be broken. The mechanism by which this happens in the Standard Model is described in the following section.

1.1.2 Spontaneous symmetry breaking and the Higgs mechanism

In the formulation of the SM as a gauge theory, all the (local) symmetries of the Lagrangian are not necessarily also symmetries of the vacuum (ground) state of the system. In fact, the evidence for massive weak force mediators (the $W^{\pm}$ and Z bosons) and massless electromagnetic and strong force mediators (the photon γ and the gluons g) implies that the vacuum state is invariant under the $U(1)_{EM} \otimes SU(3)_C$ gauge groups, but the SM

symmetry under $SU(2)_L \otimes U_Y$ is not preserved, and it is reduced to a symmetry for $U(1)_{EM}$ only.

In quantum field theory, the *Nambu-Goldstone* theorem [25, 26] states that for every symmetry in the Lagrangian which is not a symmetry of the vacuum state there exists a massless scalar boson. That is, every broken symmetry of the vacuum should manifest as a physical particle - known as *Goldstone boson* - whose interactions are dictated by the gauge symmetry of the Lagrangian itself. The $SU(2)_L \otimes U_Y$ symmetry breaking would imply the existence of three Goldstone bosons. Yet the fact that no such particles are observed must imply a way of breaking the symmetry of the ground state which takes into account the "absorption" of the Goldstone boson degrees of freedom into other observable quantities. The idea of having such a mechanism in particle physics was developed independently by Higgs, Brout, Englert, Kibble, Guralnik, Hagen in the 1960's, and is known nowadays as the *Higgs mechanism* [27, 28, 29, 30].

The Higgs mechanism postulates the existence of a scalar field (the *Higgs boson*), which can be minimally represented as a weak isospin doublet, with $Y = 1$ and electric charge $Q = 0$. This results from the addition of a $SU(2)_L \otimes U_Y$ invariant potential term in the Lagrangian for the scalar field self-interaction, with a non-zero vacuum expectation value v (VEV) for the field itself:

$$\varphi_0 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix}, \tag{3}$$

which implies the field to be massive.

The value for the vacuum expectation value can be inferred from the precise measurement of the Fermi constant G_F through muon lifetime, and it is $v = (\sqrt{2}G_F)^{-1/2} \approx 246$ GeV. This sets the energy scale at which the electroweak symmetry is broken. At this scale, electromagnetic and weak forces become distinct, and the bosons mediating the interactions acquire a mass. The Higgs boson mass can be expressed as $m_H = \sqrt{2\lambda}v$, where λ is the coupling strength of the four-point Higgs self-interaction. Since the latter is unknown a priori, the Higgs mass remains a free parameter of the theory.

Expanding the Higgs field as excitations around the minimum ($\varphi = \frac{1}{\sqrt{2}} \left(\begin{smallmatrix} 0 \\ v+h \end{smallmatrix} \right)$) eventually leads to mass terms in the Lagrangian for three out of the four vector bosons of the electroweak interaction, leaving one massless. Also, the three massless bosons predicted by the Nambu-Goldstone theorem can be interpreted as being effectively absorbed as longitudinal polarisation modes

of the three massive vector bosons of the weak interaction, in accordance with the *Goldstone boson equivalence* theorem.

The Higgs boson couplings g_V to the $W^{\pm}$ and Z bosons are proportional to the square of their masses:

$$= ig_V = i2\frac{m_V^2}{v}g^{\mu\nu} \quad (V = W^{\pm}, Z), \tag{4}$$

which in turn depend on the $SU(2)_L$ and $U(1)_Y$ coupling constants g, g' and the VEV:

$$m_W = \frac{1}{2}vg \quad m_Z = \frac{v}{2}\sqrt{g^2 + g'^2}. \tag{5}$$

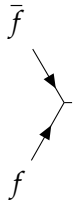
The Higgs mechanism can also account for the introduction of gauge-invariant fermion mass terms after spontaneous symmetry breaking. This is achieved by postulating a Yukawa interaction term between the Higgs doublet and the fermion fields:

$$\Delta\mathcal{L}_f = -g_f\left(\bar{f}_L\varphi f_R + h.c.\right) = -\frac{g_f v}{\sqrt{2}}\bar{f}f - \frac{g_f}{\sqrt{2}}h\bar{f}f, \tag{6}$$

where a mass-like term for fermions is identified:

$$m_f = \frac{g_f v}{\sqrt{2}}, \tag{7}$$

A linear relationship between the Higgs-fermion Yukawa coupling g_f and the fermion mass thus holds:



$$= -ig_f = -i\sqrt{2}\frac{m_f}{v}. \tag{8}$$

Like the Higgs boson mass, also the fermion masses are free parameters of the SM, as it does not provide a prediction for the fermionic Yukawa couplings.

The Higgs boson has no tree level coupling to massless vector bosons, i.e. gluons and photons. However, an effective coupling arises from virtual loops involving quarks and massive vector bosons, as later displayed in Figures 6a and 7.

1.2 HIGGS PHYSICS AT COLLIDERS

Attempts to a direct observation of the Higgs boson have represented one of the leading fields of physics research during the past 40 years. Based on theoretical grounds, an upper bound on the Higgs mass of $m_H \lesssim 1$ TeV exists to prevent violation of unitarity in the two-body $W^+W^- \to W^+W^-$ vector boson scattering process [31]. Stability of the vacuum further implies a lower mass bound of $m_H > 3.7$ GeV [32].

In the high energy collider experiments era (starting from the late 80's), experiments at the LEP collider [33] conducted direct searches for the Higgs through the $e^+e^- \to ZH$ process in the mass range up to ≈ 120 GeV, setting a lower bound of $m_H > 114.4$ GeV at 95% confidence level. Indirect experimental limits from precision measurements of the electroweak parameters such as the top mass and the W boson mass - which depend logarithmically on the Higgs boson mass through radiative corrections - were also extracted, predicting a mass of the Higgs of $m_H = 94^{+29}_{-24}$ GeV [34]. Other direct searches, conducted by the CDF and D$\varnothing$ experiments at the $p - \bar{p}$ Tevatron collider [35], lead to exclusion of the $100 < m_H < 103$ GeV and $147 < m_H < 180$ GeV ranges of mass.

The summer of 2012 eventually marked a milestone in the history of particle physics, after both the ATLAS [36] and CMS [37] experiments at the LHC announced the observation of a particle consistent with the Standard Model Higgs boson with more than 5σ significance [38, 39] at a mass of approximately 125 GeV. The observation came from the analysis of the final states of the Higgs decay to bosons, i.e. $H \to WW^* \to \ell\nu\ell\nu, H \to ZZ^* \to 4\ell, H \to \gamma\gamma$.

With more data collected by the two experiments in the following years, evidence of the Higgs decay to fermions ($b\bar{b}$, $\tau^+\tau^-$) has also been established [40, 41], and measurements of the Higgs boson properties - including its mass [42], spin and parity [43], production rates and couplings [44] - have been carried out. These measurements provide additional tests of the Standard Model predictions, and are sensitive to new physics effects. In the following paragraphs more details on the Higgs coupling measurements will be given, starting with a review of the Higgs production modes at the LHC and the accessible decay modes.

1.2.1 Higgs boson production and decays at the LHC

The Large Hadron Collider at CERN has been built with the main purpose of searching for new phenomena beyond the SM - such as new particles predicted by supersymmetry (SUSY) or other exotic models - as well as testing the predictions of the SM itself with improved precision. As for the latter, the aim was firstly to prove (or disprove) the hypothesis of the Higgs mechanism. At the beginning of its activity it operated at a centre-of-mass energies of the collision of $\sqrt{s} = 7,8$ TeV, subsequently raised to $\sqrt{s} = 13$ TeV since early 2015, reaching a peak instantaneous luminosity of $13.8 \times 10^{33}\text{cm}^{-2}\text{s}^{-1}$. More details on the LHC will be provided in the next chapter.

Depending on the value of m_H, the Higgs boson production cross section and decay branching ratios can be precisely calculated. At the LHC, there are four main channels through which the SM Higgs boson can be produced: gluon-gluon fusion (ggF), vector boson fusion (VBF), associated production with W/Z bosons (VH), and associated production with a top-antitop quark pair ($t\bar{t}H$). Their Feynnman diagrams and the most up-to-date predictions for the cross sections are summarised in Table 1. In more detail:

- Gluon-gluon fusion (ggF)
  The Higgs boson cannot couple directly to massless particles, such as gluons. Nevertheless, this coupling is allowed via a virtual fermion loop with a heavy quark dominance. Figure 3 shows the parton luminosities as a function of the mass of the two-parton system $\sqrt{\hat{s}} = m_\alpha$ at a centre-of-mass energy of the proton-proton collision of $\sqrt{s} =$ 14 TeV [47] for different parton type pairings. For a final state of the partonic interaction corresponding to a Higgs boson mass of $\sqrt{\hat{s}} \approx$ 125 GeV, the parton-parton luminosities involving gluons in the hard scattering process largely dominate over the quark-only luminosity. Because of this, the ggF production channel has the largest cross section at the LHC.

- Vector boson fusion (VBF)
  This production mode features two quarks in the initial state emitting a pair of heavy gauge bosons (W or Z), which annihilate to produce a Higgs boson. The final state is characterised by a distinct topology, with the presence of two energetic, forward-directed jets from the hadronisation of the two initial quarks. Due to the smaller q-q luminosity, and the fact that it is an electroweak process at LO, its cross section is smaller than the ggF production.

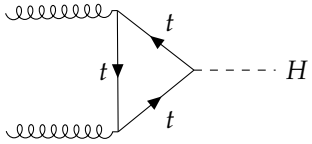
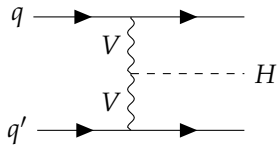
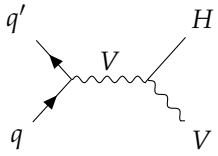
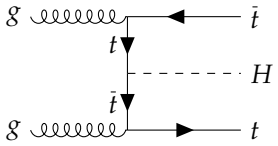
Production mode	LO diagram	σ_X [pb]	$\frac{\sigma_X}{\sigma_{pp\to H}}$	Order in perturbation theory
ggF		$46.87^{+7.9\%}_{-8.7\%}$	87%	N3LO QCD, NLO EW
VBF		$3.78^{+2.14\%}_{-2.12\%}$	7%	NNLO QCD, NLO EW
VH		$1.37^{+1.96\%}_{-2.03\%}$ $(pp \to WH)$ $0.884^{+4.12\%}_{-3.49\%}$ $(pp \to ZH)$ $0.094^{+1.89\%}_{-1.93\%}$ $(W^+(\ell^+\nu)H)$ $0.059^{+2.04\%}_{-2.12\%}$ $(W^-(\ell^-\bar{\nu})H)$ $0.029^{+4.12\%}_{-3.49\%}$ $(Z(\ell^+\ell^-)H)$	5%	NLO QCD, NLO EW
ttH		$0.507^{+6.82\%}_{-9.88\%}$	$\lesssim 1\%$	NLO QCD, NLO EW

Table 1: Higgs production modes at the LHC and their predicted cross sections. The cross sections are calculated for $m_H = 125$ GeV, and $\sqrt{s} = 13$ TeV, at the order in QCD/Electroweak perturbation theory specified in the rightmost column (NLO: *next-to-leading order*. NNLO: *next-to-next-to-leading order*...). Uncertainty on the cross section include the contributions from renormalisation and factorisation scale dependence, and uncertainties on the PDFs and the perturbative QCD coupling constant α_s [46, 45]. The second column from the right displays the relative size of the production cross section of each process with respect to the total $pp \to H$ cross section at the LHC.

- Associated production with W/Z bosons (VH)
  The third most common production channel at the LHC is the "Higgsstrahlung" mechanism (VH), where the Higgs boson is radiated off a virtual gauge boson. These events can be characterised by one or two isolated and high transverse momentum charged leptons yielded by the decay of the vector boson in the final state.

- Associated production with a top-antitop quark pair ($t\bar{t}H$)
  Our process of interest, the $t\bar{t}H$ production mechanism, has the smallest cross section with respect to the channels previously discussed due to the heaviness of its final state, despite being initiated by a pair of gluons. A very rare tH production mode (through the $gb \to WtH$ and $qb \to tHq'$ processes) is also accessible at the LHC, with a cross section of only $\approx 14\%$ of $\sigma_{t\bar{t}H}$ due to destructive interference of diagrams in the SM. However, this suppression is heavily dependent on the relative sign of the t and W Higgs coupling, thus could be enhanced in some BSM models with non-zero pseudoscalar, CP violating Higgs couplings [48].

For a Higgs boson mass of 125 GeV, a large number of final states are accessible for study, whose branching ratios are shown in Table 2. The main decay mode at $m_H = 125$ GeV is into $b\bar{b}$ pairs, since b quarks are the heaviest particles that can be pair-produced on-shell from a Higgs boson decay. The second largest decay mode into fermions is the $\tau^+\tau^-$ mode. In spite of a similar branching ratio, decays to $c\bar{c}$ pairs are less accessible at the LHC from the experimental point of view, due to the overwhelming background from QCD dijet production. The same argument holds against the Higgs decay to gg. Decays to pairs of muons have a much smaller branching ratio, albeit providing very high dilepton mass resolution. Decays to pairs of W,Z bosons do not dominate in this Higgs mass region because they cannot be both produced on-shell. However, the large Higgs coupling to these massive particles makes such decays still accessible through production of one off-shell boson in the pair. The $H \to WW^*$ branching ratio is larger than $H \to ZZ^*$ owing to the larger accessible phase space from combinatorics of $W^\pm W^\mp$. The dibosonic $H \to Z\gamma$, $H \to \gamma\gamma$ decays have a small branching ratio, yet take advantage of the very high experimental photon energy resolution. The $H \to \gamma\gamma$ decay also profits from a large signal over background ratio, thanks to the clean experimental signature of only two energetic photons.

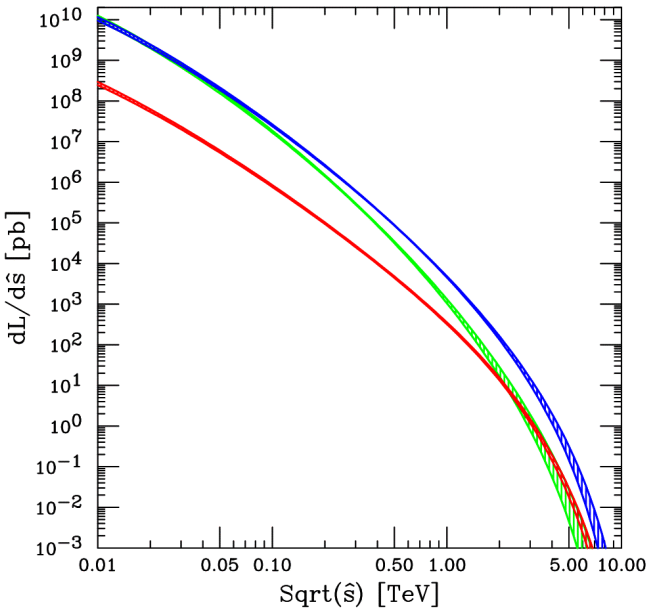


Figure 3: Parton luminosities $\frac{1}{s}\frac{d\mathcal{L}}{d\tau}$ for different parton-parton interactions at $\sqrt{s} =$ 14 TeV [47]. Widths of the curves estimate PDF uncertainties. The PDFs are taken from the CTEQ6L1 set [49]. In green, the gg luminosity; in blue the $gq + g\bar{q} + qg + \bar{q}g$ luminosity; in red, the $q\bar{q} + \bar{q}q$ luminosity.

Decay channel	Branching ratio ($m_H = 125$ GeV)
$H \to b\bar{b}$	$5.824 \times 10^{-1}{}^{+1.2\%}_{-1.3\%}$
$H \to WW^*$	$2.137 \times 10^{-1}{}^{+1.5\%}_{-1.5\%}$
$H \to \tau\tau$	$6.272 \times 10^{-2}{}^{+1.6\%}_{-1.6\%}$
$H \to ZZ^*$	$2.619 \times 10^{-2}{}^{+1.5\%}_{-1.5\%}$
$H \to \gamma\gamma$	$2.270 \times 10^{-3}{}^{+2.1\%}_{-2.1\%}$
$H \to \mu\mu$	$2.176 \times 10^{-4}{}^{+1.7\%}_{-1.7\%}$

Table 2: Higgs decay branching ratios and relative uncertainty, for $m_H = 125$ GeV [46].

1.2.2 Current results on Higgs boson properties

Analysis of Run 1 data has provided plenty of insight into the properties of the new observed particle. The ATLAS and CMS collaborations have performed a combined measurement of its mass of $m_H = 125.09 \pm 0.23$ GeV in the $ZZ^* \to 4\ell$, $\gamma\gamma$ final states [42], corresponding to an experimental precision of 0.2%. Measurement of the spin-parity of this particle - performed in the $WW^*, ZZ^*, \gamma\gamma$ final states - shows that alternative hypotheses to the SM prediction of a CP-even scalar Higgs boson ($J^P = 0^+$) are excluded at 99.9% confidence level [43].

Extensive studies of the production cross section and decay rates have been carried out combining experimental results of multiple independent analyses of both ATLAS and CMS, targeting different final states, and are documented in [44, 50, 51]. Specifically, for a given production process i and decay channel f, a fit to data is performed to determine the *signal strength* parameter:

$$\begin{aligned} \mu_i^f &= \frac{\sigma_i \times BR_f}{(\sigma_i \times BR_f)_{SM}} = \mu_i \times \mu_f, \\ \mu_i &= \frac{\sigma_i}{(\sigma_i)_{SM}}, \\ \mu_f &= \frac{BR_f}{(BR_f)_{SM}}. \end{aligned} \tag{9}$$

Since for each independent final state only the product $\mu_i \times \mu_f$ can be measured from the data, the *decay signal strength* of each analysis is the combined result of the measurements for different production modes under the assumption that all the production cross sections are the ones from the SM, that is, all μ_i's are equal to one. Similarly, the *production signal strengths* of different modes can also be fitted, under the assumption that the different Higgs decay branching fractions are the ones predicted by the SM. This approach implies that results are subject to the signal theoretical uncertainties in the inclusive cross sections and branching ratios for the various production and decay processes.

Figure 4 shows the observed production and decay signal strengths for several Higgs decays channels and production modes. A global Higgs signal strength is also defined, assuming SM values for the ratio of cross sections of different production modes, and for the ratio of branching ratios of different decay channels. Its value as obtained from the combination of the ATLAS and CMS Run 1 measurements is $\mu = 1.09^{+0.11}_{-0.10}$, and currently represents

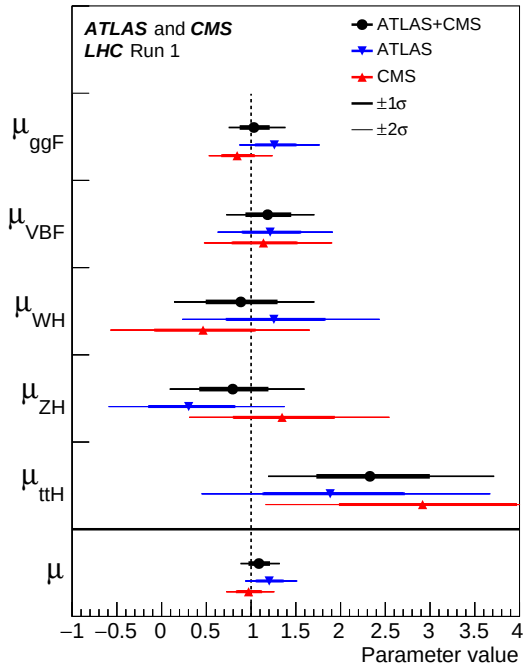
the most precise experimental measurement related to Higgs couplings. The error quoted includes the statistical uncertainty, the systematic uncertainties from background estimates and other experimental uncertainties, and the theory uncertainties on the SM Higgs cross section and branching ratios, as well as signal acceptance and kinematics/topology modelling uncertainties.

More general parametrisations are explored as well, in order to release some of the previous assumptions. One example is to consider *ratios* of cross-sections for different production modes and the *ratios* of branching fractions (equal to the ratio of partial decay widths Γ_f), normalised by the product of the ggF cross section and the $H \to ZZ^*$ width. The latter indeed represent the single measurements currently characterised by the smallest uncertainties. Results for this parametrisation are illustrated in Figure 5. This effectively establishes the existence of the SM ggF and VBF production processes with more than 5σ significance, and shows evidence of the VH mode at more than 3σ standard deviation over the background-only hypothesis.

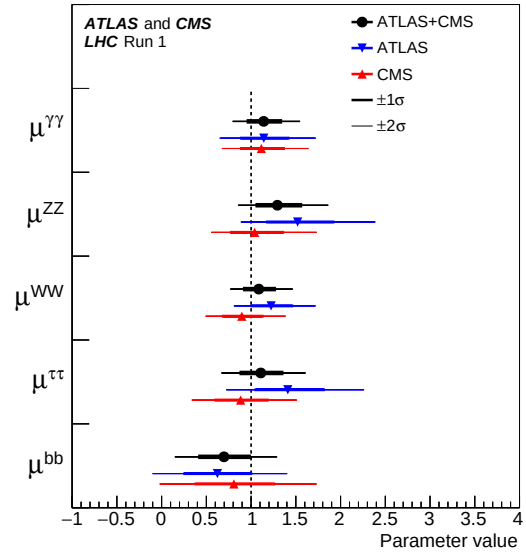
No sufficient statistical power is achieved with the Run 1 dataset to prove evidence of the SM $t\bar{t}H$ production, and this is reflected by an *expected* significance of only about 2σ. Interestingly though, the SM-normalised observed best fit value of $(\sigma_{ttH}/\sigma_{ggF}) = 3.3^{+1.0}_{-0.9}$ corresponds to an excess of approximately 3σ above the SM prediction, leading to an *observed* combined significance of 4.4σ against the background-only hypothesis. The combined global $t\bar{t}H$ signal strength under the assumption of SM branching ratios is found to be $\mu_{ttH} = 2.3^{+0.7}_{-0.6}$. Whilst all the other signal strength measurements show high compatibility of the observed new particle with the SM Higgs boson, the $t\bar{t}H$ measurement creates somewhat a tension with the SM expectation.

1.2.3 Constraints on Higgs couplings

Unlike the signal strength case, in the measurement of the Higgs couplings to bosons and fermions, production and decay modes cannot be treated independently, as each observed process involves at least two different couplings. In order to interpret the couplings as they appear in both production and decay modes in a most model-independent way, a generic framework to quantify possible deviations of the Higgs couplings from the SM is used [45]. In such framework, it is postulated that the width of the Higgs boson resonance is negligible compared to the current experimental resolution. Under this assumption, the cross section (σ_i) times branching ratio (Γ_f) for a given set of intial and final states $\sigma(i \to H \to f)$ can be factorised as:



(a) μ_i



(b) μ_f

Figure 4: 4a : best fit results for the *production* signal strengths for SM decay modes, i.e., assuming $\mu_f = 1$ in Eq. (9). The measurements of the global signal strength are also shown. 4b : best fit results for the *decay* signal strengths for SM production modes, i.e., assuming $\mu_i = 1$ in Eq. (9). Both figures are taken from [44].

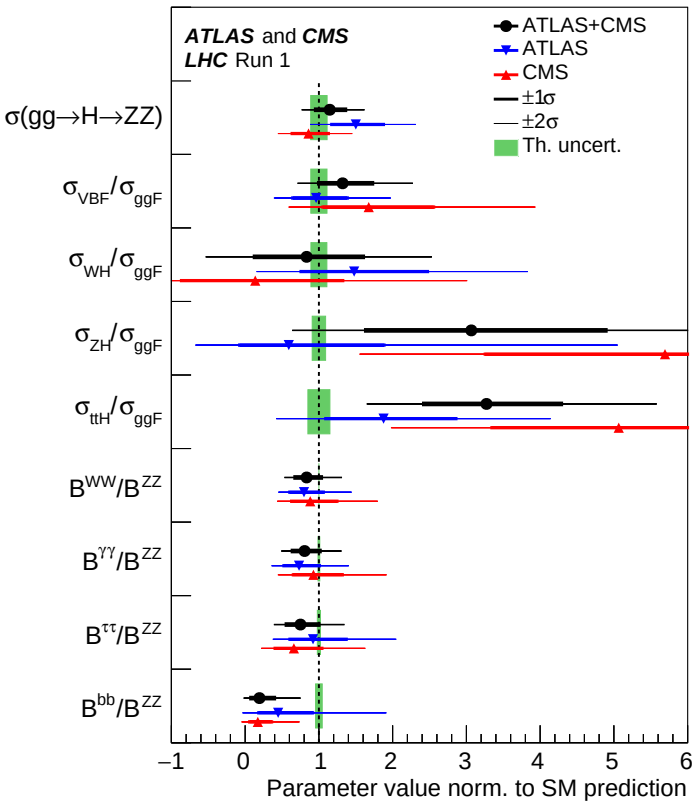


Figure 5: Best fit values of the $\sigma(ggH \to ZZ^*)$ cross section and of ratios of cross sections and branching fractions [44]. The error bars indicate the 1σ (thick lines) and 2σ (thin lines) intervals. The fit results are normalised to the SM predictions for the various parameters and the shaded bands indicate the theoretical uncertainties in these predictions.

$$
\begin{aligned}
&\sigma(i \to\ H \to\ f) = \frac{\sigma_i(\kappa_j) \cdot \Gamma_f(\kappa_j)}{\Gamma_H(\kappa_j)}, \\
&\sigma_i = \kappa_j^2 \cdot \sigma_i^{SM}, \\
&\Gamma_f = \kappa_j^2 \cdot \Gamma_f^{SM}, \\
&\Gamma_H = \kappa_H^2 \cdot \Gamma_H^{SM} \qquad \left(\kappa_H^2 = \sum_j k_j^2 \cdot \Gamma_j^{SM}\right),
\end{aligned}
\tag{10}
$$

where the κ_j factors (j refers to any particle to which the Higgs can couple) parametrise the effective leading-order coupling to all the particles which could contribute to a particular production or decay mode. Referring to Eq. (4) and (8), this means:

$$
\begin{aligned}
g_V &= \kappa_V \cdot g_V^{SM}, \\
g_f &= \kappa_f \cdot g_f^{SM}.
\end{aligned}
\tag{11}
$$

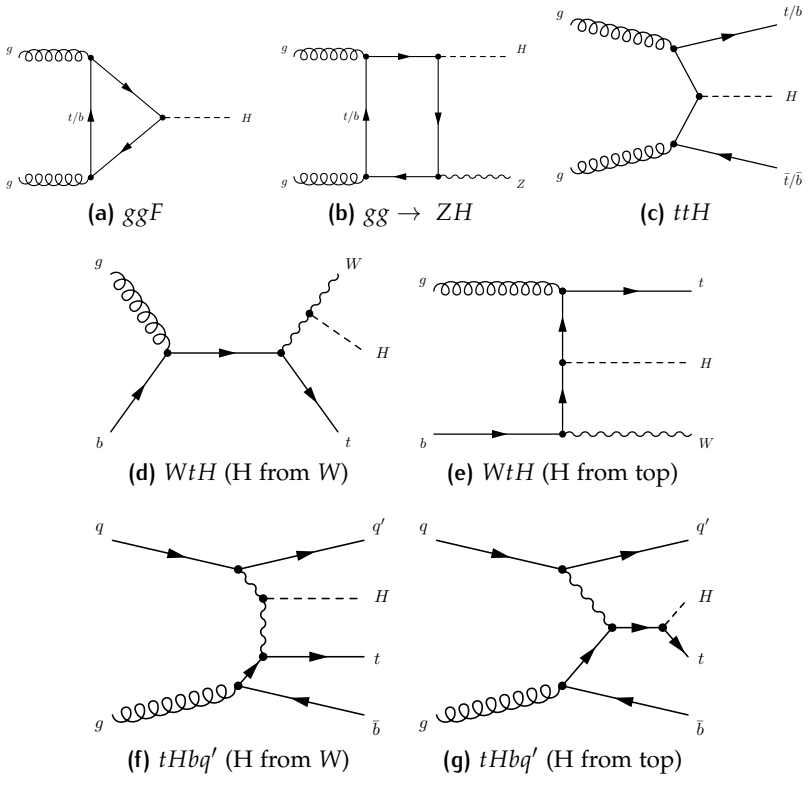
It follows that a value of $\kappa_j = 1$ implies the SM value for the couplings, hence for σ_i, Γ_f, Γ_H. The quantity Γ_H represents the Higgs boson total decay width.

Several processes are sensitive to the top coupling strength scale factor κ_t, and are highlighted in Table 3, and their LO Feynman diagrams are displayed in Figures 6,7. Notably, the tree-level $t\bar{t}H$ production can directly probe κ_t. The tH tree-level processes are suppressed in the SM due to destructive interference of the diagrams where the Higgs is radiated from a W or a top quark, as in the SM κ_W and κ_t have the same relative sign. However, if their relative sign is flipped and $\kappa_W \cdot \kappa_t$ becomes negative, the interference becomes constructive and the WtH ($tHbq'$) cross section increases by a factor 6 (13) with respect to the SM value. This would make the tH process sensitive to the relative sign between the W boson and top quark couplings.

All the previous description relies on the assumption that the Higgs boson can decay only to SM particles to determine Γ_H, and it is also assumed that only SM particles can couple to the Higgs boson in the loops, such as in the ggF or $H \to \gamma\gamma$ case. Anyway, these model assumptions can be relaxed taking into account e.g., Higgs decays to BSM invisible or undetected particles to contribute to the total decay width, and describing the loops in gg production and $\gamma\gamma$ decay mode by means of *effective coupling scale factors* (κ_g, κ_γ) which embed the possible effect of BSM couplings.

Production	Loops @ LO	Interference	κ parametrisation of σ, Γ
σ_{ggF}	✓	b-t	$\sim 1.06\kappa_t^2 + 0.01\kappa_b^2 - 0.07\kappa_t \cdot \kappa_b$
$\sigma_{gg\to ZH}$	✓	Z-t	$\sim 2.27\kappa_Z^2 + 0.37\kappa_t^2 - 1.64\kappa_t \cdot \kappa_Z$
σ_{ttH}			$\sim \kappa_t^2$
$\sigma_{gb\to WtH}$		W-t	$\sim 1.84\kappa_t^2 + 1.57\kappa_W^2 - 2.41\kappa_t \cdot \kappa_W$
$\sigma_{qb\to tHq'}$		W-t	$\sim 3.40\kappa_t^2 + 3.56\kappa_W^2 - 5.96\kappa_t \cdot \kappa_W$
Decay width			
$\Gamma_{\gamma\gamma}$	✓	W-t	$\sim 1.59\kappa_W^2 + 0.07\kappa_t^2 - 0.66\kappa_t \cdot \kappa_W$
$\Gamma_{Z\gamma}$	✓	W-t	$\sim 1.12\kappa_W^2 + 0.00035\kappa_t^2 - 0.12\kappa_t \cdot \kappa_W$

Table 3: Higgs boson production cross sections and decay widths as a function of coupling strength scale factors κ_j, for which a coupling to the top quark is relevant [50]. Expressions in the rightmost column are derived as described in [45]. For processes where multiple amplitudes contribute, the rate may depend on multiple Higgs boson coupling-strength scale factors, and interference terms may give rise to scalar product terms $\kappa_i \cdot \kappa_j$, that allow the relative sign of the coupling-strength scale factors κ_i and κ_j to be determined. The values for the coefficiencts are calculated for $\sqrt{s} = 8$ TeV and $m_H = 125.36$ GeV, and the parametrisation include higher-order QCD and EW corrections to the inclusive cross section and branching ratios.



(a) ggF **(b)** $gg \to ZH$ **(c)** ttH

(d) WtH (H from W) **(e)** WtH (H from top)

(f) $tHbq'$ (H from W) **(g)** $tHbq'$ (H from top)

Figure 6: LO Feynman diagarms for Higgs boson production modes involving coupling to the top quark.

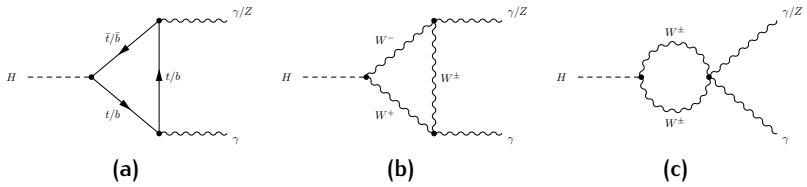


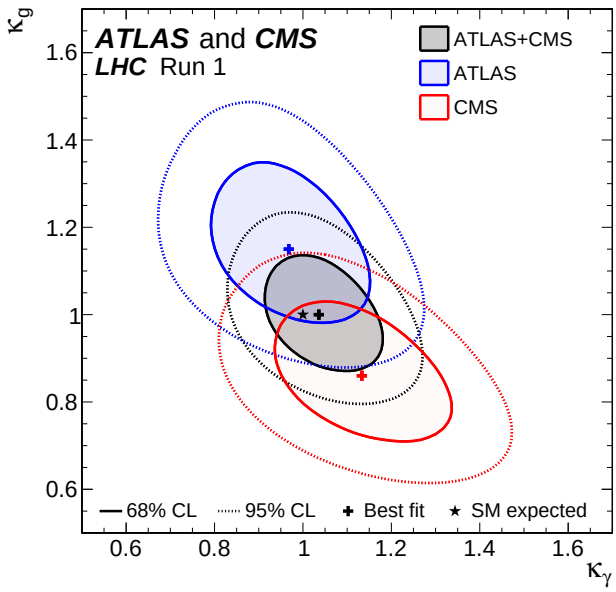
Figure 7: LO Feynman diagrams of Higgs boson decays to a pair of photons, or to a photon and a Z boson.

The results of the fit for the effective coupling modifiers, with all other coupling modifiers fixed to their SM values of unity, are shown in Figure 8a. The SM prediction of $\kappa_g = \kappa_\gamma = 1$ lies within the 68% confidence level region, and the p-value of the compatibility between the data and the SM predictions is 82%. Supported by this finding, when assuming no BSM particles in loops as well as in the Higgs decay products, the fit to the coupling modifiers produces the results shown in Figure 8b. It can be observed that all the κ's are found to be compatible with the SM expectation, and their relative sign probed by the interference terms in the parametrisation is consistent to the SM as well. In particular, the top coupling modifier is measured to be $\kappa_t = 0.87^{+0.15}_{-0.15}$ in the combination. This result is mostly constrained by the measurement of the ggF cross section and of the $H \to \gamma\gamma$ decay width. However, when allowing BSM couplings in the loops, a value of $\kappa_t = 1.40^{+0.24}_{-0.21}$ is extracted. The incompatibility of the two values thus reveals a quite strong model dependence of the interpretation, given the precision of the LHC Run 1 measurements.

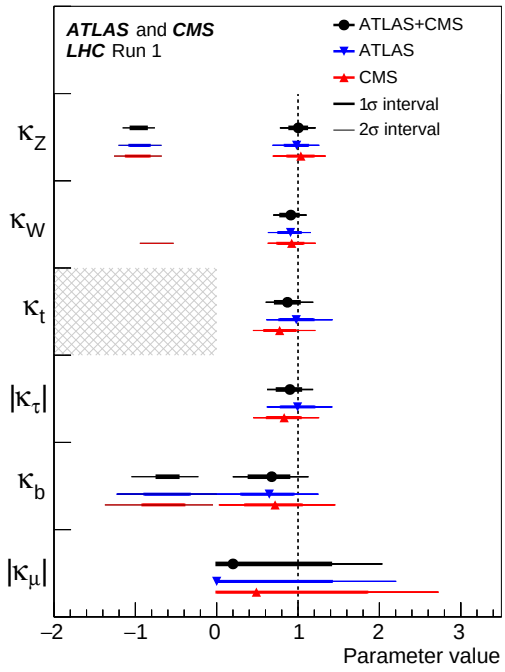
Using the coupling modifiers, new parameters c_f, c_V that are explicitly linearly dependent on the weak bosons and fermions masses are also defined, remembering Eq. (4) and (8):

$$\begin{aligned} c_f &= \kappa_f \cdot \frac{g_f}{\sqrt{2}} = \kappa_f \cdot \frac{m_f}{v}, \\ c_V &= \sqrt{\kappa_V} \cdot \sqrt{\frac{g_V}{2v}} = \sqrt{\kappa_V} \cdot \frac{m_V}{v}. \end{aligned} \tag{12}$$

Figure 9 shows the best fit values for c_f and c_V as a function of the particle mass, extracted under the hypothesis of no BSM contributions in loops. It can be seen that such linear scaling as a function of the particle masses is effectively found, indicating the compatibility of the measurements with the SM when no BSM couplings in the loops are permitted.



(a)



(b)

Figure 8: 8a: negative log-likelihood contours at 68% and 95% CL in the $(\kappa_g, \kappa_\gamma)$ plane for the combination of ATLAS and CMS and for each experiment separately, as obtained from the fit to the parameterisation allowing BSM contributions to the ggF and $H \to \gamma\gamma$ loops, while constraining all the other coupling modifiers to their SM values [44]. No BSM decay modes of the Higgs boson are considered. 8b: best fit values for the coupling parameters under the assumption of no BSM particles in the loops [44]. The κ_t parameter is assumed to be positive, without loss of generality. The κ_μ, κ_τ parameters are not affected by interference, therefore their absolute value is shown.

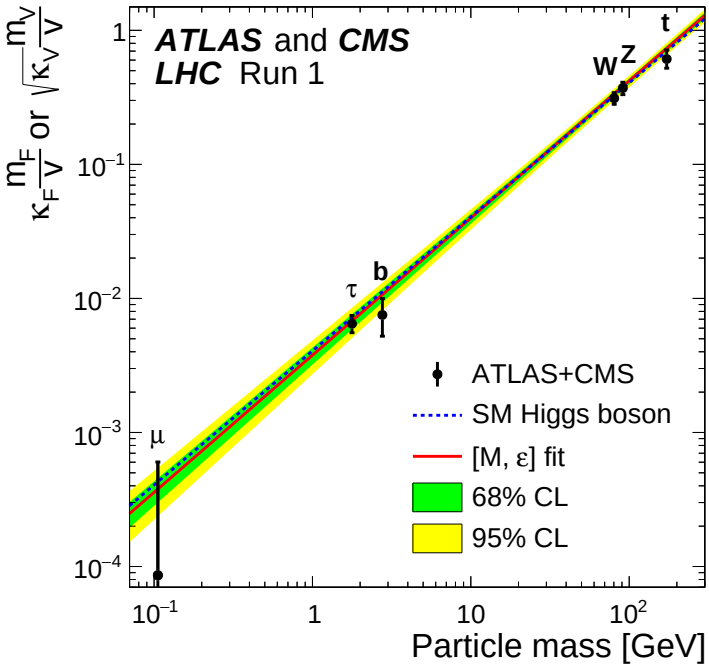


Figure 9: Best fit values for the fermion and vector boson masses parametrised via coupling modifiers κ_V, κ_f, extracted under the hypothesis of no BSM contributions in loops [44]. The dashed blue line indicates the predicted dependence on the particle mass in the case of the SM Higgs boson. The red line represents another possible expression of the coupling modifiers as a function of a mass scaling parameter, based on a phenomenological model [52], with the green and yellow bands representing the 68%, 95% CL bands.

1.3 CONCLUSIONS ON THE TOP YUKAWA COUPLING AND THE SEARCH FOR $t\bar{t}H$

As previously highlighted, parametrisations for the measurement of the Higgs couplings have to come under some assumptions, either on the contributions to the loops in production and decay modes, or on the particles to which the Higgs boson can decay. As for the coupling to the top, at the moment the strongest experimental constraint on its measurement comes from the ggF production mechanism. In the interpretation of the results, this introduces a large degree of dependence on the hypothesised physics contributions to the loop, which is naturally absent in the $t\bar{t}H$ case. Increasing the sensitivity of the $t\bar{t}H$ search is therefore crucial in order to be capable of directly measuring the top Yukawa coupling.

Furthermore, the nature of the top Yukawa coupling itself makes it of special interest among the other couplings. The heavy mass of the top quark ($m_t = 173.34 \pm 0.76$ GeV [53]) results in a coupling $\approx$ 35 times larger than the one to the second heaviest fermion, the b quark. From Eq. (7), it also appears that the top Yukawa coupling g_t is the only one "naturally" close to the scale of the EW symmetry breaking. These features have deep implications, as it has been shown they can influence both the scale of appearance of new physics, and the stability of the vacuum in our universe [54]. Ruling out any of the possible scenarios predicted by these theoretical models requires as input an extremely precise value for g_t - of order 10^{-4} accuracy - which is currently beyond our experimental sensitivity. Yet, such studies reveal how far reaching are the implications of the top Yukawa coupling.

2 THE ATLAS EXPERIMENT AT THE LARGE HADRON COLLIDER

The search for new fundamental physics at the TeV scale and the precision tests of the Standard Model require the analysis of an enormous amount of data to allow any process of interest to be distinguished from the overwhelming expected backgrounds and ensure sufficient sampling. This can be achieved by colliding particles accelerated to $\mathcal{O}(\text{TeV})$ energies with very high rate, and designing detector apparatuses with high granularity and wide spatial coverage to capture the signature of particles produced in the collisions. This chapter is dedicated to the description of the most powerful particle collider ever built so far - the Large Hadron Collider - and of the ATLAS detector, whose data have been analysed in this thesis.

2.1 THE LARGE HADRON COLLIDER (LHC)

The Large Hadron Collider (LHC) is a 27 km long storage ring built about 100 m underground at the European Organisation for Nuclear Research (CERN) facility near Geneva, Switzerland. It is hosted in the same tunnel of the previous LEP collider, and shares several parts of the already existing pre-accelerators chain. It has been designed to provide proton-proton (p-p) and heavy ions (Pb-Pb) collisions at unprecedented rate and energy, and it stems from a 30 year-long effort of a worldwide collaboration in projecting, realising and commissioning.

It is designed to reach a centre-of-mass energy of $\sqrt{s} = 14$ TeV for p-p collisions, with a nominal instantaneous luminosity of 10^{34} cm^{-2}s^{-1}. Full details about the underlying technology and performance can be found elsewhere [18, 55, 56].

The LHC started its operations in September 2008 at $\sqrt{s} = 7$ TeV, subsequently raised to 8 TeV until early 2013, which concluded the "Run 1" operations of the machine. The machine was shut down for about two years to allow for upgrades in the accelerator chain to increase the energy and rate of collisions for the "Run 2" data taking. It finally resumed operations in April, 2015, reaching a centre-of-mass energy of $\sqrt{s} = 13$ TeV. The dataset collected for this work corresponds to the years 2015 and 2016 of LHC operations.

Four main experiments have been built along the accelerator infrastructure at four interaction regions: ATLAS [36], CMS [37], LHCb [57] and ALICE [58]. ATLAS and CMS are multi-purpose detectors whose major goals are to test the Standard Model predictions in the electroweak and QCD sectors and to search for new physics beyond the Standard Model. More details about the ATLAS experiment is covered in the following section. The LHCb experiment is dedicated to precision measurements of CP violation and rare decays of b-hadrons, whereas ALICE is designed to study heavy ion collisions. In this work, however, we will concentrate exclusively on the proton acceleration chain and proton collision-related physics programme.

2.1.1 The particle acceleration chain

A schematic view of the LHC ring and the layout of the CERN accelerator complex is shown in Figure 10.

Before being injected into the LHC, protons are brought up to the injection energy of 450 GeV by a chain of accelerators. Protons are extracted from hydrogen gas, and injected in the LINAC2 linear accelerator, reaching

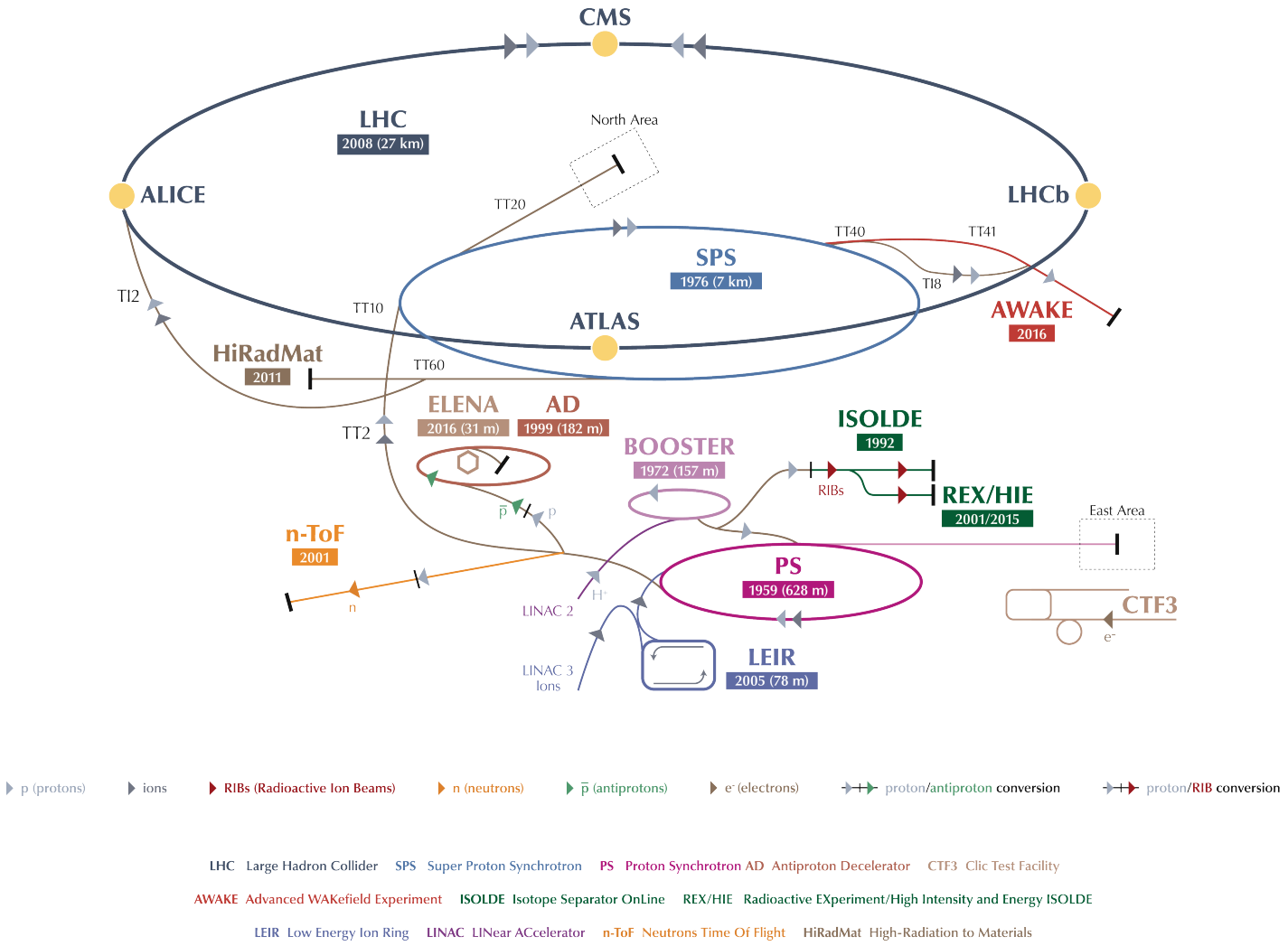


Figure 10: Schematic view of the CERN accelerator chain and the location of the four main experiments along the LHC ring, indicated by the yellow blobs. Image courtesy of CERN.

an energy of 50 MeV. Subsequently, they are sent through the Proton Synchrotron Booster (PSB) and are brought to 1.4 GeV. After that, protons are injected into the Proton Synchrotron (PS), where they are accelerated to 26 GeV. From the PS, they reach the Super Proton Synchrotron (SPS), which increases their energy up to 450 GeV.

Protons from the SPS are finally injected in the LHC as particle *bunches* making up two opposite-directed *beams*, which are kept in circular orbit by means of a 7.7 (8.3) T magnetic field for $\sqrt{s} = 13$ (14) TeV collsion energy, provided by 1392 superconducting Nb-Ti dipole magnets operating at 1.9 K. Beams are focused by 392 quadrupole magnets with strongest focusing power near the interaction points to maximise the collision rate. The beam energy is boosted up to the nominal collision energy through 16 superconducting radiofrequency (RF) cavities displaced around the LHC.

The moment when bunches start being injected into the LHC ring marks the beginning of a *fill*. After the desired number of bunches circulating in the LHC has been reached, and as soon as the beams are in a stable condition, collisions begin taking place. Usually a LHC fill lasts for several hours, up to the point when the beam intensity is reduced down to a critical threshold due to the inelastic collisions. At that point, the beams are dumped from the machine to allow for a new cycle to start.

2.1.2 Beam energy, luminosity and collision parameters

The two main parameters which influence the reach of physics searches and measurements at the LHC experiments are the beam centre-of-mass energy ($\sqrt{s}$) and the instantaneous luminosity (L). The former sets the accessible mass range of the particles produced as final state of the collisions, and it is driven by the features of the RF accelerator system and the maximum bending power of the LHC dipole magnets. The instantaneous luminosity is defined as the proportionality constant between the event rate dN_α/dt for some process of interest α and its cross section σ_α:

$$L = \frac{dN_\alpha}{dt} \cdot \frac{1}{\sigma_\alpha}. \tag{13}$$

Therefore, luminosity is expressed in units of $\mathrm{cm}^{-2} \times \mathrm{s}^{-1}$, and when integrated over time it represents a measurement of the collected data sample size.

Beam/collision parameters	2015	2016	Nominal design (if available)
Centre-of-mass energy ($\sqrt{s}$) [TeV]	13	13	14
Bunch spacing [ns]	50 - 25	25	25
Bunch revolution frequency (f_r) [kHz]	11.245	11.245	11.245
Max. number of bunches/beam (n_b)	2232	2208	2808
Max. charge per bunch colliding (p/bunch)	1.21×10^{11}	1.31×10^{11}	1.15×10^{11}
Peak instantaneous luminosity [$\mathrm{cm}^{-2}\mathrm{s}^{-1}$]	5.02×10^{33}	1.38×10^{34}	10^{34}
Max. average nr. of interactions per bunch crossing	28.2	52.2	
Longest stable beams fill duration [h]	24.3	37.03	

Table 4: Some collision parameters of the LHC in 2015 and 2016 as taken from the ATLAS detector measurements [59], compared with the machine design values.

The luminosity is completely determined by the parameters of the circulating beams, and by the collider magnets optics. Assuming Gaussian transverse profiles of the beams, it holds:

$$L = \frac{f_r n_b N_b^2}{4\epsilon\beta^*}, \tag{14}$$

where N_b (the number of protons in each colliding bunch, assuming them equal), n_b (the number of circulating bunches), f_r (the bunch revolution frequency) and ϵ (the normalised transverse beam emittance, a quantity related to the transversal dispersion of the beam particles) are set by the injection chain. The β^* is the so-called amplitude function at the interaction point, and it is a property set by the configuration of the focusing magnets (quadrupoles and higher order magnets).

The product $\epsilon\beta^*$ is directly proportional to the beam area transverse to the beam direction: the smaller the emittance and amplitude function, the narrower the beam cross-section will be and the higher the event rate. Additional luminosity reduction effects, such as a nonzero bunch crossing angle and longitudinal bunch length, are neglected in the above formula.

Details about the collision parameters for the years 2015 and 2016, and comparison to the design values are given in Table 4. The performance of the LHC machine have been improving throughout the years of operation, and some of such parameters already exceed the design value. In particular, the maximum value for the instantaneous luminosity has surpassed the 10^{34} $\mathrm{cm}^{-2}\mathrm{s}^{-1}$ threshold during the 2016 data taking.

2.2 THE ATLAS EXPERIMENT

At the energy scale of the LHC, the cross sections for rare processes like the Higgs boson production are several orders of magnitude smaller than QCD multi-jet production cross section, as Figure 11 reveals. Consequently, it is a crucial requirement to be able to detect and identify with high efficiency the experimental signatures characteristic of those processes, such as high transverse momentum light leptons, missing transverse energy and secondary vertices.

Furthermore, the high luminosity needed to pursue the objectives of the LHC physics programme makes it very likely that a large number of protons undergo inelastic scattering simultaneously with the hard interaction. Such scattering events are soft; they are characterised by low four-momentum transfer, hence the final state products have on average low transverse momentum, of order $p_T \approx 0.5$ GeV. Soft interactions - which are more frequent than hard scattering ones - are known as *minimum bias* events, and the overlap of soft inelastic scattering events to the typically unique hard collision is referred to as *pile-up* (PU). This is more precisely defined as the *in-time* pile-up, as it is an effect of multiple collisions happening simultaneously in a single proton bunch crossing. There is also an *out-of-time* pile-up effect, due to detector response extending over several bunch crossings. Pile-up represents a formidable experimental challenge, as it degrades significantly the performance of most physics object reconstruction algorithms by introducing spurious activity in the detectors. Also, pile-up interactions represent a major source of background for numerous searches.

In order to fulfil its ambitious physics goals, the ATLAS detector has been designed in accordance to some key requirements [36]:

- Fast response electronics, to be able to detect interesting physics signatures minimising the effect of overlaying interactions.
- Resistance of all detector components to high radiation doses, both in terms of operation and ageing.
- High detector granularity, necessary to handle high particle fluxes.
- Efficient trigger system, to cope with the limited bandwidth for event storage and select only interesting collision events for offline physics analysis.
- Maximal solid angle coverage around the interaction point.

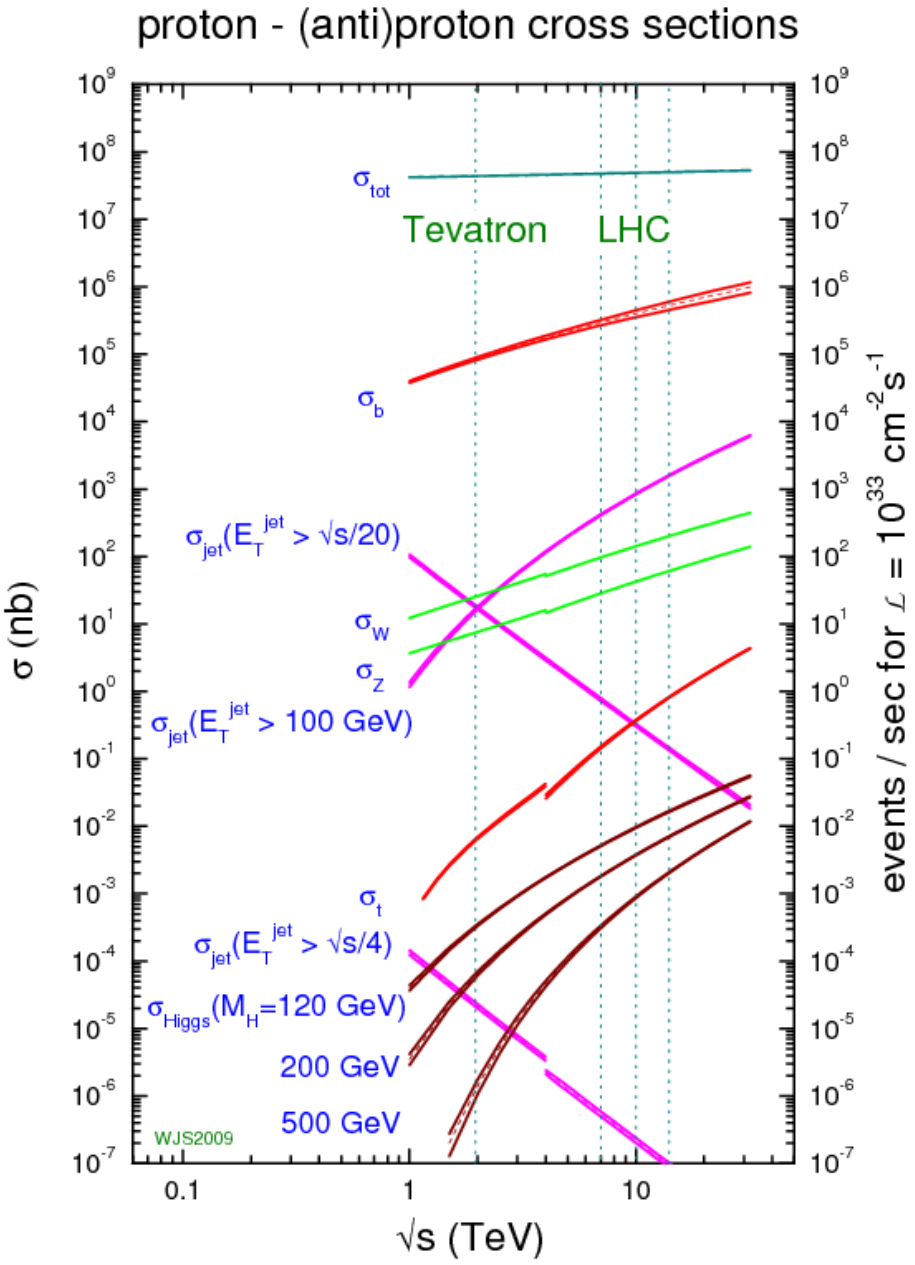


Figure 11: Production cross sections calculated at next-to-leading order in perturbation theory for signal and background processes at hadron colliders as a function of the centre-of-mass energy. σ_{tot} represents the prediction for the total inelastic cross section [60].

- Efficient particle reconstruction and identification, robust against pile-up.

The ATLAS detector surrounds the interaction region at Point 1 of the LHC collider. It is 25 m in height and 44 m in length, with an overall weight of approximately 7000 tonnes. A sketch of the detector is displayed in Figure 12, and a description of its component is provided in the following, starting from the innermost subsystems outwards. First, a brief explanation of the ATLAS coordinate system is presented.

2.2.1 ATLAS coordinate system

The origin of the coordinate system used in ATLAS is defined by the nominal interaction point (IP). A right-handed scheme is adopted: the direction of the beam pipe defines the z-axis, with the x-axis pointing towards the centre of the of the LHC ring, and the y-axis pointing upwards. The ATLAS detector has to good approximation a cylindrical symmetry along the z axis. The azimuthal angle ϕ is defined in the transverse (x, y) plane around the beam axis between $-\pi$ and π, and the polar angle θ is defined between 0 and π, with $\theta = 0$ lying on the positive z axis.

Since the interacting partons carry a random fraction of the incoming proton momentum, the centre-of-mass system of the two partons has a nonzero, unknown Lorentz boost in the laboratory frame along the z direction. However, variables involving only the transverse components such as ϕ, transverse momentum $p_T = \sqrt{p_x^2 + p_y^2}$, and transverse energy $E_T = E \sin\theta = \sqrt{p_T^2 + m^2}$ are invariant under longitudinal boosts. Furthermore, differences in pseudorapidity - where the latter is defined as $\eta = -\ln\left(\tan\frac{\theta}{2}\right)$ in the range $-\infty < \eta < \infty$ - are boost-invariant as well in the limit when $m \ll E$. It is therefore common to express the angular distance between physics objects as $\Delta R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2}$.

2.2.2 Inner Detector

The Inner Detector (ID) [61] is 7 m long, with an outer radius of 1.15 m and it is the subdetector closest to the interaction point. The layout of the ATLAS Inner Detector system is presented in Figure 13. It provides the measurement of the position of electrically charged particles which fly through the ID layers. Individual measurements of the position in each subsequent layer can be combined to form *tracks*, which can be further combined to

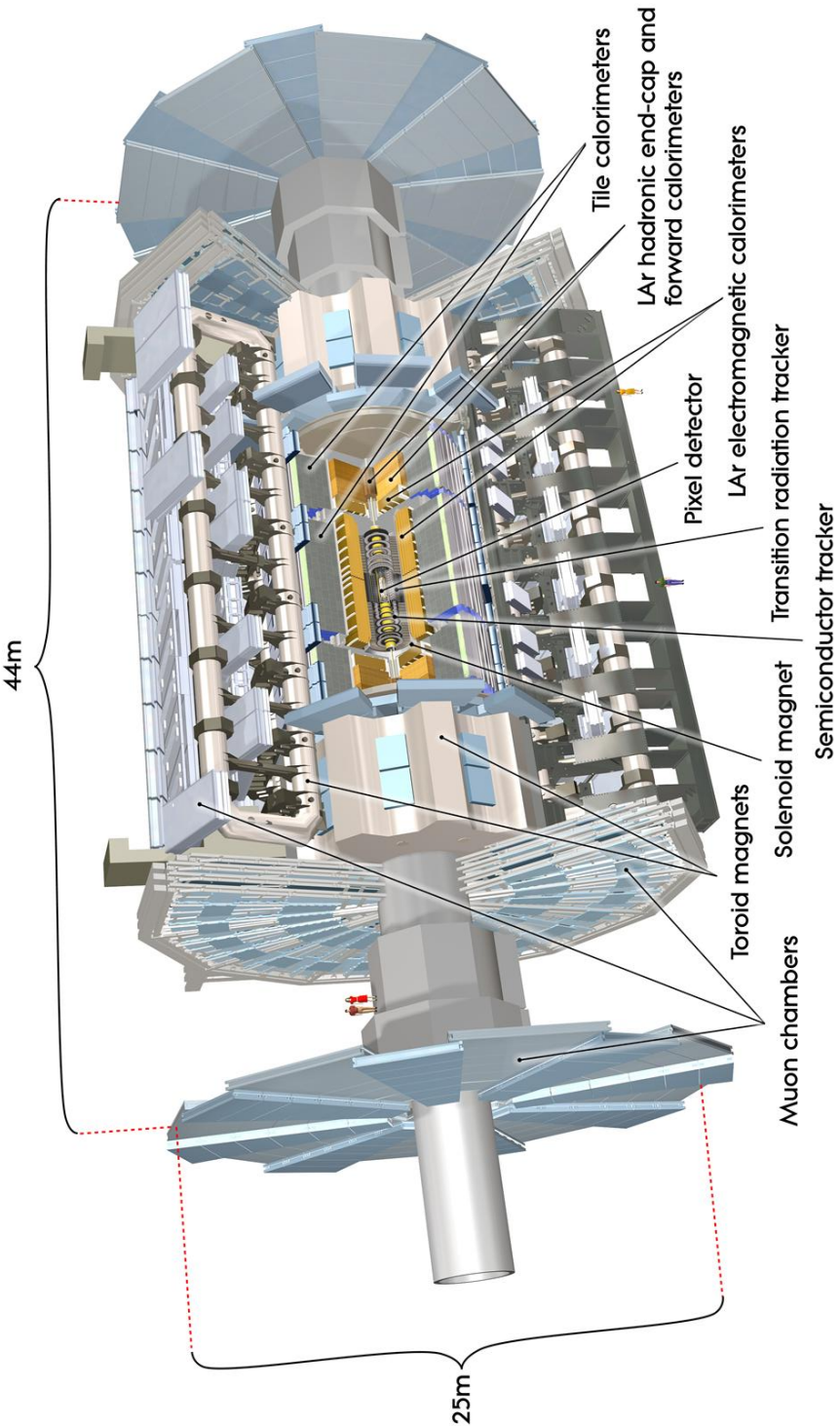


Figure 12: Schematic view of the ATLAS detector. Image courtesy of the ATLAS Experiment.

determine primary or secondary vertices of the interaction. Particle momentum can be measured from the curvature of the tracks, which are assumed to describe a helical trajectory due to 2 T solenoidal magnetic field provided by a superconducting magnet which completely surrounds the ID.

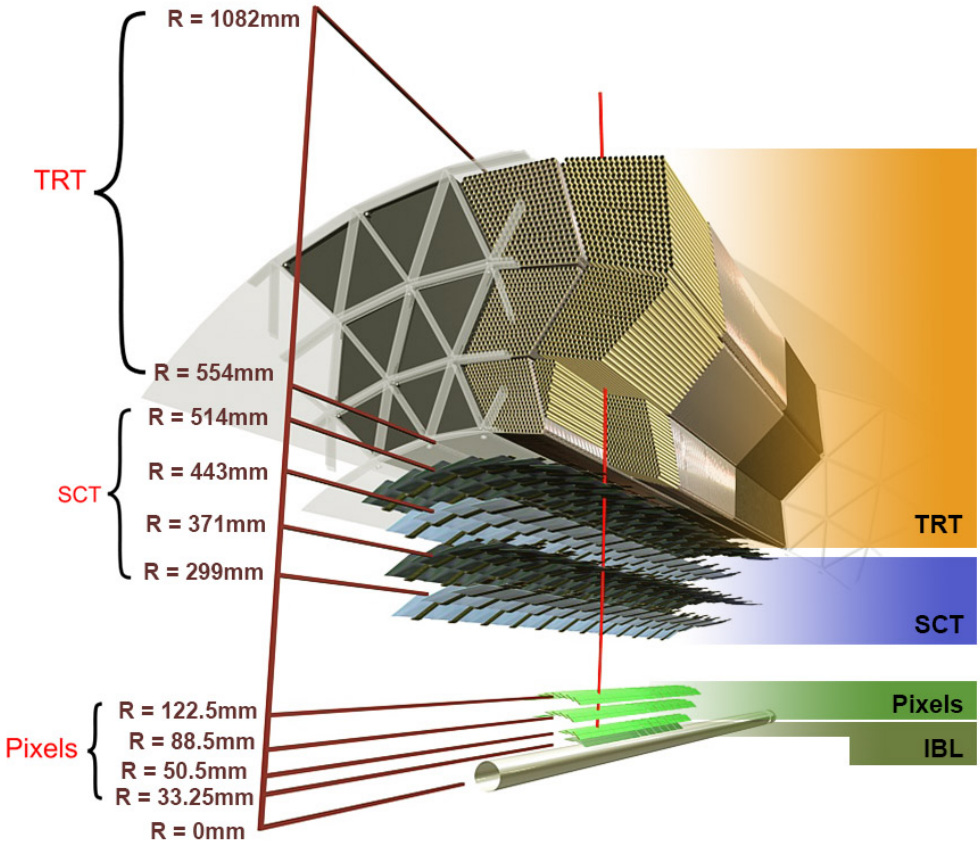
The ID system is composed of three sections. The Pixel Detector and the Strip (SCT) Detector are closest to the beam; they are both silicon-based detectors and characterised by high granularity to cope with higher occupancy near the interaction point. The outermost section - the Transition Radiation Tracker - is a straw-tube gas detector which provides additional track measurements and particle identification. Along the z direction, the ID is divided into a barrel and two end-cap regions. In the barrel, layers are arranged in concentric cylinders around the beam pipe, whereas in the end-caps they are mounted as disks perpendicular to the beam axis. A total coverage of $|\eta| < 2.5$ by the Pixel and SCT detectors is achieved.

The Pixel Detector consists of 4 barrel layers and 3 end-cap disks. It underwent a major re-design during the long machine shutdown between 2013 and 2015, with the addition of an innermost fourth layer at a radius of 3.3 cm (the Insertable B-Layer, IBL), the adoption of a new, thinner beam pipe, improved readout electronics and repairs for damaged modules [64].

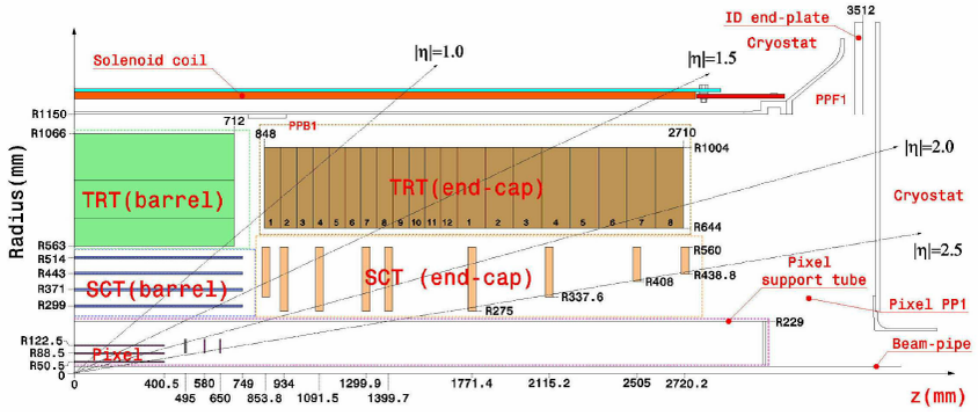
Pixels are arranged into modules, with a total number of 1744 modules in the barrel, and 288 in the end-cap sections. Each module has 46080 elements, accounting for 80 million pixels in total with a typical size in $(r-\phi) \times z$ of 50×400 μm^2, and a module thickness of 250 μm. The Pixel Detector has an intrinsic accuracy of 10 μm in $(r-\phi)$ and 115 μm in z. It provides on average four space-point measurements per track, and mostly determines the impact parameter resolution and vertexing performance. The Pixel Detector two-dimensional fine segmentation comes at the cost of very complex electronics, with about 80.4 million read-out channels in total.

Moving outwards from the interaction point, the Semiconductor silicon Strip Detector (SCT) completes the high precision tracking system. The barrel comprises four cylindrical layers of silicon microstrips modules, each made up of two sensors at a 40 mrad stereo angle in order to provide a three-dimensional measurement, with a strip pitch of about 80 μm. The end-cap SCT is composed of 9 disks each. The intrinsic accuracies of the SCT are 17 μm on the $(r-\phi)$ direction and 580 μm along z. On average, the SCT adds additional 8 space point measurements per track. The total number of read-out channels is approximately 6.3 million.

The Transition Radiation Tracker (TRT) is made up of $\approx 300,000$, 4 mm-diameter straw drift tubes filled with a mixture of 70% Xe, 20% CF_4, 10%



(a) $(r-\phi)$



(b) $(r-z)$

Figure 13: View of the ATLAS ID 13a in the $(r-\phi)$ plane and 13b in the $(r-z)$ plane [62, 63]. In the barrel, a charged particle track produced in the primary interaction sequentially crosses the beryllium beam-pipe, the four cylindrical silicon pixel layers (the innermost one, known as the IBL, included since 2015), the four cylindrical silicon microstrip (SCT) double layers (one axial and one with a stereo angle of 40 mrad) of pitch 80 mm, and approximately 36 axial straws of 4 mm diameter contained in the transition radiation tracker (TRT) modules. The $(r-z)$ section does not show the recently added IBL pixel layer.

CO_2 gas, arranged parallel to the beams in the barrel region and radially in the end-cap, with a total number of readout channels of about 351,000. Charged particles travelling across the straw ionise the gas, and the resulting free electrons drift under an applied voltage and get collected by a wire at the centre of the straw. In the barrel, only the measurement on the $(r-\phi)$ plane is provided, and on $(r-z)$ in the end-cap, with an intrinsic accuracy of 130 μm per straw. On average, a track within $|\eta| < 2.0$ has 36 measurements in this detector.

The TRT contributes also to charged particle identification. TRT tubes are interleaved with layers of polypropylene fibres and foils: a charged particle that passes through the boundary region between materials with a different refraction index emits X-ray radiation whose intensity is proportional to the relativistic γ factor of the particle itself. Electrons produced by photoelectric effect of such X-rays in the straw tube gas give a signal which has a higher amplitude than the signal originating from particles passing by. Given their light mass, electrons start producing transition radiation when their momentum is close to 1 GeV, while pions start to radiate only in the $\mathcal{O}(100)$ GeV momentum range. This piece of information is relevant for the electron identification algorithms.

The combined information of the Pixel, SCT and TRT subsystems allows measurement of track parameters for charged particles with p_T down to 400 MeV, with a design relative resolution for the measurement of transverse momentum described by [36]:

$$\frac{\sigma_{p_T}}{p_T} = 0.05\% p_T \oplus 1\%. \tag{15}$$

2.2.3 Calorimetry

The ATLAS calorimetry system [36] comprises several components with diverse technology and granularity, covering a very large range of pseudorapidity ($|\eta| < 4.9$). All components are sampling calorimeters, with passive, inactive material to generate the electromagnetic/hadronic shower alternated to active material layers to detect a fraction of the incoming particle energy. The energy calibration has been carried out through test beam [65, 66] and validated in collision data. The layout of the ATLAS calorimeters is sketched in Figure 14.

The innermost section right outside the solenoid is the electromagnetic calorimeter (EM), devoted to the energy measurement of electrons and photons. It uses liquid argon (LAr) as active medium, which is characterised

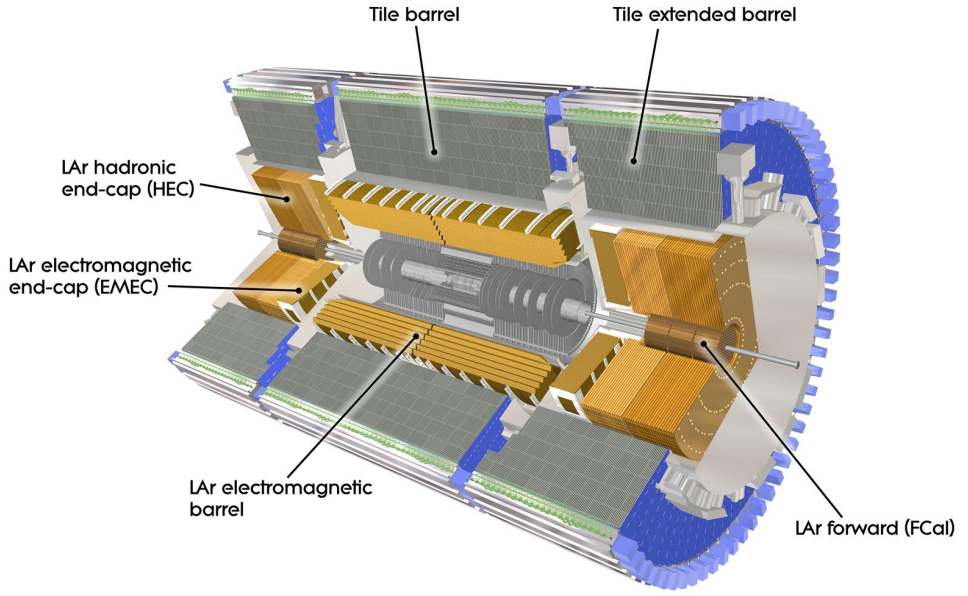


Figure 14: Schematic view of the ATLAS Calorimeter system. Image courtesy of the ATLAS Experiment.

by excellent radiation hardness and energy resolution, and lead to generate the EM showers. The electrodes are capton-plated copper plates segmented into strips which constitute the read out cells. The barrel section of the EM calorimeter covers up to $|\eta| = 1.4$, and the end-cap system extends to $|\eta| = 3.2$. It is divided into 3-4 layers, with an accordion-like geometry to provide full ϕ coverage without dead regions, and features a fine segmentation into $\Delta\eta \times \Delta\phi$ down to a value of 0.025×0.025 within the ID acceptance ($|\eta| < 2.5$) to resolve $\pi_0 \to \gamma\gamma$ decays efficiently. The thickness of the EM calorimeter is $> 22\ X_0$ radiation lengths in the barrel, and $> 24\ X_0$ in the end-cap region to ensure good containment of the electromagnetic cascade. The transition region $1.37 < |\eta| < 1.52$ between barrel and end-cap is expected to have poorer performance because of the lack of instrumented material, since most of the service infrastructure is located there. The energy resolution achieved in the EM calorimeter is described as [36]:

$$\frac{\sigma_E}{E} = \frac{10\%}{\sqrt{E}} \oplus 0.7\%. \tag{16}$$

The hadronic calorimeter (HCal) encloses the EM one, and is dedicated to the energy measurement of hadrons such as neutrons, pions and kaons. The barrel region $|\eta| < 1.7$ is serviced by a steel-plastic scintillator sampling calorimeter (TileCal), with 3 layers providing a total thickness of 9.7 λ

interaction lengths at $\eta = 0$. This minimises the punch-through effect of hadronic showers decay products reaching the outer muon spectrometer. The hadronic end-cap calorimeter (HEC) matches the $1.5 < |\eta| < 3.2$ range of the end-cap EM calorimeter and shares the same liquid argon technology, although using copper as passive material and presenting a planar geometry. The granularity of the hadronic calorimeter is coarser than for the EM, of order 0.1×0.1 $\Delta\eta \times \Delta\phi$, and the energy resolution is worse due to the inelastic nature of the hadronic interactions with the material nuclei [36]:

$$\frac{\sigma_E}{E} = \frac{50\%}{\sqrt{E}} \oplus 3\%. \tag{17}$$

Finally, a 3-layered forward calorimeter (FCal) ensures hermeticity by covering the high-pseudorapidity range $3.1 < |\eta| < 4.9$, to contribute to the measurement of missing transverse energy and forward jets. It exploits LAr technology with copper and tungsten as absorber material for the first and last two layers, respectively.

2.2.4 Muon Spectrometer

Final states with high energetic muons are among the most important signatures for several physics searches at ATLAS, including the $t\bar{t}H$ search. Muons are unique among all detectable particles, being the only ones able to travel beyond the hadronic calorimeters. ATLAS has a dedicated muon detection system (Muon Spectrometer, MS) [36], which surrounds the other ATLAS sub-detectors hermetically. Figures 15, 16 depict a full scheme of the ATLAS muon spectrometer.

The MS is based on the deflection on the $(r - z)$ plane of the trajectories of muons by means of air-core super conducting toroid magnets, instrumented with trigger and high precision tracking chambers. The field in the $|\eta| < 1.4$ region is provided by a large 8-coil barrel toroid with a bending power of ≈ 2.5 Tm. Two smaller end-cap magnets provide a ≈ 6 Tm bending power field in the region up to $|\eta| = 2.7$. This configuration provides excellent stand-alone muon momentum resolution, minimising the multiple scattering degradation effect due to material crossing [36]:

$$\frac{\sigma_{p_T}}{p_T} = 10\% \ (p_T = 1 \text{ TeV}) \qquad \frac{\sigma_{p_T}}{p_T} = 3\% \ (p_T = 100 \text{ GeV}) \tag{18}$$

Measurements in the bending plane for muon track reconstruction are provided by Monitored Drift Tubes (MDT) precision chambers, arranged in 3 cylindrical layers around z in the barrel in between the 8 coils of the

magnet system, and in 3 layers perpendicular to the beam in the end-cap, with an intrinsic resolution of 35 μm per chamber. In the $2.0 < |\eta| < 2.7$ region a higher particle flux is expected, therefore additional Cathode Strip Chambers (CSC) with higher granularity are used in the innermost layer.

Within $|\eta| < 2.4$, the MS is equipped to provide fast muon detection for triggering purposes as well as to measure the track coordinate in the non-bending plane. Resistive Plate Chambers (RPC) are used in the barrel ($|\eta| < 1.05$) and Thin Gap Chambers (TGC) in the end-cap ($1.05 < |\eta| < 2.4$). Both are characterised by lower spatial resolution ($\mathcal{O}(5-10)$ mm) than the precision tracking chambers, but have much better time resolution (≈ 1.2 ns) to quickly determine the passage of a muon.

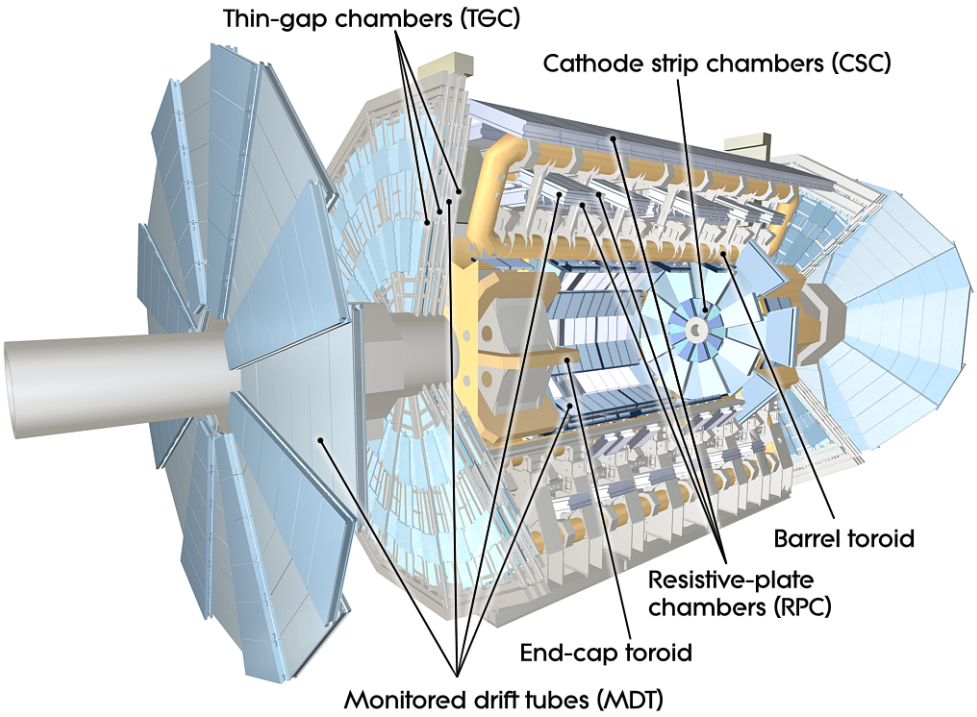


Figure 15: Cut-away view of the ATLAS Muon Spectrometer (MS). Image courtesy of the ATLAS Experiment.

2.2.5 Luminosity measurement and data quality

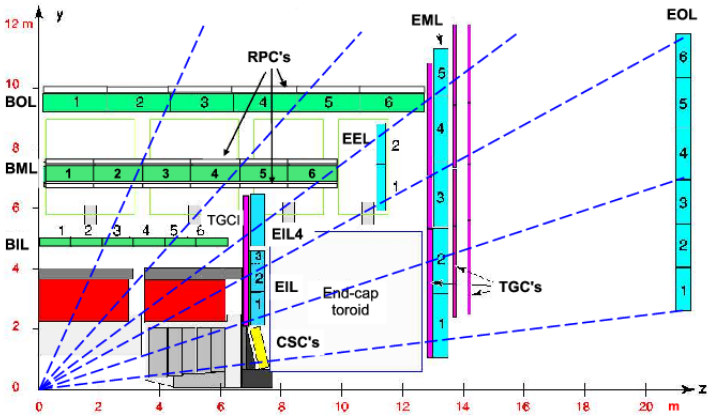
As introduced in Sec. 2.1.2, having an accurate measurement of the luminosity is crucial in order to reliably convert event rates into a cross section. The procedure followed by ATLAS [68] consists in measuring the total p-p inelastic scattering rate within the detector acceptance, and calibrating to its visible cross section:

$$L = n_b \cdot \frac{f_r \mu^{vis}}{\sigma_{inel}^{vis}}, \tag{19}$$

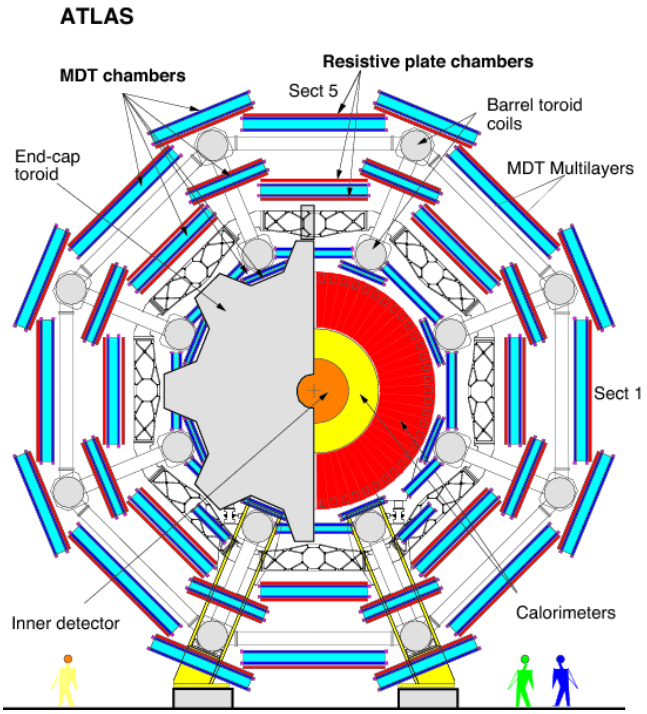
where the parameter μ^{vis} is the observed average number of inelastic interactions per bunch crossing (i.e., the pile-up), f_r is the bunch revolution frequency, n_b the number of circulating bunches, and σ_{inel} is the p-p inelastic scattering cross section to which the luminosity is calibrated.

The inelastic cross section is obtained exploiting the van der Meer technique [69], by performing beam separation scans to determine the beam parameters and therefore the luminosity scale which combined with μ^{vis} in the above formula yields σ_{inel}^{vis}. These scans are performed only every few months during LHC operations, and the beam conditions during van der Meer scans are different from those in normal physics operation, with lower bunch intensities and only a few tens of widely spaced bunches circulating. These conditions are optimised to reduce various systematic uncertainties in the calibration procedure.

Several independent luminometers are used in ATLAS to ensure different acceptance phase spaces are covered, and systematic uncertainties on the luminosity measurement are controlled. The two primary systems are the LUCID (LUminosity measurement using a Cherenkov Integrating Detector) and BCM (Beam Conditions Monitor) detectors. LUCID [70] is a Cherenkov detector, consisting of 16 PMT surrounding the beam pipe on each side of the IP at a distance of 17 m, covering the $5.6 < |\eta| < 6.0$ range, which read out the Cherenkov light produced in their own quartz window. BCM [71] is made up of four 8×8 mm^2 diamond sensors arranged around the beam pipe in a cross pattern at $\pm$ 1.84 m on each side of the ATLAS IP. A different method is based on charged track multiplicity monitoring in randomly selected colliding-bunch crossings [68], therefore relying on the Inner Detector performance.



(a) $(r-z)$ plane



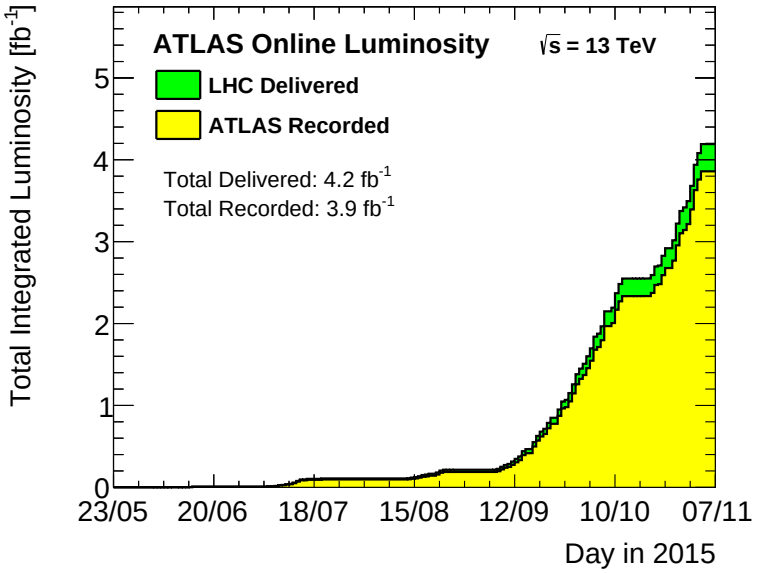
(b) $(r-\phi)$ plane

Figure 16: 16a: section of the muon spectrometer in the bending $(r-z)$ plane. The barrel MDT chambers are shown in green, the endcap MDT chambers are blue. In the barrel (endcaps), the RPC (TGC) chambers are shown outlined in black (solid purple). The straight dashed lines would indicate trajectories of infinite-momentum muons. 16b: section of the muon spectrometer in the non-bending $(r-\phi)$ plane, showing 3 concentric cylindrical layers of 8 large and 8 small chambers [36, 67].

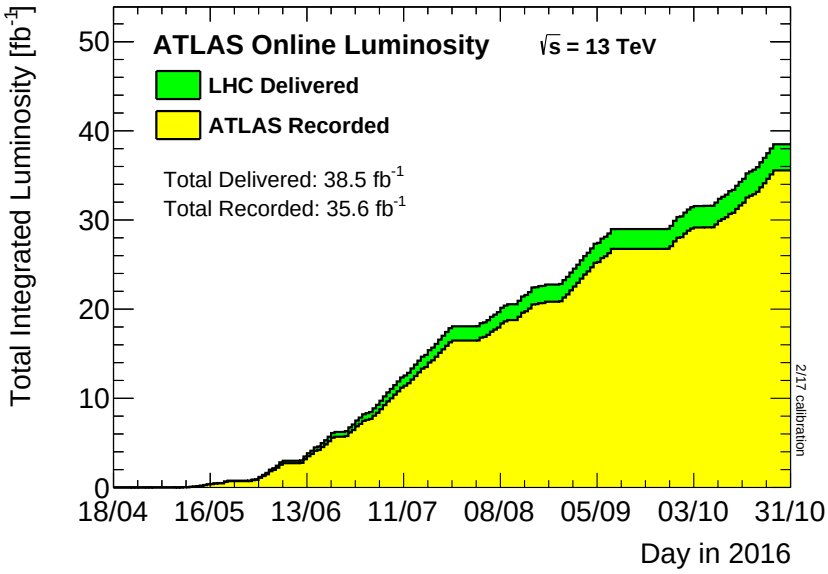
The basic time unit for storing ATLAS luminosity information for physics use is the *luminosity block* (LB). The boundaries of each LB are defined to ensure data taking conditions such as trigger prescales and the luminosity itself are constant over the LB length, and in general the duration of each LB is approximately one to two minutes. For each LB, the instantaneous luminosity from each detector and algorithm, averaged over the luminosity block, is stored, along with a variety of general ATLAS data-quality information. The time period across which both the beam and detector conditions are stable is called a *run*. To define a data sample for physics, quality criteria are applied to select LBs within a run where detector conditions are acceptable. Then the instantaneous luminosity in each LB is multiplied by the LB duration to provide the integrated luminosity delivered in the LB. Additional corrections can be made for trigger dead-time and trigger prescale factors. Adding up the integrated luminosity delivered in a specific set of luminosity blocks provides the integrated luminosity of the entire data sample.

Figure 17 shows the amount of integrated luminosity delivered by the LHC during the years 2015 and 2016, the fraction recorded by ATLAS and the effective fraction usable by physics analyses. The inefficiency of ATLAS in recording data provided by the LHC is mostly due to occasional dead-time from temporary faults of some detector components, and because of the time required by the Inner Detector to ramp up the high voltage and become fully operative from the moment when stable beams are declared by the LHC operation team. The amount of valid data collected by ATLAS for the 2015 and 2016 operations used in the context of this thesis corresponds to an integrated luminosity of 36.1 fb^{-1}

Figure 18 shows the luminosity-weighted distribution of the mean number of interactions per crossing μ measured by the ATLAS experiment during the 2015 and 2016 data taking. The mean number of interactions per crossing corresponds the mean of the Poisson distribution of the number of interactions per crossing for each bunch. It is calculated from the instantaneous luminosity per bunch according to Eq. (14). The inelastic cross section is taken to be $\sigma_{inel}^{vis} = 80$ mb, based on van der Meer beam-separation scans performed in 2016. An average value of $\langle\mu\rangle = 23.7$ interactions per crossing is obtained, with a peak mean value recorded of 52.2, which gives a scale of the pile-up effect.



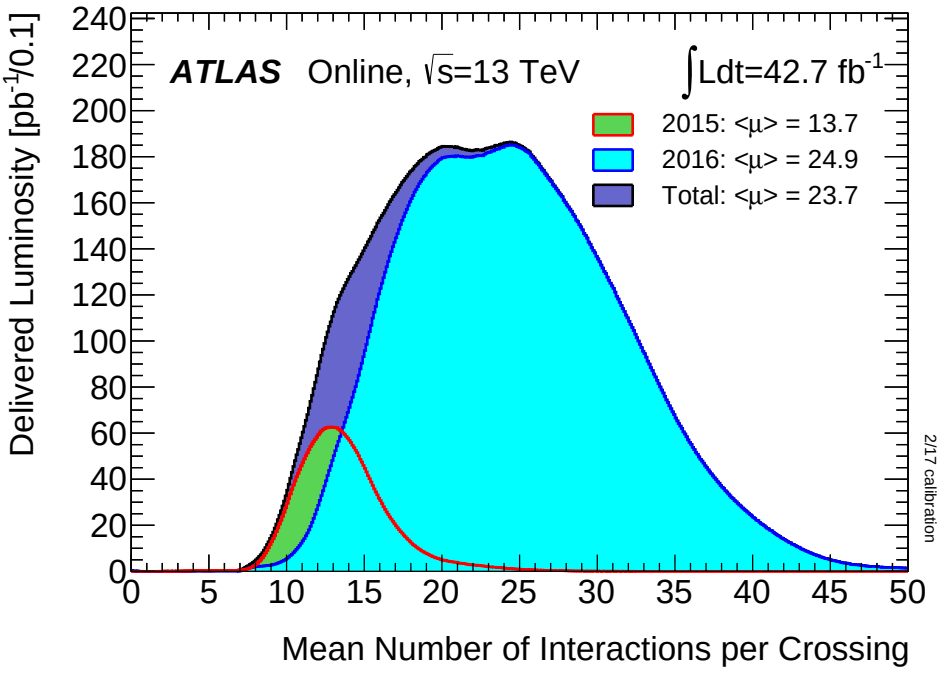
(a) 2015 data taking



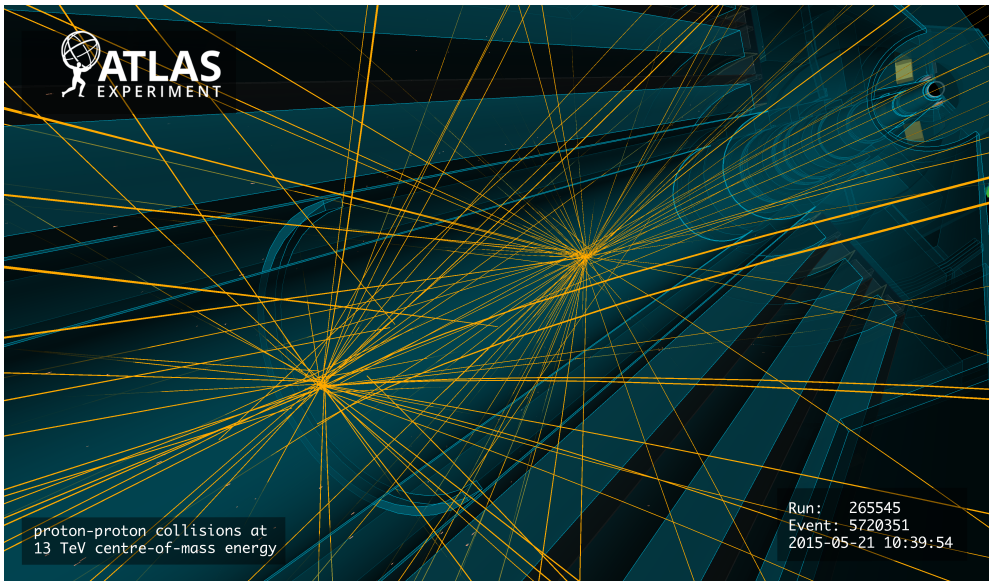
(b) 2016 data taking

Year	$\int L\,dt$ [fb^{-1}] used for physics analysis
2015	3.2
2016	32.9
Total	36.1

Figure 17: Cumulative luminosity versus time delivered by the LHC and recorded by ATLAS during stable beams for p-p collisions at $\sqrt{s} = 13$ TeV in 2015 (17a) and 2016 (17b) [72]. The effective size of the data sample used for physics analysis after data quality assessment is indicated in the bottom table.



(a) μ



(b) ATLAS event at $\sqrt{s} = 13$ TeV - 21 May 2015

Figure 18: 18a: luminosity-weighted distribution of the mean number of interactions per crossing for the 2015 and 2016 p-p data taking [72]. 18b: display of a p-p collision event recorded by ATLAS on 21 May 2015, at 13 TeV collision energy. Tracks reconstructed from hits in the inner tracking detector are shown to originate from two interaction points, indicating a pile-up event [73].

2.2.6 Trigger

With 25 ns bunch spacing, the nominal bunch crossing rate is 40 MHz, leading to an approximate proton inelastic scattering rate of 1 GHz, given $\langle\mu\rangle = 23.7$. Considering the limits of the electronic read-out systems and data storage capabilities, the ATLAS trigger and data acquisition system (TDAQ) [74] is a fundamental component of the detector, being responsible for deciding whether or not to save an event from a bunch crossing for later offline study.

It consists of two levels: the hardware-based Level-1 trigger (L1), which accepts events down to a 100 kHz rate limit defined by the ATLAS detector read-out capability, and the software-based High-Level Trigger (HLT) system. The HLT operates on a standard computing farm of 40,000 CPUs, and selects events at an average rate of 1 kHz during one LHC fill, the maximum that can be handled by the offline computing model and storage.

The L1 trigger comprises the Central Trigger Processor (CTP), which processes inputs from the L1 calorimeter (L1Calo) and L1 muon (L1Muon) trigger subsystems. The L1Calo sources information directly from the calorimeters. The L1Muon exploits measurements from the RPC in the inner barrel ($|\eta| < 1.0$) and TGC in $1.0 < |\eta| < 2.4$ subsystems of the Muon Spectrometer. To cope with the higher event rate and efficiently select relevant events from a physics point of view, a new element known as the L1 Topological Processor (L1Topo) [75] has been added to the L1 since 2016. The L1Topo system sources information from both L1Calo and L1Muon, and allows computation of high-level quantities such as invariant masses to be considered for the L1 decision. With a latency of the L1 electronics of 2.5 μs, the CTP also applies preventive dead-time, by setting a minimum time between two consecutive accept decisions by L1 to avoid overlapping readout windows, and restricting the number of L1 accept decisions allowed in a given number of bunch-crossings to avoid front-end buffers from overflow.

After L1 trigger accept, events are processed by the HLT using finer-granularity calorimeter information, precision measurements from the MS and tracking information from the ID, which are not available at L1. To maximise the efficiency, the HLT software is tuned to make the algorithms and selections as close as possible to the offline reconstruction. Events accepted by the HLT are finally stored locally at the experiment site, and exported to the CERN computing centre for offline reconstruction. As needed, the HLT can process information either from Regions of Interest (RoIs) identified at

L1, or from the full detector. A complete scheme of the ATLAS trigger system is shown in Figure 19.

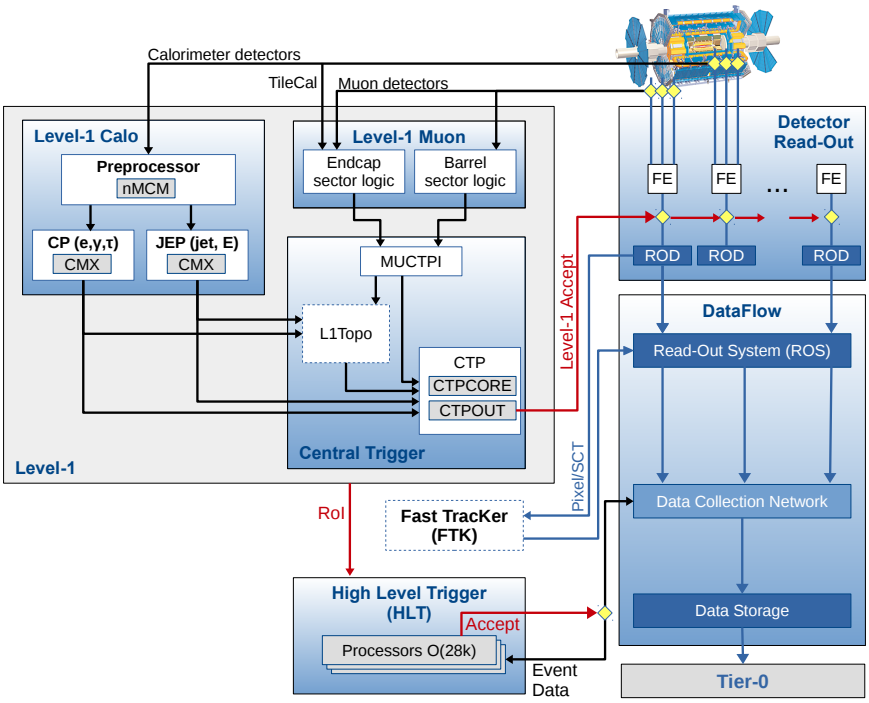


Figure 19: Block diagram of the ATLAS Trigger and Data Acquisition (TDAQ) system adopted since 2015 [74]. Note that the FTK subsystem was not operative during the data taking period relevant to this thesis.

The trigger decisions at L1 and HLT steps are the result of a set of requirements on physics object multiplicity and selection algorithms. These define the so-called *trigger menu*. The primary triggers in the menu cover all the signatures required by the diverse ATLAS physics searches, including electrons, muons, photons, tau leptons, jets and missing transverse energy. The trigger menu composition and trigger thresholds are optimised for several luminosity ranges in order to maximise the physics output of the experiment and to fit within the rate and bandwidth constraints of the ATLAS detector. The main signature of many ATLAS analyses, including the $t\bar{t}H$ search, is the presence of electrons or muons in the final state. Therefore, the trigger items in the menu requiring the presence of at least one lepton in the event

with $p_T \gtrsim 25$ GeV take up a large fraction of the available bandwidth, as Figure 20 shows.

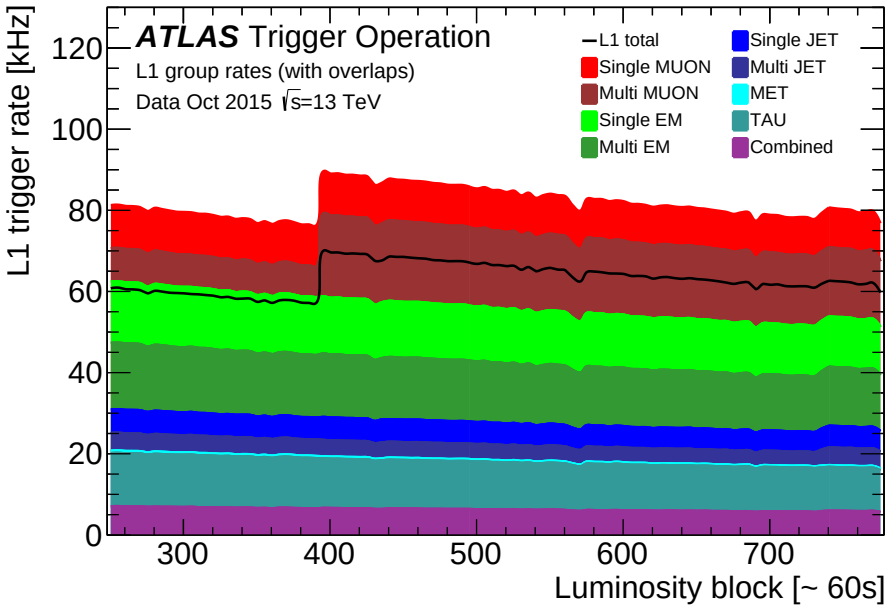
2.3 PARTICLE RECONSTRUCTION AT ATLAS

The process of converting low-level electronic space-point or energy measurements in each subsystem of the detector to a set of high-level information associated to particles produced in the collision is known as *reconstruction*. Reconstruction in ATLAS is a complex multi-step procedure, carried out by numerous algorithms organised in a centralised software framework [76, 77], whose final output are standard collections of *physics objects* (also referred to simply as "objects") which serve as inputs for the various physics searches. Such classes of objects include electrons, photons, muons, jets and hadronic tau candidates, as well as missing transverse energy. The basic steps through which they are reconstructed will be outlined in the following, with more focus on the objects that are relevant to the $t\bar{t}H$ search presented in this thesis.

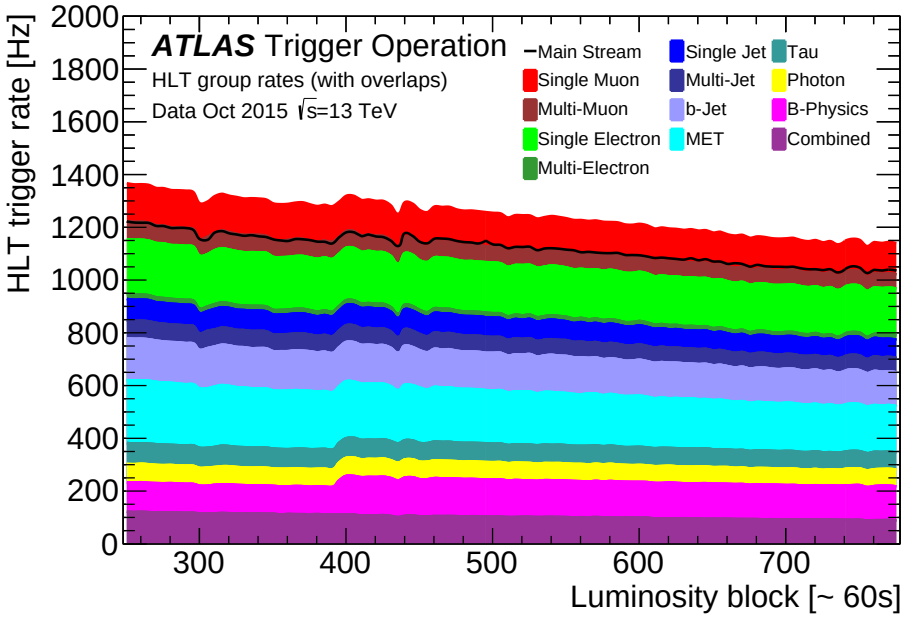
2.3.1 Tracks and energy clusters

The basic common inputs of particle reconstruction are *tracks* and *energy clusters*.

Tracks describe charged particle trajectories as measured in the ID and independently, in the MS. The initial seeds of the track reconstruction chain [78] are three-dimensional space-point measurements in the silicon detector (Pixel and SCT). These are combined by means of a combinatorial iterative track finding algorithm to form track candidates, which are subsequently fitted to helical trajectories in the ID magnetic field, taking into account material effects such as energy losses and multiple scattering, and field non uniformities [79]. Ambiguities are then solved by ranking tracks in terms of each candidate's properties such as number of shared measurements and quality of the fit. Finally, tracks are extrapolated to the TRT, and a new fit on the full track is performed. More information about the ATLAS tracking algorithm will be provided in Chapter 3, where the specific case of tracking performance in the core of high p_T jets and hadronic taus will be described. Reconstructed tracks are then used by dedicated algorithms [80] to identify the vertex candidates in each bunch crossing - including secondary vertices



(a) L1 rates



(b) HLT rates

Figure 20: 20a L1 and 20b HLT trigger rates grouped by trigger signature during an LHC fill in October 2015. Due to overlaps the sum of the individual groups is higher than the L1 total rate and HLT output stream rate, respectively, which are shown as black lines. The combined group includes multiple triggers combining different trigger signatures such as electrons with muons, taus, jets or missing transverse energy [74].

for heavy flavour jet identification - and precisely reconstruct their actual position.

Track reconstructed in the MS are joined to ID tracks to define muon candidates' trajectories. The MS track finding algorithm [81] starts by searching for hit patterns in the precision chambers to form straight track segments in each layer; muon MS tracks candidates are then found by a segment-seeded combinatorial algorithm, and the hits associated to each candidate are fitted and eventually accepted/rejected according to the global fit χ^2 score.

Energy clusters from the calorimeters represent the other basic input to particle reconstruction in ATLAS. Electrons, photons and hadrons crossing through the calorimeters lose energy by creating cascades of secondary particles which interact with the active material, leaving a signal spread over several calorimeter cells both longitudinally and in the lateral direction. Showers from electrons and photons are mostly contained within the EM calorimeter, whereas hadrons generally generate showers in the HCal. The clustering algorithms [82] group cells together, with different logics according to the type of particle the clusters will be used as input for the reconstruction, and sum the total deposited energy within each cluster, determining the cluster position as well. Energies are calibrated to account for deposits outside the cluster and in dead material, depending on the incoming particle type (electrons and photons, or hadronic jets).

Combining tracks and calibrated calorimeter clusters finally allows the definition of the physics objects here described. An illustrative sketch in the $(r-\phi)$ plane of the typical patterns of charged particle tracks and energy deposits in ATLAS - which allow the identification of different types of particles - is presented in Figure 21.

2.3.2 Electrons and photons

Electrons and photons leave very similar signatures in the EM calorimeters. The main difference between the two objects is the presence in the electron case of an ID track consistent with the primary interaction vertex that matches a calorimeter EM cluster. Photons represent the final state of one of the dedicated $t\bar{t}H$ searches in ATLAS, and they are characterised by very good energy resolution. However, photons are not found in any of the final states of the $t\bar{t}H$ multileptonic analysis, therefore their reconstruction will not be covered at this stage. Details of photon reconstruction can be found in Ref. [83].

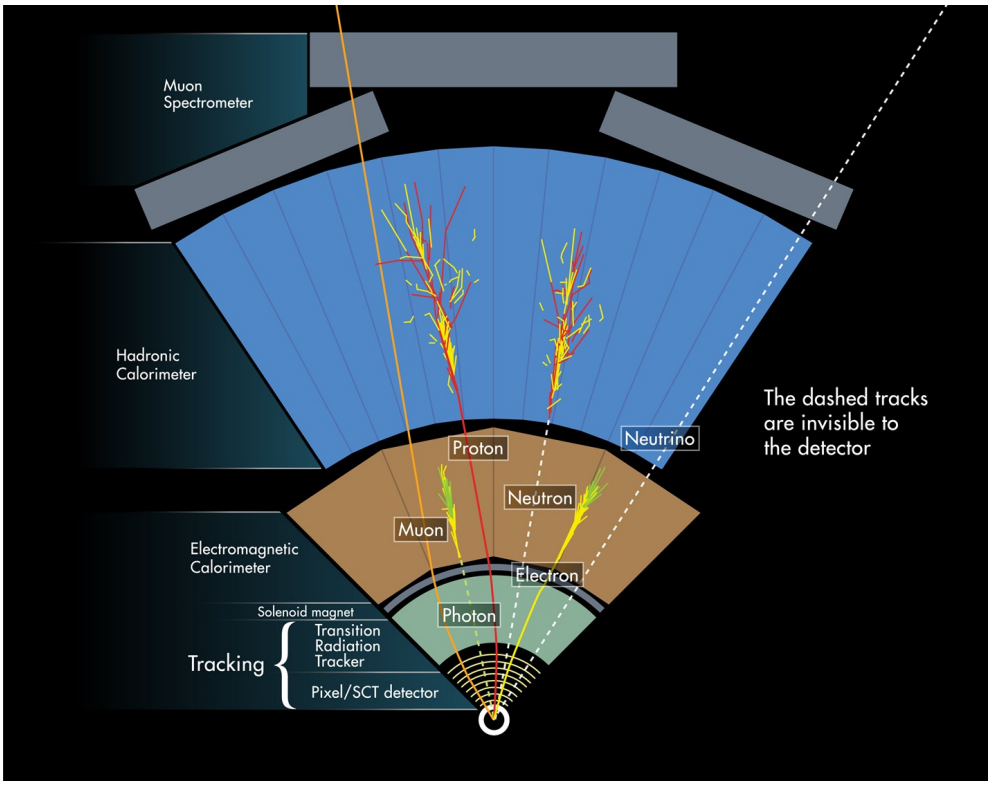


Figure 21: A $(r - \phi)$ view of a slice of the ATLAS detector, showing the typical signature of the physics objects reconstructed by the ATLAS software algorithms obtained by combining tracks and energy clusters. Image courtesy of the ATLAS Experiment.

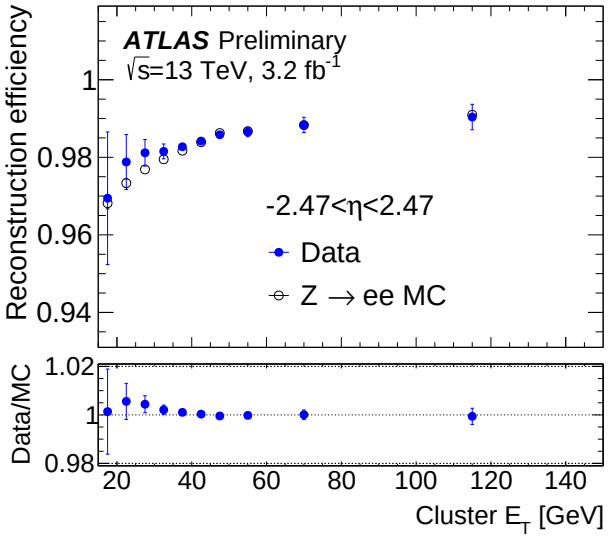
For electrons, clustering proceeds via a sliding window algorithm [82], which searches for an initial cluster seed as the tower[1] with maximum energy within a fixed size window of 3×5 cells of 0.025×0.025 unit in (η, ϕ). Electron track candidates are reconstructed through the steps described in the previous paragraph. However, a dedicated pattern recognition and fit hypothesis are used, to take into account for the typically large energy losses through bremsstrahlung in the ID material. Resulting tracks loosely matched in (η, ϕ) to an EM cluster are re-fitted using a Gaussian Sum Filter [84], to better take into account non-linearities in the bremsstrahlung process. Then, the matching of the re-fitted tracks to the clusters is enforced upon stricter conditions, such as minimal requirements on the number of silicon space-points and the presence of at least one in the innermost pixel layer to reject photons converted in the ID. The electron energy is given by the EM cluster - whose final calibration is based on Monte Carlo simulation [85] - while the η and ϕ coordinates are given by the corresponding track parameters.

Electron are reconstructed only in the $|\eta| < 2.47$ range and for $E_T > 7$ GeV, with an efficiency ranging between 97 to 99 % for $E_T > 15$ GeV in the endcap and barrel, respectively. The reconstructed efficiency as a function of η and E_T is measured in $Z \to ee$ and $J/\psi \to ee$ events [86]. Results are shown in Figure 22. Differences in the efficiencies between data and simulation are found to be of less than $\mathcal{O}(1)\%$, and are applied to simulation as a correction at the analysis stage.

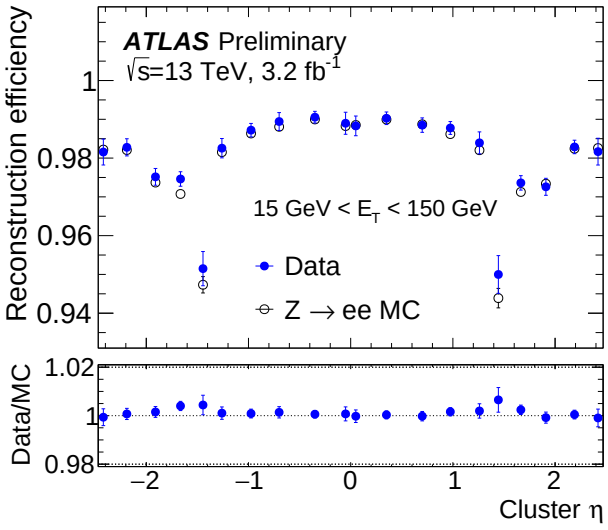
Electron identification (ID) is performed by means of a likelihood-based multivariate discriminant [86]. Input variables for the algorithm include calorimeter shower shapes, likelihood probability based on transition radiation in the TRT, track-cluster matching related quantities, track properties such as total number of measurements per track and number of measurements in the IBL, and variables measuring bremsstrahlung effects. Three identification operating points are defined - denominated *Loose*, *Medium* and *Tight* - ordered in terms of increasing background rejection. Depending on the operating point, the signal and background efficiencies for electron candidates with $E_T = 25$ GeV are in the range 78-90%, 0.3-0.8 %, respectively. Good stability of the algorithm performance against pile-up activity in the event is achieved, as Figure 23 shows.

The $Z \to ee$ and $J/\psi \to ee$ events are used also to measure the electron trigger efficiency, for which corrections to account for differences between data and simulation are extracted as well [86]. Fig. 24 shows the trigger ef-

1 A tower in the EM calorimeter represents the set of cells contained in a 0.025×0.025 (η, ϕ) unit spanning longitudinally over the 3(4) calorimeter layers.



(a) $\varepsilon_{reco}(E_T)$



(b) $\varepsilon_{reco}(\eta)$

Figure 22: Electron reconstruction efficiency in 2015 data and simulated $Z \to ee$ events, as a function of E_T and η [86].

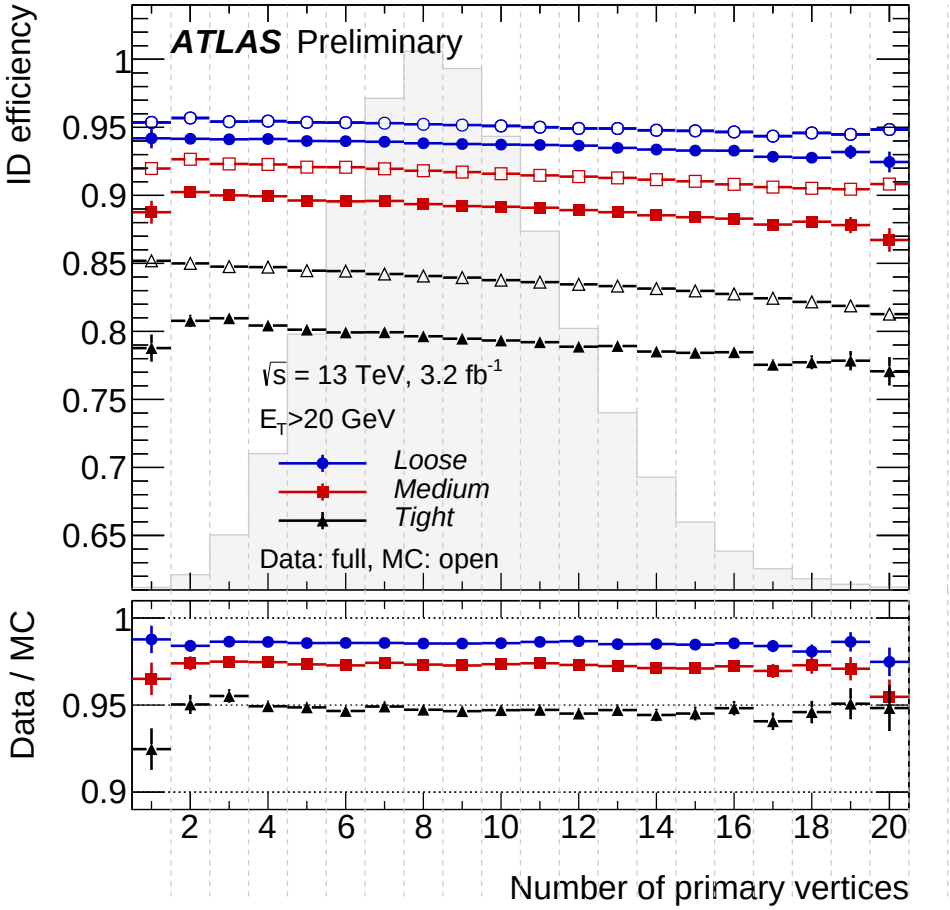


Figure 23: Efficiencies in 2015 data and simulated $Z \to ee$ events for the different electron ID likelihood operating points as a function of the number of reconstructed primary vertices (whose distribution is overlaid in background) [86]. The latter is an alternative measurement to $\langle \mu \rangle$ of the pile-up effect, and is directly sensitive to charged particle activity in the ATLAS detector.

ficiency for the primary single electron trigger used in 2015 [2] as a function of E_T and η, with respect to the Medium identification working point. The trigger effciency reaches a plateau of $\approx 95\%$ for $E_T \gtrsim 28$ GeV, where the difference between data and simulation is within 5%. The largest discrepancy is observed closer to the online E_T selection cut (24 GeV), and originates from different thresholds in the L1 seeds in data and simulation used for the comparison. These differences are corrected for in simulation.

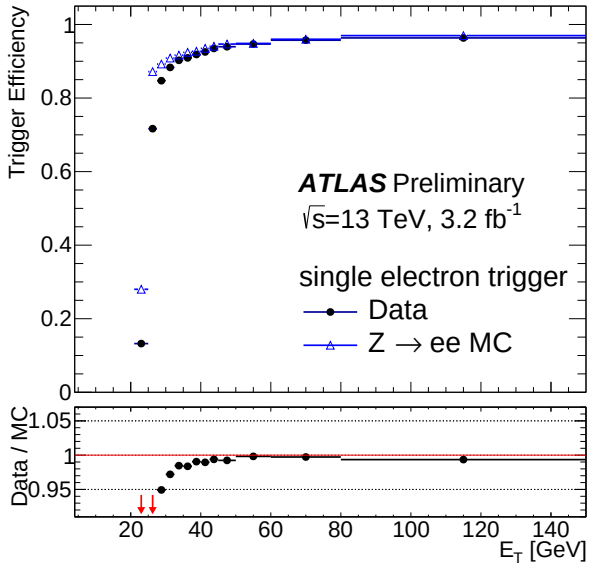
2.3.3 Muons

The reconstruction of muons [87] combines track information from the ID and the MS, as well as calorimetric information. This leads to four categories of muon objects:

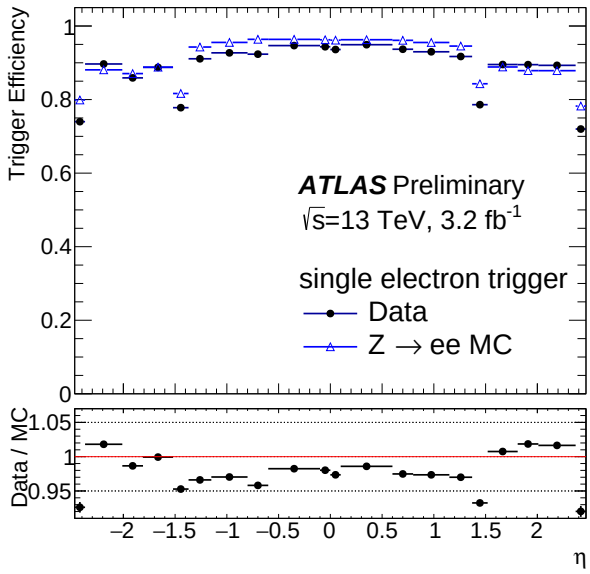
- Combined muons (CB). These muons are obtained from a global fit of hits belonging to tracks reconstructed in the ID and the MS. During the global fit procedure, MS hits may be added to or removed from the track to improve the fit quality.
- Segment tagged muons (ST). This class of muons represents the case when a full ID track is matched only to a muon track segment in the MDT or CSC, either because of low p_T or because they are found in regions of the MS with reduced acceptance.
- Calorimeter tagged muons (CT). High momentum muons from electroweak decays deposit only little energy when traversing the calorimeters, compatible with a minimum ionising particle regime loss. When an ID track matches one such cluster, a CT muon candidate is defined. These muons are characterised by higher fake rate, but allow for acceptance recovery in the $|\eta| < 0.1$ poorly instrumented MS region.
- Extrapolated muons (ME). In order to increase the acceptance for the $2.5 < |\eta| < 2.7$ region not covered by the ID, muons can be reconstructed as a MS-only track extrapolated back to the interaction point with loose compatibility requirements, taking into account the estimated energy loss in the calorimeters.

Overlaps between different muon types are resolved before defining the final muon collection to be used by physics analyses. Muons are reconstructed with a p_T down to 4 GeV. Three main muon quality working points

2 HLT_e24_lhmedium_L1EM20VH



(a) $\varepsilon_{trig}(E_T)$



(b) $\varepsilon_{trig}(\eta)$

Figure 24: Electron trigger efficiency in 2015 data and simulated $Z \to ee$ events for the main single electron trigger chain of 2015, as a function of E_T and η [86].

are defined: *Loose*, *Medium* and *Tight*. They correspond to the inclusion or exclusion of some of the above categories of reconstructed muons in the object definition: Tight muons take into account exclusively CB muons, with additional track quality requirements to suppress muons from in-flight decays, whereas Medium muons also include ME muons for $|\eta| > 2.5$ seeded by at least two track segments. Finally, the Loose muon definition also allows for ST and CT candidates.

Muon reconstruction efficiency in the $|\eta| < 2.5$ region is assessed in a similar way as for electrons, from $Z \to \mu\mu$ and $J/\psi \to \mu\mu$ events [87]. For Loose and Medium muons, a reconstruction efficiency of more than 98 % is achieved, and it is about 95 % on average for Tight muons. Differences between simulation and data of order few percent, mostly localised in the barrel region because of a poorly aligned MDT chamber. Results for the muon reconstruction performance in 2016 are summarised in Fig. 25.

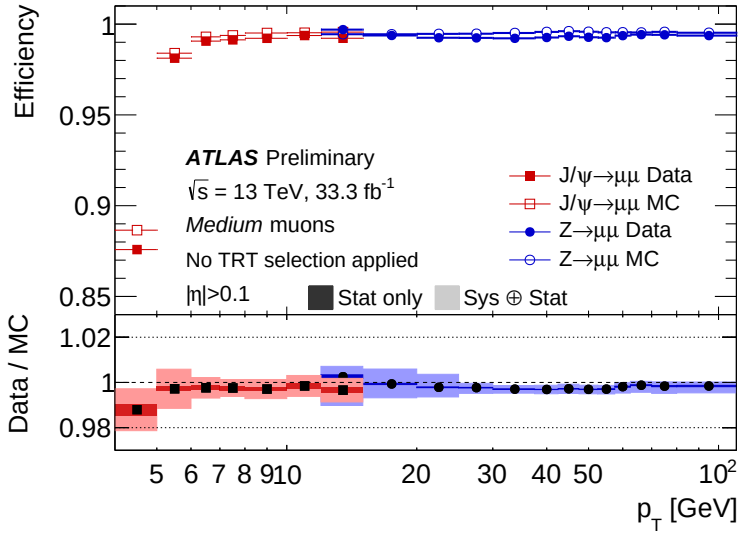
Performance of the muon trigger efficiency [89] is estimated similarly to electrons. Fig 26 shows the efficiency at L1 for the MU20 seed - which requires one muon candidate with $p_T > 20$ GeV to be measured in the MS trigger chambers - and the absolute and relative efficiency for the primary single muon trigger of 2016 at HLT [3], in the barrel ($|\eta| < 1.05$) and end-cap ($1.05 < |\eta| < 2.4$) regions. The L1 efficiency in the barrel is $\approx$ 70% at plateau, and $\approx$ 90% in the end-cap. The worse performance in the barrel is due to the reduced geometrical acceptance of the RPC compared to the TGC system, and to a larger number of local inefficiencies in the RPC chambers. For muon candidates passing L1, the HLT trigger efficiency at plateau is around 100 %. Differences in trigger efficiency between data and simulation are found to be of order few percent.

The muon transverse momentum scale and resolution in simulation is calibrated to collision data as according to Ref. [87]. The relative resolution is found to be within 1.7 % and 2.9 %, and after applying data-driven momentum corrections, the p_T resolution in data and simulation agree to better than 5% for most of the pseudorapidity range.

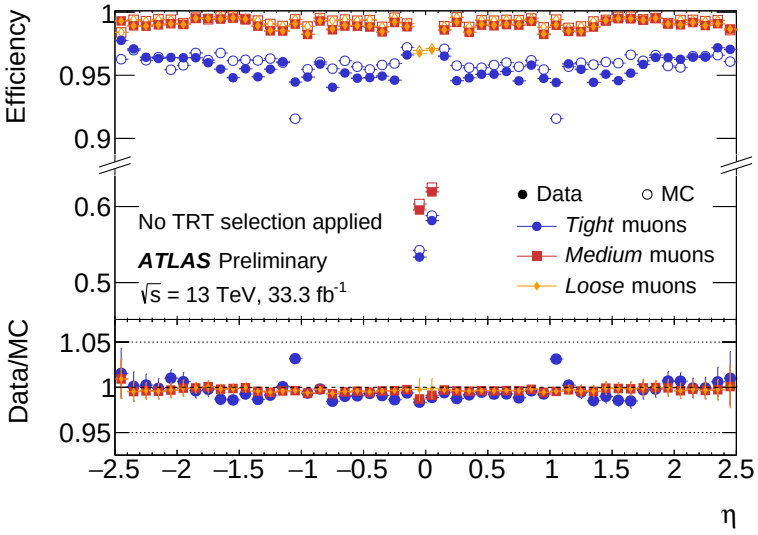
2.3.4 Jets

Quarks and gluons involved in the initial or final states of the hard-scattering generate a *parton shower* (PS) through perturbative radiation of virtual gluons. When the energy transfers in the parton showering process fall down to a scale approximately smaller than the typical hadron size $\Lambda_{QCD} \approx 200$ MeV,

3 `HLT_mu26_ivarmedium OR HLT_mu50`

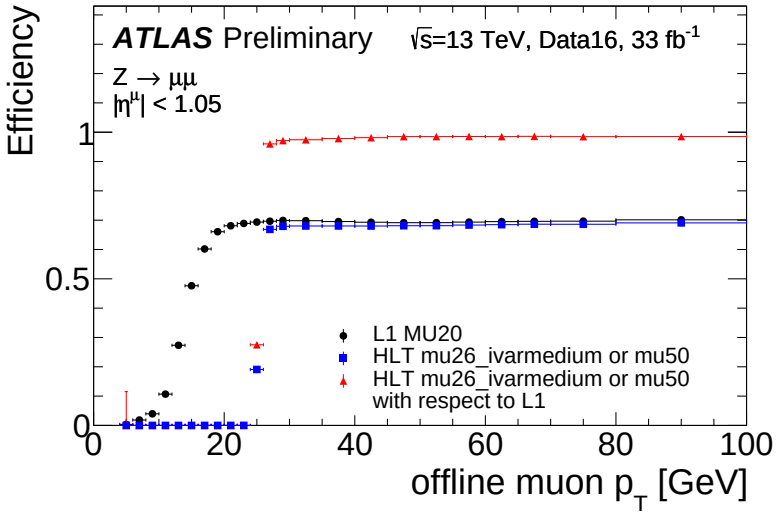


(a) $\varepsilon_{reco}(p_T)$

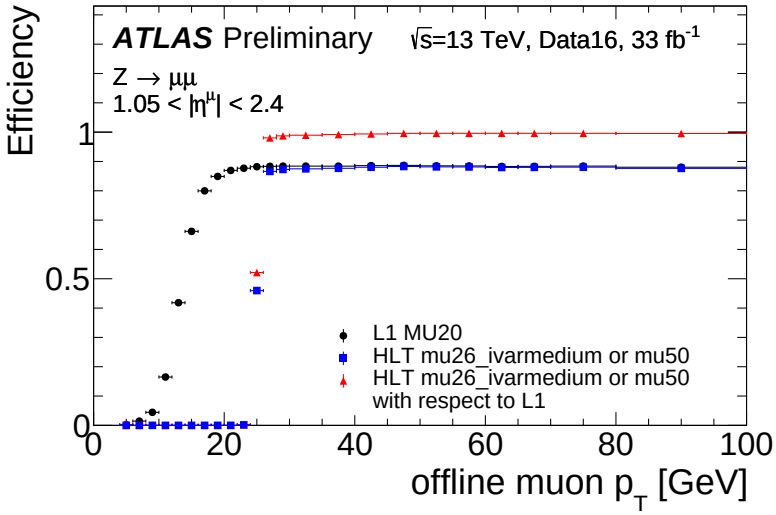


(b) $\varepsilon_{reco}(\eta)$

Figure 25: Muon reconstruction efficiency in 2015, 2016 data and simulated $Z \to \mu\mu$, $J/\psi \to \mu\mu$ events, as a function of p_T (for *Medium* muons only, and inclusively for $|\eta| > 0.1$) and η [88]. The recovery in efficiency in $|\eta| < 0.1$ for the Loose working point is due to the inclusion of calorimeter-tagged muons in the object definition.



(a) $\varepsilon_{trig}(p_T)$, $|\eta| < 1.05$



(b) $\varepsilon_{trig}(p_T)$, $1.05 < |\eta| < 2.4$

Figure 26: Muon trigger efficiency in 2015, 2016 data events for the L1 MU20 seed and for main single muon HLT trigger chain of 2016, as a function of p_T, for the barrrel and end-cap regions, respectively [89]. The online selection at HLT includes a requirement of low activity around the muon candidate (*isolation*).

the QCD perturbative regime breaks down and the *hadronisation* process begins, in which colourless, bound states of partons are created. Hadronisation is a consequence of colour confinement, an empirical evidence of the absence of colour-charged particles in a free state [14].

At detector level, the result of hadronisation is the creation of a narrow spray of hadrons and their decay products known as a *jet*, whose energy and direction is related to the one of the originating parton. Inputs to the ATLAS jet reconstruction algorithms are topologically connected clusters (*topoclusters*), defined by neighbouring cells grouped together as long as the signal in the cells is significant compared to noise [90]. Jets considered in the $t\bar{t}H$ search are built from topoclusters calibrated at the electromagnetic energy scale (EM), which correctly measures the energy deposited by electromagnetic showers in the calorimeter [91].

Jets are reconstructed from topoclusters using the anti-k_T algorithm [92] with a distance parameter of $R = 0.4$. The energy scale and resolution of the reconstructed jets with $p_T > 25$ GeV is eventually calibrated to the truth parton constituents energy with simulation-based techniques validated with data-driven ones [93, 94].

In order to suppress the contamination of jets originating from pile-up interactions for $p_T < 60$ GeV and $|\eta| < 2.4$, an algorithm based on association of jet tracks to the hard-scattering vertex known as Jet Vertex Tagger (JVT) [95] is used. For the algorithm tuning chosen in the $t\bar{t}H$ analysis, the average association efficiency per jet is 92%. The JVT efficiency is calibrated in data as a function of jet p_T using $Z(\rightarrow \mu\mu)$ + jets events, resulting in correction factors of order $1 - 5\%$ to be applied on simulated events.

The top quark, due to its very high mass, does not hadronise before decaying into b quarks. Jets originating from b quarks typically contains heavy flavoured hadrons like $B^{\pm}$ mesons, which travel on a macroscopically detectable distance before decaying ($c\langle\tau_{B^{\pm}}\rangle \approx 0.5$ mm). This property is a powerful handle to *tag* b-originating jets at collider experiments. These jets represent a key signature for the $t\bar{t}H$ search. At ATLAS, a multivariate-analysis tagging algorithm denominated MV2 [96, 97] is used, based on the combination of several discriminants which exploit information on the impact parameter of tracks associated to the jet, the reconstructed secondary vertices in the event, and the topology of the electroweak b, c hadron decays. The specific algorithm used in the $t\bar{t}H$ to multileptons analysis - denoted as MV2c10 - is trained on $t\bar{t}$ events assuming background is a mixture of 93 % light-flavoured jets and 7 % c-flavoured jets. The selected working point corresponds to a b-jet efficiency of 70 % and a c-jet (light jet) rejection fac-

tor of 12 (381). Calibration of the b-tagging efficiency to data is performed combining results from several methods, similarly to what previously done in the Run 1 analysis [98], and is measured as a function of jet p_T, η, and other physics observables quantities that are sensitive to close-by jet activity. Corrections to the simulation are found to be within 5 %, with global uncertainties of maximum $\approx 10\%$.

2.3.5 Hadronic tau leptons

Because of its large mass ($m_\tau \approx 1.77$ GeV), tau leptons decay in $\approx 36\%$ of the cases to electrons or muons, and in the other $\approx 64\%$ of the times into a combination of neutral and charged pions or - more rarely - kaons, with two or one additional neutrino. In the latter case, the vast majority of the decays is into either one charged pion (*1-prong*), or three charged pions (*3-prong*), plus additional neutral hadrons. The very short mean decay path of $c\langle\tau\rangle \approx 87$ μm implies most decays happen inside the LHC beam pipe. Thus, only its visible decay products can be exploited for tau reconstruction and identification. While for leptonic decay no attempts are made in ATLAS to distinguish the tau decay product from a prompt electron or muon, in the hadronic decays the peculiar 1-3 prong signature and the relatively collimated clustering of tracks and energy deposits allow for a dedicated hadronic tau (τ_{had}) reconstruction. These features are exploited in tau identification algorithms to increase rejection against QCD jet background. Despite being a signature of several $t\bar{t}H$ multileptonic categories, the focus of this work will be solely on the channels where hadronic taus are vetoed. Details about tau reconstruction and identification in ATLAS can be found in Ref. [99].

2.3.6 Missing transverse energy

Momentum conservation in the transverse plane can be used as a constraint at hadron colliders to define the missing transverse momentum vector:

$$\mathbf{E}_T^{miss} = \mathbf{E}_x^{miss} + \mathbf{E}_y^{miss} = -\sum \mathbf{E}_T^{vis}, \tag{20}$$

which accounts for all undetected particles such as neutrinos or new weakly interacting particles produced in the collision, as well as for particles flying

through inactive regions of the detector. The magnitude of missing transverse momentum is indicated as *missing transverse energy* E_T^{miss}:

$$E_T^{miss} = \sqrt{(E_x^{miss})^2 + (E_y^{miss})^2} \tag{21}$$

$$E_{x(y)}^{miss} = -\sum E_{x(y)}^{vis}. \tag{22}$$

The ATLAS E_T^{miss}reconstruction algorithm [100] considers all the reconstructed and calibrated objects as inputs to the right side of Eq. (22):

$$\begin{aligned} E_{x(y)}^{miss} = E_{x(y)}^{miss,e} + E_{x(y)}^{miss,\gamma} + E_{x(y)}^{miss,\tau} + E_{x(y)}^{miss,jets} + \\ + E_{x(y)}^{miss,TST} + E_{x(y)}^{miss,\mu}, \end{aligned} \tag{23}$$

where each term is the negative sum of the x,y transverse momentum components of physics objects. Whenever possible, this is measured from the reconstructed tracks in order to achieve better stability against pile-up, otherwise calorimetric transverse momentum measurement is used. The "TrackSoftTerm" (TST) contribution accounts for energy of tracks which are not associated to any of the objects definition, such as activity outside the hard scattering collision and soft radiation. Performance of the E_T^{miss}reconstruction are documented in detail in [100].

2.4 EVENT SIMULATION

In the context of a physics search at colliders, the observed data from the high energy collisions must be eventually compared to some hypothesised model which includes the signal of interest, as well as the background processes. Often the available collision data statistics is not large enough to predict all the backgrounds with sufficient precision. This is particularly true when looking at narrow regions of the collision final state phase space as the ones analysers are typically interested in. Given that many of the backgrounds are well described by the SM, simulating them represents an effective way to enhance the event sample size.

Simulation at ATLAS is a multi-tiered procedure which makes use of Monte-Carlo (MC) event generator software libraries to simulate the final state of each process in full detail, down to the level of individual stable particles [101]. Each step in the chain is carried out by specific programs, interfaced to each other to eventually produce datasets for each process of

interest which share the same format and physics objects content as the collision datasets.

- The first stage consists of integrating the probability distribution for a process of interest over the relevant phase space (kinematics, spin). The defintion of the probability density embeds the product of the parton-level matrix element - calculated at some perturbative order in QCD and QED - and the input PDFs, as from Eq. (2).

- Dedicated programs are used to simulate the radiation of gluons by partons in the initial and final state via a parton showering (PS) approach. This uses perturbative QCD to model the particle cascade, typically up to next-to-leading-logarithm accuracy (NLL), until the process gets stopped at a non-perturbative cut-off energy scale of order $\Lambda_{QCD} \approx 200$ MeV. Similarly to PS, QED radiation of photons is also simulated, employing an equivalent mathematical framework.

- The formation of hadrons, which cannot be modelled by perturbative QCD, is a process-independent step described by phenomenological models based on experimental inputs, as they are not currently understood from first principles. The followed approach can differ substantially between simulation software implementations.

- The *underlying event* (UE) is simulated as well. It represents the hadronisation of the collided proton remnants, as well as all spectator interactions outside the hard scattering event of interest. These processes can belong to both perturbative and non-perturbative QCD regimes. The minimum bias events leading to the pile-up effect are then also overlaid to the main hard scattering event. As PS, hadronisation and UE simulations imply either approximations to higher order perturbative QCD calculations (PS), or phenomenological descriptions for the soft scale processes (hadronisation and UE), each model features a set of free parameters which are optimised in a procedure known as *tuning* to produce a suitable description of physical obsevrables.

- The final part of the simulation process is about modelling the interactions with the detector of all the stable particles generated in the previous steps. This is carried out by the GEANT4 software [102], interfaced to a full description of the ATLAS detector material budget and geometry [103]. Being this the most computing resource-consuming step of the whole chain, faster detector simulation techniques are sometimes also employed [104].

- Detector hits are then reproduced and digitised. Finally, fully simulated events are processed through the same particle reconstruction algorithms as for collision data, and persistified in the same file format for analysis. The full trigger decision based on the chosen trigger menu is also simulated.

As previously described, corrections derived in collision data are applied at the analysis stage to adjust efficiencies for object reconstruction, identification and trigger. Data-driven corrections to energy scales and calibrations are applied as well. The simulated pile-up modelling strongly depends on the hypothesised collision conditions; since a full simulation campaign is produced before or during a given data taking period, by construction only a best-guess of the data pile-up conditions can be put into the simulation software. Consequently, there is the additional need at the analysis level to correct the simulated pile-up conditions to match what is actually found in the collision data.

3 TRACKING IN DENSE ENVIRONMENTS AT ATLAS

After the increase to $\sqrt{s} = 13$ TeV in the centre-of-mass energy of the LHC collisions in Run 2, physics signatures characterised by high p_T jets and hadronic tau leptons become more interesting. Such topologies are of particular importance for numerous physics analyses in ATLAS, such as searches for exotic heavy resonances decaying to boosted particles [105, 106]. One of the challenges posed by these final states is the presence of several highly collimated charged particles in the core of jets and hadronic taus, with relative distances comparable to the sensor granularity of the Inner Detector. This effect can impair the b-jet and tau identification, as well as the jet energy calibration and jet mass reconstruction. The work presented in this chapter is dedicated to improving the ATLAS tracking algorithm in order to enhance the track reconstruction performance in environments with high density of tracks. A review of this topic - including a substantial contribution from the author - is covered in depth in Ref. [107, 108].

3.1 THE ATLAS TRACKING ALGORITHM

A basic outline of ATLAS ID tracking algorithm [78, 79] has been briefly introduced in Chapter 2. In this section, a more thorough description of its different steps will be given, prior to the changes I applied with respect to Run 1 algorithm. Focus will be put only on the tracking in the silicon detectors (Pixel and SCT). In fact, due to their proximity to the interaction point, they mostly determine the performance for the track topology of interest, where high density of tracks is expected.

The logic of the track reconstruction in ATLAS follows an inside-out, staged pattern recognition approach. It starts from the reconstruction of track candidates, seeded by a combination of measurements in the high granularity silicon-based subdetectors. This is followed by a fit of the candidates' trajectory, and an iterative ambiguity solving procedure, where bad quality tracks and tracks from incorrect assignment of measurements are rejected. The resulting tracks are then extended to incorporate consistent TRT drift circles, are re-fitted and selected. This leads to the final collection of tracks used as input to the particle reconstruction.

3.1.1 Clustering in the silicon detectors and space-point definition

Track reconstruction begins with the conversion of the raw data from the Pixel and SCT detectors into three-dimensional measurements, or *space-points*.

The Pixel system is able to detect the amount of charge deposited in a sensor by measuring the time the signal pulse exceeds a given threshold, using the Time-over-Threshold (ToT) technique [109]. Often a particle leaves a charge deposit spread over more than one single pixel. Pixels within a sensor which collected charge above the read-out threshold of 3.5 ke and share an edge or corner are grouped together via a connected component analysis to form pixel *clusters*. Since the total deposited charge is proportional to the path length in the sensor and therefore to the incidence angle of the particle, the coordinate of the particle intersection in the cluster can be determined from the cluster composition using a charge interpolation technique [110]. Similarly, neighbouring SCT strips which recorded the passage of a particle within a sensor define a SCT cluster. Unlike the Pixel detector, the SCT sensors are connected to binary readout chips, which do not provide information on the amount of charge collected by each strip.

Whilst a pixel cluster corresponds to a single measurement, a SCT 3D space-point must be defined by the combination of two clusters on both sides of a module in a layer. In fact, the precise measurement on a single SCT sensor can only be given orthogonally to the strip direction. However, owing to the 40 mrad stereo angle displacement between silicon sensors on each side of an SCT module, a measurement in the other local direction can be provided.

Often in dense environments like the core of high p_T jets the distance between charged particles is of the same order of magnitude as the pixel granularity, which is typically 50 μm (400 μm) on the transverse (longitudinal) direction. This is illustrated in Figure 27. In such cases, charge deposits from multiple sources may end up being grouped together in a single, *merged* cluster, an effect pictorially represented in Figure 28.

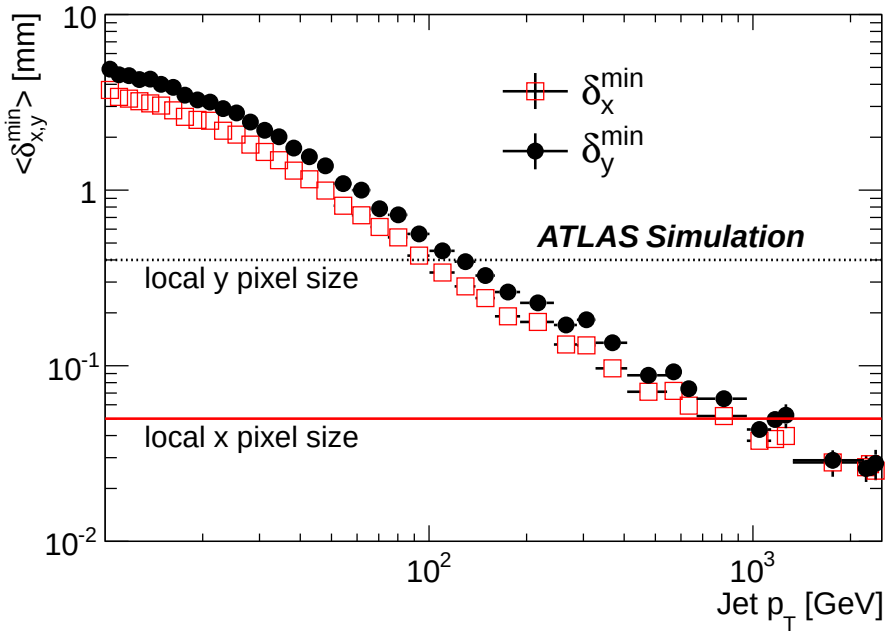


Figure 27: Average distance between the two closest charged particles in second layer of the Pixel detector as a function of jet p_T, as taken from simulation [111]. The distances on the x (transverse) and y (longitudinal) projections of the local 2D coordinate system of the pixel module are shown in red and black, respectively. The picture also shows on the vertical axis the typical transverse and longitudinal pixel size.

3.1.2 Seeding and combinatorial track finding

From the silicon detector space-point collection, a set of track seeds is formed by taking combinations of three measurements. A first fit to a perfect

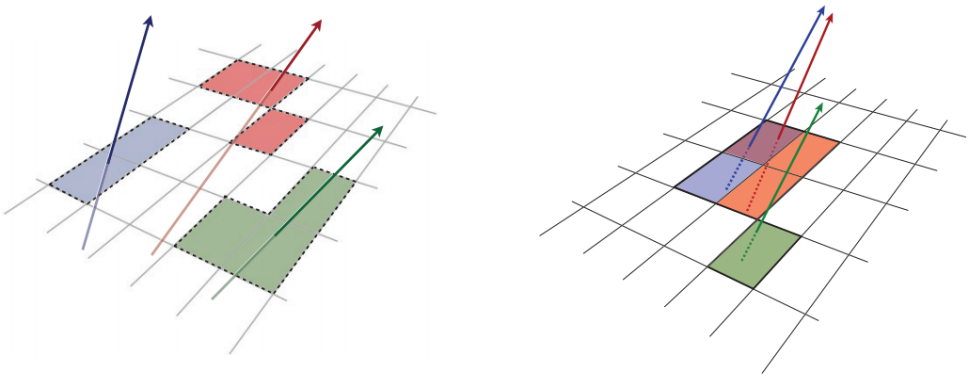


Figure 28: On the left side, a representation of three resolved clusters in a pixel sensor, corresponding to the passage of three charged particles (indicated by the arrows). On the right, a merged pixel cluster resulting from charge deposition of two close-by particles separated by a distance comparable to the pixel granularity ($d \sim \mathcal{O}(\sqrt{50^2 + 400^2})$ μm), together with a single, resolved cluster.

helical trajectory in a uniform magnetic field is performed, in order to get a raw estimate of the track seed properties. After the fit, a track is completely defined by 5 parameters, according to the track perigee parametrisation [1] adopted in ATLAS:

- The polar and azimuthal angles (θ, ϕ).
- The charge divided by the track momentum q/p.
- The transverse (d_0) and longitudinal (z_0) distance of closest approach of the trajectory to the centre of the interaction region, known as *impact parameters*.

Seeds are subsequently selected based on loose requirements on transverse momentum, impact parameter, number of associated measurements, and number of missing measurements where the fitted trajectory would predict one (referred to as *holes*). At this stage, the purity of seeds - defined as the probability for a seed to lead to a good quality track - is already quite high, owing to the preliminary seed selection. Depending on the seed space-point composition (pixel and/or SCT), the achieved purity in simulation ranges from 86% (70%) for SCT-only seeds in absence of pile-up (with a pile-up of $\langle\mu\rangle$=40), to 79% (39%) for pixel-only seeds [112]. This level of

1 The perigee parameters define a track at its point of closest approach (the *perigee*) to a given coordinate system. Typically, the centre of the interaction region (referred to as the *beam spot*) is used as reference point.

purity still leads eventually to a $> 99\%$ track reconstruction efficiency for primary muons [113].

Track candidates are then obtained by performing a search for further space-points within a road via a combinatorial Kalman filter [114], following the direction of the seed. After this step, several track candidates can be associated to a single seed, and a cluster on a track can belong to multiple track candidates, in which case it is referred to as a *shared* cluster.

3.1.3 Ambiguity solving

After the seed selection and combinatorial pattern recognition stage, the majority of *fake* track candidates resulting from random combinations of measurements from multiple particles are successfully rejected. However, tracks defined by assignment of spurious measurements or track duplicates are still present.

In order to further suppress fakes, an *ambiguity solving* process is implemented. In this process, track candidates are ranked based on their likelihood to describe the actual trajectories of particles produced in the event. Each track is firstly assigned a score based on global features of the track itself, as well as on the quality of the track fit, and it is put in a track candidate list sorted in order of descending score. The score is increased when many measurements are assigned to the track candidates - with a larger weight generally assigned to the ones belonging to the more precise inner layers of the detector - and decreased in the presence of holes and a poor χ^2 of the fit. Also, a lower score is assigned to lower p_T candidates, which are more likely to include spurious measurements. If the score of a track is too low, the candidate is immediately rejected. In this context, a cluster associated to a track is also referred to as a *hit*. Figure 29 shows an example of three track candidates considered in the ambiguity solving stage, as they appear on an $(r - \phi)$ projection of the SCT barrel.

After the first ranking of the track collection, the ambiguity solving algorithm starts processing the ordered set applying strict selection criteria, removing rejected tracks from the list for the subsequent iterations. A candidate is rejected if it fails any of the the following requirements:

- A minimum of 7 clusters from combination of the Pixel and SCT.
- In a given ID layer, a maximum of one shared pixel cluster, or two shared SCT clusters.
- A maximum of two holes, and no more than one in the Pixel.

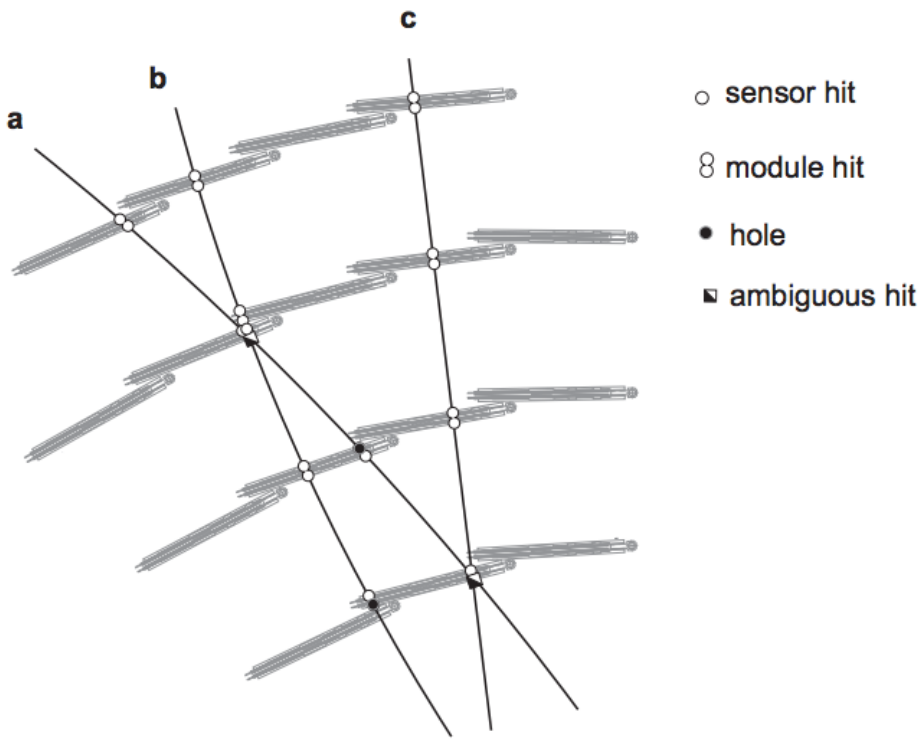


Figure 29: An example of three track candidates from the combinatorial Kalman filter at the ambiguity solving stage, in the SCT barrel layers [78]. In this example, a" module hit" on track, i.e., a 3D measurement in the SCT silicon detector, will be scored relatively higher than two "sensor hits" on a single side of an SCT module. Hits in a overlap region between modules, as in the case for track *b*, are highly scored, whereas holes on track give a penalty in the ranking.

- $p_T > 400$ MeV.
- $d_0 < 2.0$ mm.
- $z_0 \times sin\theta < 3.0$ mm.

If a candidate passes the selection, a full resolution fit is performed, taking into account also material effects such as multiple scattering, the score is re-evaluated and the track is processed again. The choice of delaying the fit up to this stage minimises the average number of calls to the fitter algorithm per event, reducing the CPU resources consumption. The track is eventually accepted and taken out of the list if it passes also the second iteration. As a result, full tracks will be preferred over small track segments.

At this step of the reconstruction chain, tracks may have several shared measurements. This is in general a feature to be avoided, as it degrades the track parameters resolution. Nevertheless, tracks in dense environments tend to share many measurements by construction. After the selection, the ambiguity solving algorithm checks the hit content of the track under exam, and can change the track cluster composition in case shared hits are involved. Specifically, a shared cluster is removed from a track candidate if more than one previously accepted track already compete for it, or if it would cause an accepted track to fail that cap for the same cluster. As a result, tracks sharing too many measurements will not make it through the ambiguity solving. A simplified scheme of the logic of the ambiguity solving algorithm is presented in Figure 30.

It is therefore important to find an optimal treatment of shared hits on a track to improve the reconstruction efficiency for close-by tracks, without increasing the fake rate more than achieved during Run 1, on average around 5%. The underlying guide principle is that a shared cluster should not actually be penalised if it is truly resulting from two collimated charged particles. In other words, merged clusters must be efficiently and correctly identified.

3.1.4 Identification of merged clusters

A tool consisting of a set of three artificial neural networks [111] has been developed and employed during the Run 1 data taking at the clustering stage to identify merged clusters in the Pixel detector, as well as to precisely determine the cluster position and its uncertainty. One neural network is trained to assess the compatibility of a cluster with the hypothesis of being produced by 1, 2, or 3 charged particles. Inputs to the algorithm include the cluster size and information on the charge deposited per pixel via the ToT,

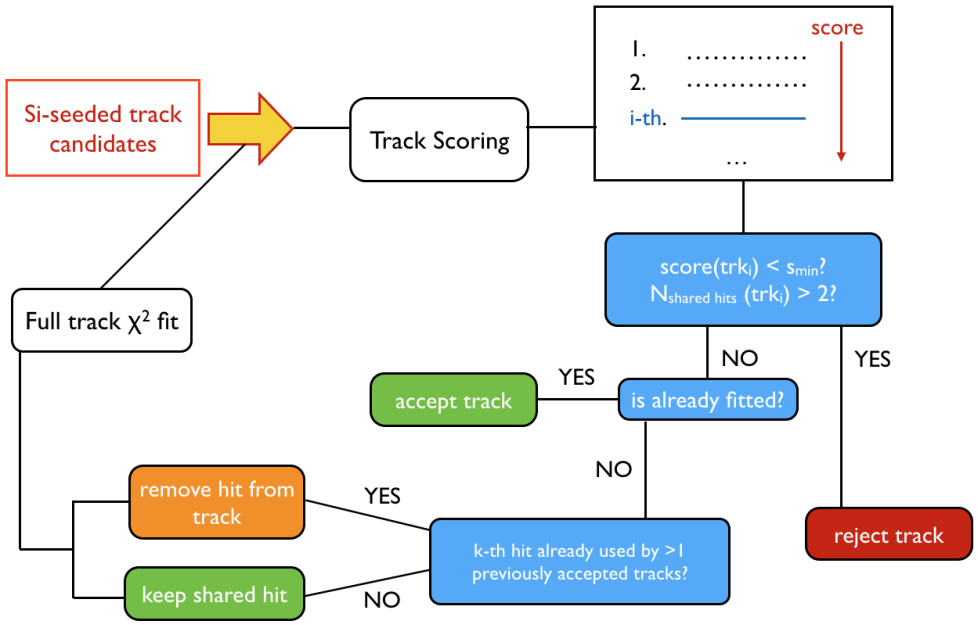


Figure 30: A simplified flowchart of the ambiguity solving algorithm. Starting with track candidates from the combinatorial search, each is assigned a score based on cluster content of the track and quality of the fit, and listed in descending order of score. Tracks with too low score, or not satisfying requirements on the number of hits, including shared ones, are rejected and removed from the list. The algorithm then checks shared hits on track against previously accepted candidates, and can remove them from the candidate cluster content under certain conditions. A track is then fully fitted, scored again, and reprocessed by the algorithm. If it passes the selection, then it is accepted and removed from the set. The algorithm ends when no more candidates are left in the ordered list.

and the incidence angle - estimated roughly at the clustering stage by joining the cluster centroid with respect to the beam spot position with a straight line. The cluster position is estimated using a second set of neural networks, and third set of neural networks is used to estimate the position residuals.

During Run 1, clusters identified as merged by the neural networks were subsequently *split* before being considered by the combinatorial road finder. The usage of this algorithm at the clustering step leads to a reduction in the average number of total shared measurements per event of about a factor three, and the higher precision of the cluster position estimate achieved generally improves the resolution of the track parameters.

An example of the splitting algorithm is shown in Figure 31. This represents the case when a merged cluster produced by two particles is successfully split by the neural network into two separate clusters, whose position and error are also estimated via the additional set of networks. Due to the intrinsically random nature of the material interactions in the ID components, the neural networks cannot achieve perfect results. In particular, the emission of energetic δ-ray electrons can lead to larger clusters with more charge deposited than expected from a single particle.

Performance of the splitting algorithm is illustrated in Figure 32. It displays the fraction of incorrectly split clusters associated to a single particle against the inefficiency in splitting clusters associated to two particles, using truth information. The chosen working point of the algorithm for Run 1 is highlighted with a star. The performance is particularly sensitive to the precise determination of the angle of incidence. In fact, when exploiting the more refined measurement of the trajectory from the fit after the combinatorial track finding, a reduction of about 60% of the fraction of non-split two-particle clusters is obtained, for the same fraction of split one-particle clusters. This corresponds to an increase of $\approx 15\%$ in efficiency for identifying merged clusters. However, since in Run 1 the neural network is applied at clustering stage and the Kalman filter is a highly CPU-intensive process, no attempt was made at re-running the cluster splitting algorithm with improved information and performing the road search again.

As a consequence, many nearby tracks in dense environments remain lost at the ambiguity stage because of many non-identified merged clusters being shared by multiple candidates. For instance, the efficiency for true 3-prong hadronic tau decays to be reconstructed with 2 charged tracks only [105] dominates for τ_{had} candidates with $p_T \gtrsim 500$ GeV, as illustrated in Figure 33. This called for an overhaul of the strategy of application of the neural network decision.

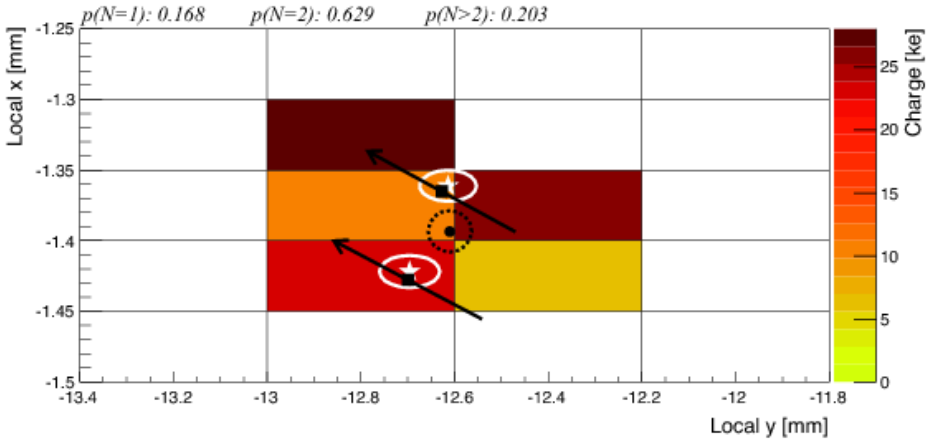


Figure 31: Example of a merged pixel cluster from a pair of collimated charged particles as identified by the NN clustering algorithm in simulation [111]. The true path of particles through the module is represented by the black arrows, with black squares indicating the true intersection with the module plane. The black dot is the original cluster centroid, and the white stars represent the split clusters positions as estimated by the algorithm, with the surrounding circles being the estimated uncertainties.

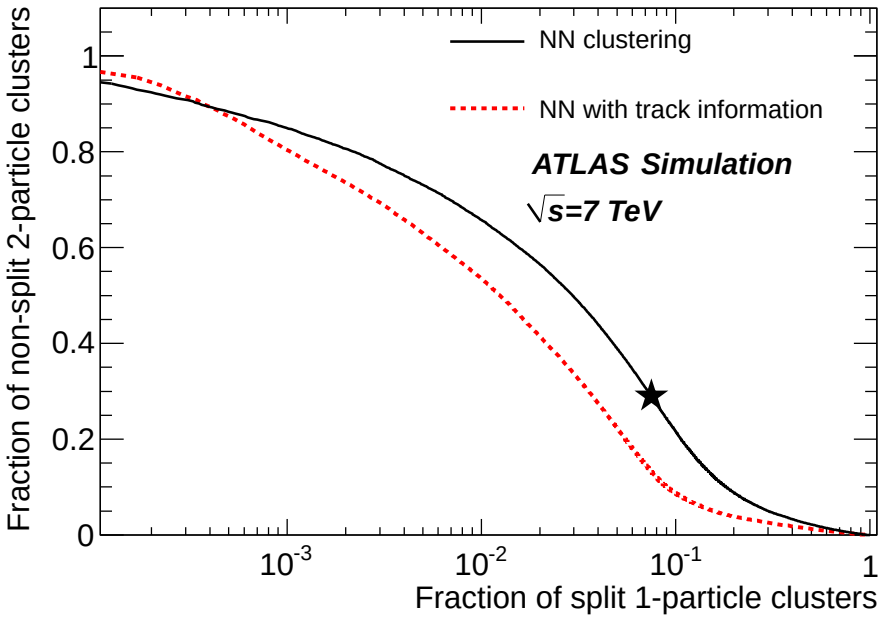


Figure 32: Fraction of non-split two-particle clusters versus the fraction of incorrectly split one-particle clusters in simulation [111]. The chosen working point of the NN clustering for Run 1 - indicated by the star on the black line - correctly splits about 71% of the clusters generated by two particles. Besides, 7.5% of the clusters that arise only from one particle are incorrectly split. The red line represents the improvement in performance when including more refined track incidence angle information as input to the network.

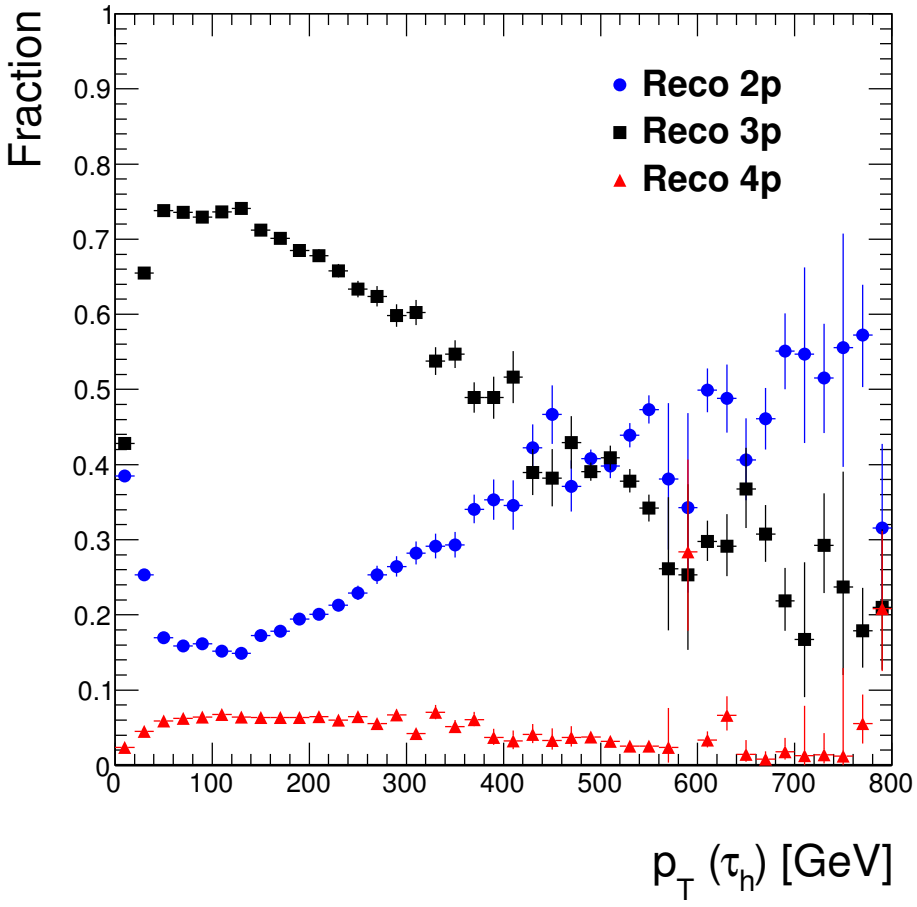


Figure 33: Efficiency for true 3-prong hadronic tau decays to be reconstructed with 2,3,4 charged tracks in simulated $Z' \rightarrow \tau_{had}\tau_{had}$ ($m_{Z'} = 1\text{TeV}$) events [105]. The decrease in 3-prong tau reconstruction efficiency is effectively due to miscounting the number of tracks at high p_T.

3.2 IMPROVEMENTS TO THE TRACKING ALGORITHM

3.2.1 Changes at the ambiguity solving stage

Based on the previous observations, the first change I implemented consists in delaying the usage of the neural networks at the ambiguity solving stage, and only for clusters that are shared. Such clusters are no longer split; on the other hand, their likelihood of being merged as from the network output is used to decide whether they can be kept without penalty in the cluster content of the track, or treated similarly to what done in Run 1 for generic shared clusters. The cluster position and uncertainties are also re-defined by the networks, and used in the full resolution track fit during the algorithm iteration.

This change has a two-fold advantage. Firstly, it allows the exploitation of the refined track information from the fit to improve the network performance. Second, having less clusters reduces the number of seeds by approximately 10%, thus reducing the CPU time.

Additionally, I revisited the logic itself of the ambiguity solving to better cope with the topology of tracks in dense environments. The key observation is that the distance between collimated tracks decreases by moving inwards through the layers of the inner detector. Therefore, if a cluster on a track at a given layer is identified as merged, the cluster on the previous layer is expected to be merged as well. To take this into account, if the same track candidates compete for clusters on two consecutive pixel layers and the cluster on the outer layer is marked as merged by the neural network, the one on the inner layer will also be considered merged and not penalised, regardless of the score of the network for that particular cluster. Correlating information of measurements on multiple layers is beneficial, as it recovers a part of the neural network inefficiency at splitting 2-particle clusters. The reasoning for this change is sketched in Figure 34.

In order to maximise the tracking efficiency, the neural networks working points have been re-optimised to be looser with respect to Run 1. This results in an overall larger number of shared clusters. In order to keep the track fake rate at an acceptable level, a minimum number of hits is required to allow a track having shared hits. Due to the inclusion of the IBL in Run 2, the minimal hit content is set to 9 hits, compared to 8 in Run 1. On top of this, tracks are allowed to have merged clusters only when satisfying a minimal p_T cut of 1 GeV, and a minimum number of four unique hits in the SCT. These selections prevent short fake track segments from being accepted

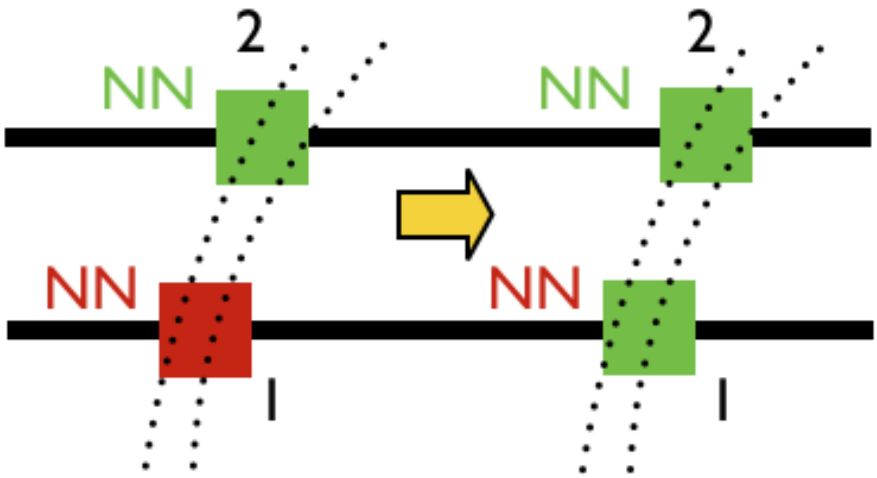


Figure 34: Schematic representation of the change in the ambiguity solving logic for the dense environments case, from the Run 1 setup (left) to the new setup (right). Two close-by tracks are competing for clusters on consecutive pixel layers. The pixel cluster in the outermost layer ("2") is flagged as merged by the neural network (green "NN" flag), and it is therefore not penalised as "shared" cluster. The preceding cluster ("1") on the layer inwards is incorrectly assigned as one-particle-like by the network (red "NN" flag). However, in the revisited logic that cluster will not be penalised as shared in the ambiguity solving logic since it is compatible with a dense-track topology.

when they are incorrectly allowed to share clusters on the inner layers. In the SCT, no simple methods to identify merged clusters are available, as no deposited charge information can be exploited by multivariate discriminators as in the Pixel detector. Therefore, a limit of maximum 2 shared SCT clusters is applied.

3.3 PERFORMANCE OF TRACKING IN DENSE ENVIRONMENTS

In order to quantitatively evaluate the changes introduced in the tracking software, a set of simulated samples with large statistics of events characterised by single particle decays to collimated charged daughters have been generated. Specifically:

- A sample of $\rho \to \pi^+\pi^-$, to check the cleanest two-track topology.
- A sample of $\tau^\pm \to \pi^+\pi^-\pi^\pm\nu_\tau$, to reproduce a more realistic three-track topology.
- A sample of B^+ decaying to an inclusive multi-particle final state, to check the performance in the case of a long-lifetime parent particle.

For all of them, the p_T spectrum of the parent is uniformly distributed between 10 GeV and 1 TeV, and $|\eta| < 1.0$. The samples are processed through a full simulation of the ATLAS detector response [103], based on the GEANT4 framework [102].

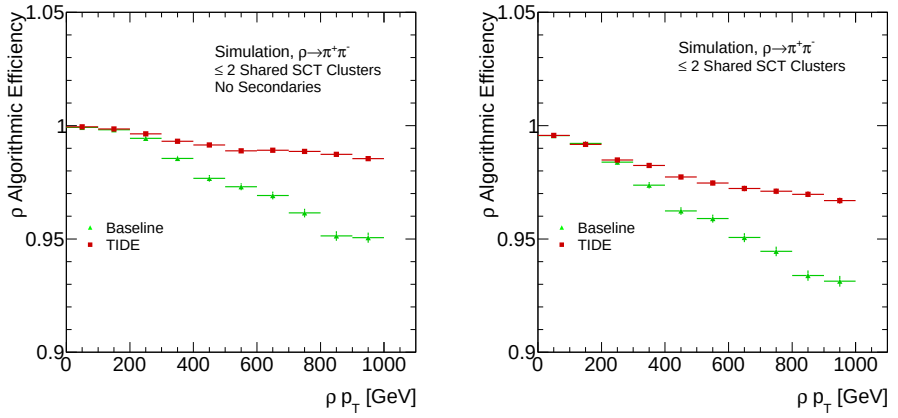
As introduced earlier, fake tracks can either originate from random combinations of clusters during the seeding, or from large mis-allocation of clusters from other particles to a given track candidate. In simulation, track candidates are classified by means of a score determined from truth information, by counting the fraction of weighted measurements on track associated to the same simulated particle. The weights are assigned according to which part of the Inner Detector each measurement belongs to, and reflect the average number of expected measurements in each subdetector: highest weights are given to Pixel measurements, lowest to TRT ones. Fake tracks are defined as having a truth-based score of less than 70%, and are excluded to prevent bias on the performance studies.

3.3.1 Reconstruction efficiency of charged daughters in single particle decays

The efficiency for reconstructing all primary tracks of a parent particle (ρ, τ) decay products (indicated in the plots as "algorithmic" efficiency) as a function of the parent p_T is shown in Figure 35. A comparison is made between the simplified case where no secondary tracks from material interactions are simulated, and the more realistic one where such effects are included. It can be observed that a sizeable improvement in efficiency is achieved for higher parent particle transverse momentum - hence closer decay products - in the new setup (labelled as TIDE, i.e. "Tracking In Dense Environments") with respect to the Run 1 algorithm setup. The efficiency for ρ is overall higher than for τ, due to the busier track environment in the latter case where merged clusters can arise from three particle contribution, and almost constant at near 100% in the simplified scenario without secondaries. The negative slope towards higher momentum is a result of inefficiencies of the clustering algorithm and the neural network classification, which leads to tracks being rejected because of too low truth-matching score. This effect is clearly exacerbated when more clusters from secondary tracks are introduced.

3.3.2 Cluster assignment efficiency, and average number of pixel clusters on track

The improvement in the identification of merged pixel clusters from a more sensible usage and revisited tuning of the neural networks leads to better performance in assigning the correct hits to a track candidate. Figure 36 illustrates the efficiency by which reconstructed tracks are properly assigned to a track in single ρ and single B^+ samples, parametrised with respect to the distance between particles. The chosen figure of merit is taken to be the minimum distance between pairs of truth decay products at the innermost pixel layer i.e., the IBL (located at $\langle r \rangle = 3.3$ cm from the nominal interaction point). The assignment efficiency is shown both for clusters at the IBL layer, and on the subsequent pixel layer outwards (Layer 1, $\langle r \rangle = 5.05$ cm). An efficiency above 95% is achieved with the new setup for the ρ case, even for distances of order of the minimal pixel separation of 50 μm. For the B^+ decay, the efficiency is worse because of the longer particle lifetime, which can cause daughter particles to be generated very close to the IBL (or Layer 1). This leads to multi-particle contributions to a single merged cluster,



(a) $\rho \to \pi^+\pi^-$, w/o secondaries

(b) $\rho \to \pi^+\pi^-$, w/ secondaries

(c) $\tau^\pm \to \pi^+\pi^-\pi^\pm\nu_\tau$, w/o secondaries

(d) $\tau^\pm \to \pi^+\pi^-\pi^\pm\nu_\tau$, w/ secondaries

Figure 35: Efficiency to reconstruct all charged decay products of single ρ or 3-prong τ_{had} particles, comparing the performance of the Run 1 tracking setup (green) to the changes implemented for Run 2 (TIDE, red). The plots on the left are obtained from samples where no secondary charged particles from material interactions in the ID are taken into account. The plots on the right side do consider such contributions instead.

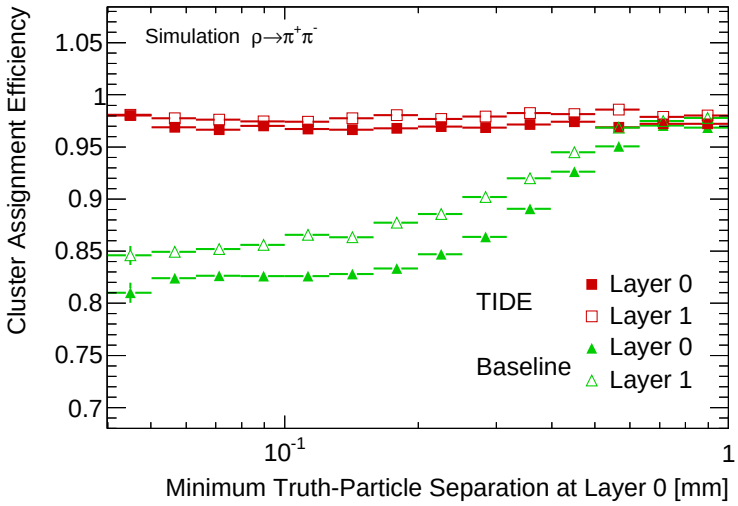
which are harder to identify. Overall, the new scheme outperforms the Run 1 tracking.

Furthermore, the average number of hits on a track after the ambiguity solving stage becomes closer to the expectation of the "ideal tracking" case. This is defined as a truth-based reconstruction, in which no combinatorial road finding is performed, and instead candidates are defined by grouping together truth-labelled clusters associated to the same truth particle. These candidates do not need to pass the ambiguity solving stage for what concerns shared clusters counting, and therefore have the ideal cluster content, given the cluster formation efficiency of the connected component analysis. Figure 37 represents the average number of pixel hits found on track pre and post-changes, compared to the ideal case, for the single ρ and τ_{had} samples as a function of parent particle p_T. The cluster content per candidate is much better reproduced after the modifications to the algorithm.

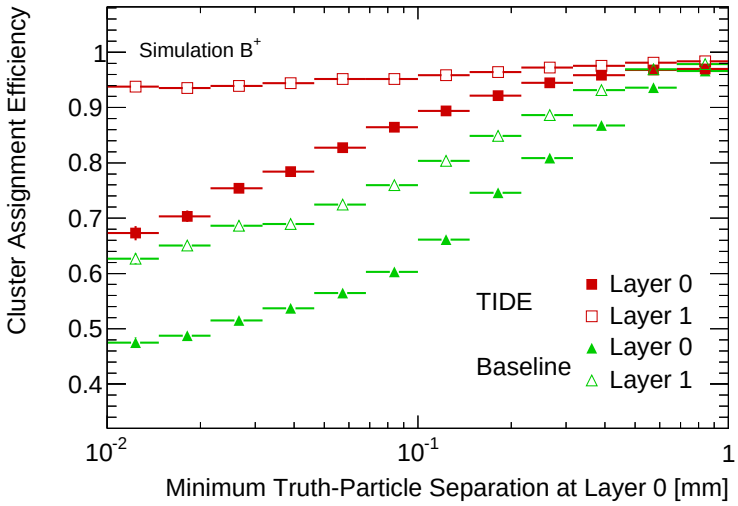
3.3.3 Performance in $Z' \to t\bar{t}$ events

The effect of the changes has been validated also on a sample of $Z' \to t\bar{t}$ events, with $m_{Z'} = 3$ TeV. Such events were simulated using the Pythia 8 generator [115] with MSTW 2008 Lo PDF sets [116]. Minimum-bias events generated with Pythia 8 are also included to simulate the pile-up effect expected during the Run 2 data taking. Jets in this sample are built from truth-level constituents using the standard anti-k_T algorithm described in Section 2.3.4, and are selected to have a minimum p_T of 100 GeV and $|\eta| < 2.5$.

For the reconstructed jet tracks, the average number of hits on the innermost pixel layer, and the reconstruction efficiency as a function of the angular distance to the jet axis are represented in Figure 38. For the efficiency, only jets with $450 < p_T < 750$ GeV are selected, and a minimal p_T of 2 GeV per track is required. Moreover, only charged particles which transverse the entire SCT detector ($r = 60.0$ cm) and originate within 10.0 cm from the beam line are considered, to take into account effects from large track displacement. Based on the flavour of the originating parton, b-jets are plotted separately from light-flavoured jets. At small distances, up to 10% more clusters on track are found with the Run 2 setup. After the changes, a large increase in efficiency is obtained, which is up to 10% (14%) higher in the core of light-flavoured (b) jets. The overall lower efficiency for b-jets is to be ascribed to the lesser cluster content due to the displaced production vertex of many of the jet tracks, as well as to merged cluster identification inefficiency as observed for the single B^+ case in Fig. 36b.

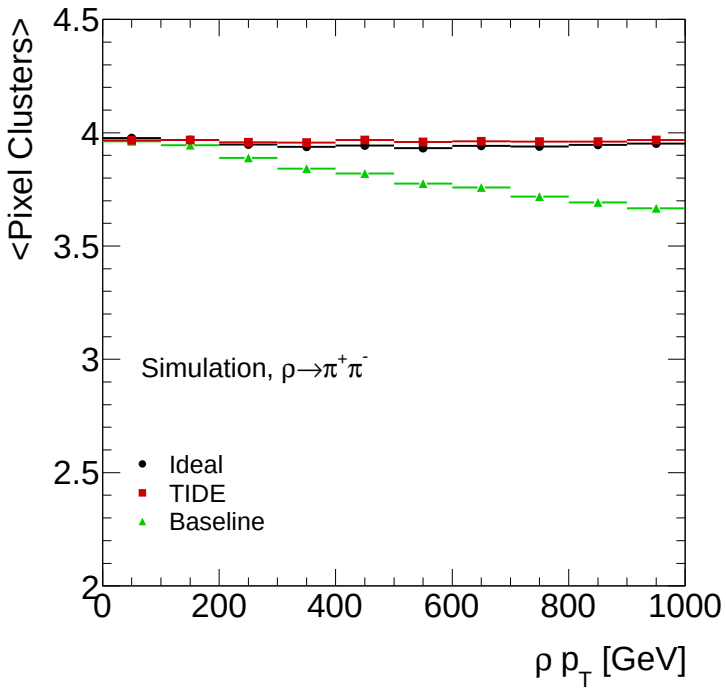


(a) $\rho \to \pi^+\pi^-$

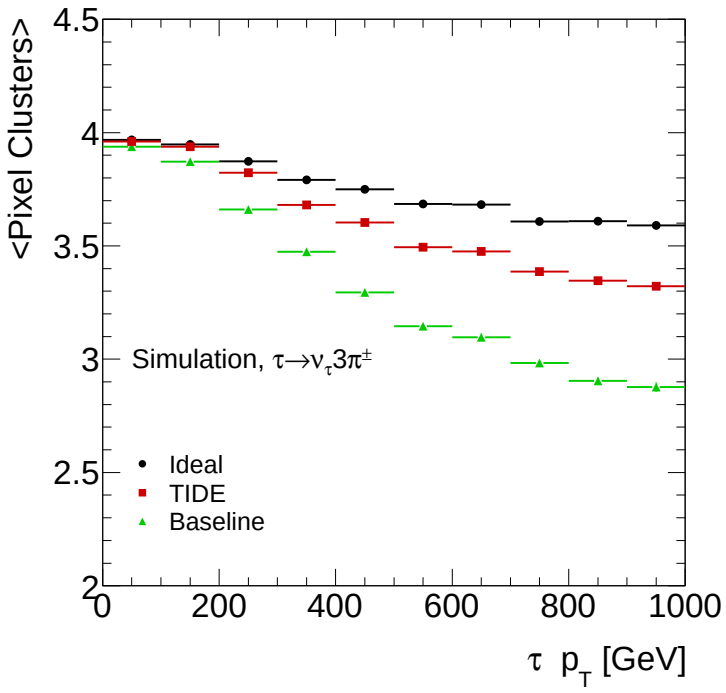


(b) B^+

Figure 36: Efficiency to assign correctly pixel clusters on a track candidate at Layer 0 (the IBL) and Layer 1, as a function of the minimal separation between pairs of truth particles at the innermost pixel layer, for single ρ and B^+ decay topologies. A comparison between the new Run 2 setup (TIDE, red) and the previous one from Run 1 (green) is displayed.

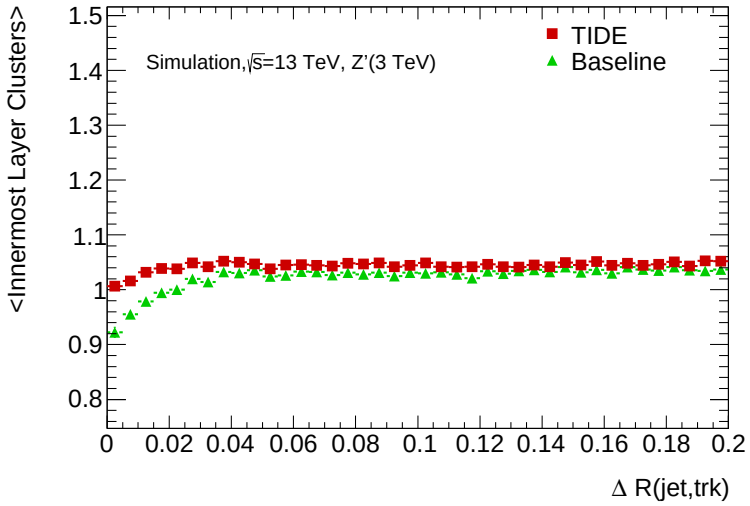


(a) $\rho \to \pi^+\pi^-$

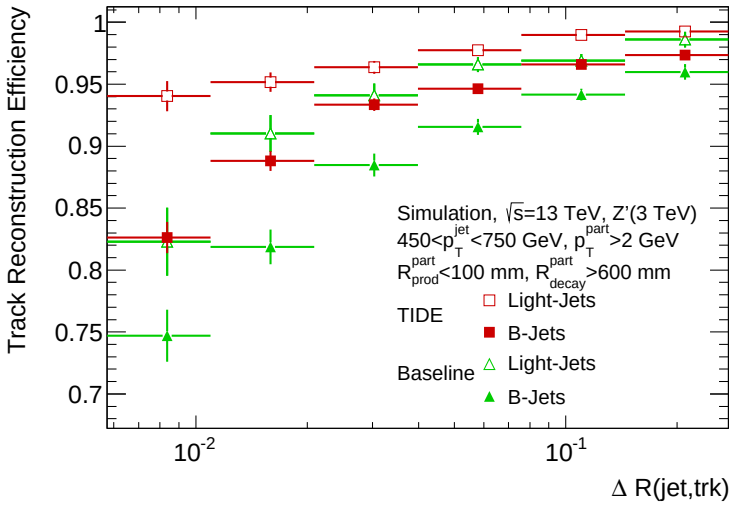


(b) $\tau^\pm \to \pi^+\pi^-\pi^\pm\nu_\tau$

Figure 37: Average number of pixel clusters on track as a function of ρ or τ_{had} p_T. The ideal case, resulting from truth-based cluster combination, is represented in black, and compared to the new Run 2 setup (TIDE, red) and the previous one from Run 1 (green).



(a)



(b)

Figure 38: Fig. 38a: the average number of innermost (IBL) pixel layer clusters on primary tracks with a production vertex before the first layer, as a function of the track distance from the jet axis. Fig. 38b: the efficiency to reconstruct charged primary particle tracks as a function of the track angular distance from the jet axis.

As introduced in Section 2.3.4, b-tagging algorithms in ATLAS rely on information about jet track impact parameters and properties of secondary vertices. The impact parameter resolution tends to worsen in presence of several shared measurements on track, and when hits are missing in the innermost pixel layers.

One of the inputs to the baseline ATLAS b-tagging algorithm is the IP3D [117] likelihood-based discriminator, which relies on impact parameter measurement. Its performance have been preliminary assessed after the changes to the Run 2 tracking. Figure 39 illustrates the b-jet identification efficiency versus the light-jet rejection factor of the IP3D discriminator, for jets selected in $Z' \to t\bar{t}$ events as previously described, comparing the Run 1 tracking configuration to the new one. For the baseline light-jet rejection factors corresponding to the standard 50, 60, 70, and 80% b-jet identification efficiency working points, the efficiency increases by 13, 11, 9, and 7%, respectively. A further improvement is foreseen, since the likelihood discriminant has not been re-trained using tracks after the Run 2 modifications for this first-pass check.

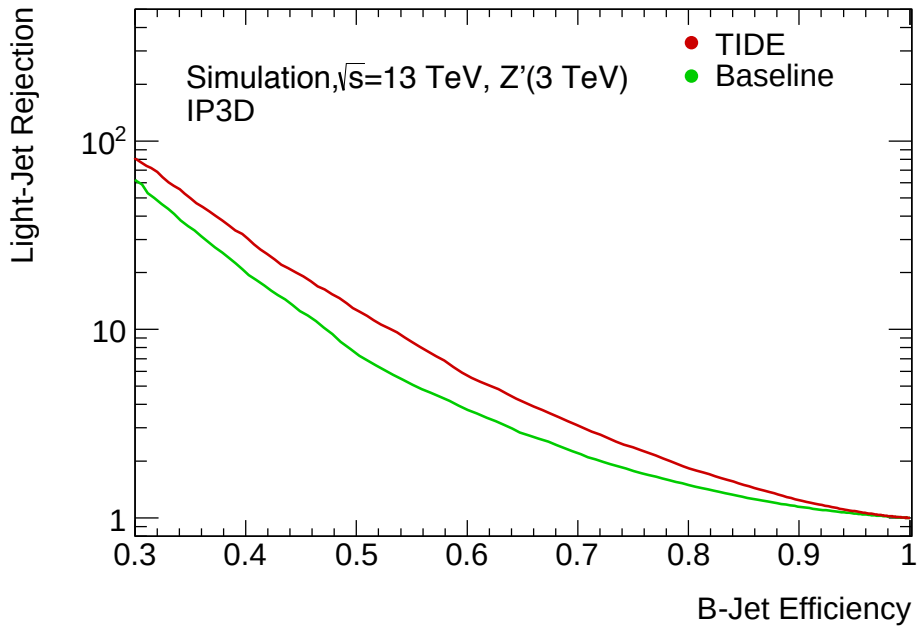


Figure 39: B-jet identification efficiency vs. light-jet rejection for the IP3D algorithm, for jets of $p_T > 100$ GeV and $|\eta| < 2.5$ in simulated $Z' \to t\bar{t}$ events.

3.4 CONCLUSIONS

In this chapter, the changes I introduced in the ATLAS tracking software to improve its performance in dense-track topologies have been presented. The new algorithm allows better identification of merged pixel clusters from non-resolved nearby charged particles. This has been achieved by making a more effective use of a set of neural network algorithms already available during Run 1, and by looking at global track properties in the ambiguity solving stage to better cope with multi-track topologies.

Performance have been successfully tested after the changes on simplified samples with single particle decays to multiple charged daughters. Better cluster-to-track assignment efficiency, and track reconstruction efficiency are reached with respect to the Run 1 setup. Validation on a more realistic Z' sample shows improved tracking performance as well, and enhanced b-tagging performance for an impact-parameter based identification algorithm.

Near-future developments in this field focus on finding a way to identify SCT merged clusters, to further relax the current track quality requirements. Another topic of research is the measurement of tracking efficiency in the core of high p_T jets in a fully data-driven way, using the charged particle energy loss (dE/dx) to identify pixel clusters originating from two particles. This information can be used to constrain the uncertainty on the tracking inefficiency in the cores of jets, which is relevant for jet energy and mass calibrations, as well as measurements of jet properties. Preliminary results on this measurement are described in Ref. [108].

4 SEARCH FOR THE $t\bar{t}H$ PRODUCTION IN THE MULTILEPTONIC FINAL STATE

This chapter aims to describe the $t\bar{t}H$ to multi-leptons search carried out on the data collected in 2015 and 2016 by the ATLAS experiment at the LHC in p-p collisions at $\sqrt{s}$ = 13 TeV.

Motivation for the analysis strategy as a substantial improvement to the Run 1 search is initially provided. I concentrate on the high-sensitivity final state with two identified light leptons (e, μ) [1] with same electric charge and no hadronically decaying taus (2ℓ SS $0\tau_{had}$). This represents the main focus of my work. Detailed information on the collision dataset and simulated samples is given, together with the event and physics object selection. A description of the multi-variate discriminators used to separate the $t\bar{t}H$ signal from the major backgrounds is provided afterwards. Subsequently, the background modelling is discussed, with a particular focus on the estimation of the background from fake electrons and muons. The impact of the main sources of systematic uncertainty on the measurement is then discussed. The statistical model for the interpretation and the results of the analysis will be presented in the following chapter.

1 In the following, light leptons will be also referred to as *leptons*, both for their particles and antiparticles.

4.1 INTRODUCTION AND MOTIVATION FOR THE RUN 2 ANALYSIS

At the ATLAS experiment, the search for the SM Higgs boson[2] produced in association with top quark pairs is performed in three independent analyses, which target different final states:

- $t\bar{t}(H \to b\bar{b})$
- $t\bar{t}(H \to \gamma\gamma)$
- $t\bar{t}H \to \text{multi} - \text{leptons}$

The $t\bar{t}(H \to b\bar{b})$ channel [118, 119] is characterised by the largest Higgs decay branching fraction, and by a peculiar final state signature with numerous b-jets. Despite the small branching ratio to photons, the $t\bar{t}(H \to \gamma\gamma)$ channel [120] can profit from the excellent diphoton mass peak resolution and high signal purity. However, the ATLAS searches targeting these channels will not be covered in this thesis.

As the name suggests, the multi-leptonic channel does not focus on an exclusive Higgs decay. Rather, it comprises several final states categorised by the multiplicity of identified light leptons and hadronic taus, and it is sensitive to several Higgs decay modes ($H \to WW^*$, $H \to \tau^+\tau^-$, $H \to ZZ^*$). This choice is driven by the current goal of constraining the rate of $t\bar{t}H$ production mechanism, rather than to provide constraints on the light fermion and boson couplings accessed through the Higgs decays. The latter is addressed by several other independent ATLAS analyses, as highlighted in subsection 1.2.2.

4.1.1 Summary of Run 1 results, and prospects for Run 2

The strategy of the Run 1 $t\bar{t}H$ to multi-leptons analysis [121] is that of a simple counting experiment. The prediction of the overall event yield from the SM $t\bar{t}H$ signal and background model is considered in several categories based on the number of leptons, taus, and the lepton flavour composition in the final state. The production signal strength parameter $\mu_{t\bar{t}H} = \sigma/\sigma_{SM}$ is then fitted to the observed data independently in each category, and results are eventually combined. This strategy has been determined by the low statistics available, due to the very small $t\bar{t}H$ cross section at $\sqrt{s}$ = 8 TeV (≈

2 For brevity, henceforth it will be also referred to as simply the *Higgs*.

130 fb) and the small size of the collision dataset, equivalent to an integrated luminosity of 20.3 fb^{-1}. The best-fit value achieved is $\mu_{t\bar{t}H} = 2.1^{+1.4}_{-1.2}$, and an upper limit at 95% confidence level of $\mu_{t\bar{t}H} < 4.7$ is also set. The expected sensitivity of the search - computed as a p-value against the SM expectation of $\mu = 1$ - is found to be equivalent to 0.9σ, with an observed value of 1.8σ when testing the background-only hypothesis.

The fitted signal strength values for each channel and their combination are summarised in Figure 40, which includes also the results of the analogous CMS search [122]. The best-fit production signal strength from the combination of all the $t\bar{t}H$ channels of ATLAS and CMS is $\mu_{t\bar{t}H} = 2.3^{+0.7}_{-0.6}$. Notably, the excess over the SM prediction is mostly driven by the measurement in the multi-leptonic channel.

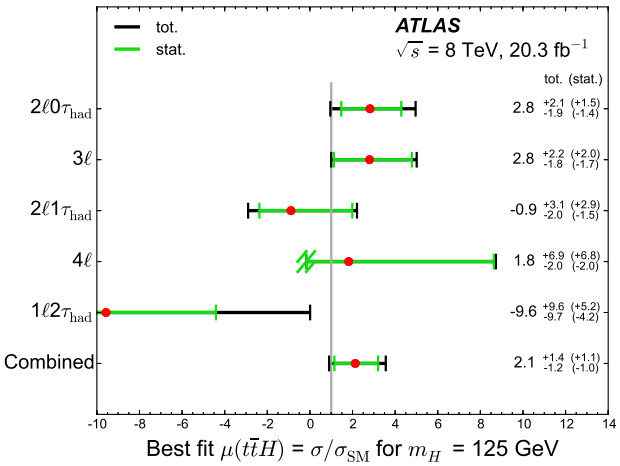
At $\sqrt{s}$ = 13 TeV, the $t\bar{t}H$ cross section grows aproximately by a factor 4 compared to $\sqrt{s}$ = 8 TeV. This is a sizeable increase compared to other Higgs production modes, and is due to the relatively larger growth of the gluon-gluon parton luminosity with respect to the quark-gluon and quark-quark luminosities. Neglecting systematic uncertainties and assuming $t\bar{t}$ as the only background, a boost in statistical significance ($S/\sqrt{B}$) of about a factor 2 for a dataset of the same size as the Run 1 one is expected. This is illustrated in Figure 41.

In summer 2016, a preliminary set of results on 13.2 fb^{-1} of data at $\sqrt{s}$ = 13 TeV were published [124]. The analysis follows the same approach as in Run 1 regarding the background modelling and the fit strategy, albeit including updates on the event and object selection to cope with the beam conditions and detector changes introduced during Run 2. A best-fit of $\mu_{t\bar{t}H} = 2.5^{+1.3}_{-1.1} = 2.5^{+0.7}_{-0.7}(\text{stat.})^{+1.1}_{-0.9}(\text{syst})$ is obtained, and the observed (expected) sensitivity of the search is quantified as 2.2σ (1.3σ). The latter result roughly matches the expectation based on the above argument for "identical" analyses, given the different integrated luminosity of the datasets.

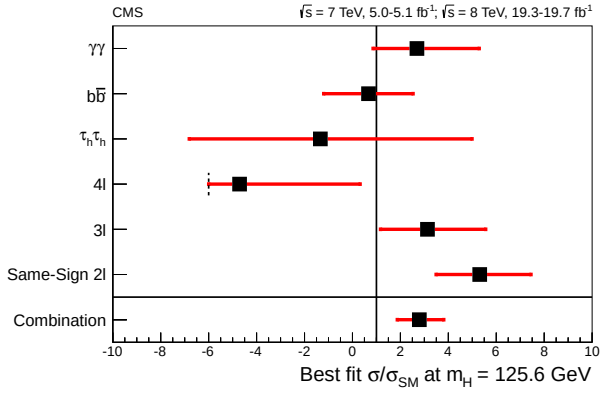
Even with only slightly more than a half of the Run 1 integrated luminosity, this preliminary Run 2 analysis reveals how the total uncertainty on μ starts being dominated by the systematics uncertainties, which are mostly ascribed to background modelling and theoretical uncertainties on the simulated samples. This is illustrated more clearly in Table 5, where the impact of the different sources of systematic uncertainty on μ in this early Run 2 analysis is presented.

With the combination of the full 2015 and 2016 datasets and an optimised analysis setup, reaching evidence for the $t\bar{t}H$ production becomes foreseeable when combining the multi-leptonic channel with others. In pursuance

Channel	$\mu_{t\bar{t}H} = \sigma/\sigma_{SM}$	
	ATLAS	CMS
$t\bar{t}H \to \text{multi} - \text{leptons}$	$2.1^{+1.4}_{-1.2}$	–
$t\bar{t}(H \to b\bar{b})$	$1.4^{+1.0}_{-1.0}$	$0.7^{+1.9}_{-1.9}$
$t\bar{t}(H \to \gamma\gamma)$	$1.3^{+2.62}_{-1.75}$	$2.7^{+2.6}_{-1.8}$
Combination	$1.9^{+0.8}_{-0.7}$	$2.9^{+1.0}_{-0.9}$
	$2.3^{+0.7}_{-0.6}$	

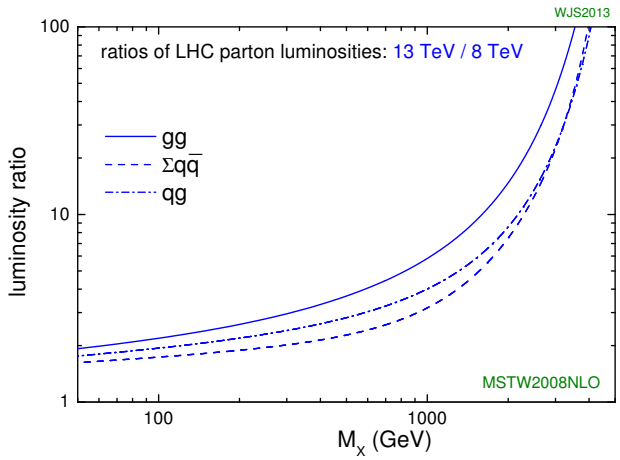


(a)



(b)

Figure 40: Run 1 ATLAS and CMS Higgs boson signal strength measurement for the $t\bar{t}H$ production, combining the $t\bar{t}(H \to b\bar{b})$, $t\bar{t}(H \to \gamma\gamma)$, and $t\bar{t}H \to \text{multi} - \text{leptons}$ channels [122, 121]. The table also presents the combination of the measurements performed by each experiment. The plots at the bottom show the best fit μ split into each multi-leptonic final state category (the CMS plot includes also the $t\bar{t}(H \to \gamma\gamma)$, $t\bar{t}(H \to b\bar{b})$ channels, and no multi-leptonic final states combination is available). The observed excess over the SM prediction is driven by the multi-leptonic $2\ell\ SS$ and 3ℓ categories.



σ [fb]	$\sqrt{s}$ = 8 TeV	$\sqrt{s}$ = 13 TeV	Ratio (13/8)
ggF	1.95×10^4	4.68×10^4	2.4
VBF	1.58×10^3	3.78×10^3	2.4
WH	6.97×10^2	1.37×10^3	2.0
ZH	3.94×10^2	8.84×10^2	2.3
$t\bar{t}H$	1.30×10^2	5.07×10^2	3.9
$t\bar{t}$	2.53×10^5	8.30×10^5	3.3
$\sigma_{ttH}/\sqrt{\sigma_{t\bar{t}}}$	0.26	0.54	2

Figure 41: In the table: Higgs and $t\bar{t}$ production cross section at $\sqrt{s}$ = 8, 13 TeV. Assuming $t\bar{t}$ as the dominant background for the $t\bar{t}H$ search, the bottom row shows the ratio $\sigma_S/\sqrt{\sigma_B}$ for $\sqrt{s}$ = 8, 13 TeV, and the increase in statistical significance obtained by taking the ratio of the two is presented in the bottom right box. The latter assumes the same integrated luminosity for the 8 and 13 TeV datasets, acceptance, and selection efficiencies. The highest sensitivity boost for $t\bar{t}H$ is reached mostly because of the relatively larger increase in g-g parton luminosity, and because of the large mass of the $t\bar{t}H$ final state. This is illustrated in the plot at the top [123], where the PDFs are taken from the MSTW2008 set [116] (here $M_X \equiv \sqrt{\hat{s}}$).

Uncertainty Source	Type	$\Delta\mu$	
Fake background modelling	Background	$+0.56$	-0.64
JVT association, pile-up modelling	Detector	$+0.48$	-0.36
$t\bar{t}W$ modelling	Background	$+0.29$	-0.31
$t\bar{t}H$ modelling	Signal	$+0.31$	-0.15
Jet energy scale and resolution	Detector	$+0.22$	-0.18
$t\bar{t}Z$ modelling	Background	$+0.19$	-0.19
Luminosity	Detector	$+0.19$	-0.15
Diboson modelling	Background	$+0.15$	-0.14
Jet flavor tagging	Detector	$+0.15$	-0.12
e, μ and τ_{had} ID, isolation, trigger	Detector	$+0.12$	-0.10
Other background modelling	Background	$+0.11$	-0.11
Total systematic uncertainty		$+1.1$	-0.9

Table 5: Summary of the effects of the main sources of systematic uncertainty (detetctor-related, signal/background-modelling-related) on the fitted value of μ, ranked from the highest to the lowest contribution, as in the early Run 2 analysis of 13.2 fb^{-1} of data at $\sqrt{s}$ = 13 TeV [124]. The sum in quadrature of the individual systematics does not sum up to the total uncertainty because of correlations between different sources of systematic uncertainty.

of this goal, the reduction of systematic uncertainties represents a pivotal point. This can be achieved through improved background estimation techniques, and by devising more refined selections to suppress the main backgrounds.

Furthermore, a change in the fit strategy to enhance the sensitivity of the search is also beneficial. Indeed, the larger size of the datasets allows to exploit information about event topology and kinematics to better separate the $t\bar{t}H$ signal from the other processes. This approach is also expected to be more effective at defining pure background-enriched control regions to be included in the fit, in order to better constrain the systematic uncertainties on the backgrounds [3]. As a result of the higher statistics and the upgrades of the analysis methodology, a sizeable improvement in sensitivity with respect to the simple cut-and-count approach of Run 1 and early Run 2 is predicted.

4.2 ANALYSIS CATEGORISATION AND STRATEGY

The best sensitivity of the search is achieved by classifying the collected events into a set of orthogonal categories, based on the number of identified

3 In the Run 2 analysis, control regions are defined exclusively for the 3ℓ $0\tau_{had}$ final state.

lepton and hadronic tau candidates. For the 2015 and 2016 dataset analysis, 8 independent multi-leptonic categories are defined:

- 2ℓ SS $0\tau_{had}$: exactly two light leptons of same electric charge, and no hadronic taus.
- 3ℓ $0\tau_{had}$: three light leptons, whose electric charges must sum to ± 1, and no hadronic taus.
- 4ℓ (Z-enriched/Z-depleted): four light leptons, with zero total electric charge.
- 2ℓ SS $1\tau_{had}$: same as 2ℓ SS $0\tau_{had}$, but with exactly one hadronic tau.
- 2ℓ OS $1\tau_{had}$: exactly two light leptons of opposite electric charge, and one hadronic tau.
- 3ℓ $1\tau_{had}$: same as 3ℓ $0\tau_{had}$, but with exactly one hadronic tau.
- 1ℓ $2\tau_{had}$ OS : one light lepton, and two opposite-sign charged hadronic taus.

The overall categorisation is sketched in Figure 42. Depending on its lepton composition, each category is more or less sensitive to different Higgs decays, and the background contamination also varies across categories. Table 6 shows the percentage of events for each Higgs decay mode in every category, as well as $S/\sqrt{B}$ (where S and B indicate the total number of signal and background events, respectively) and the signal acceptance times efficiency, $(A \times \varepsilon)_{t\bar{t}H}$. The latter includes SM Higgs and top quark branching fractions, detector acceptance, reconstruction and selection efficiency, and is computed relative to the inclusive $t\bar{t}H$ production rate using simulation. The 4ℓ category is sub-divided into Z-enriched and Z-depleted regions, depending on the presence of a same-flavour, opposite-sign lepton pair. In most categories the $H \to WW^*$ decay mode is dominant, with an increasing fraction of $H \to \tau^+\tau^-$ when explicitly requiring hadronic taus in the final state. Some sensitivity to the $H \to ZZ^*$ decay is also achieved, which grows as more light leptons are required.

4.2.1 The 2ℓ SS$0\tau_{had}$ category

The rest of the chapter focuses on the analysis of the 2ℓ SS $0\tau_{had}$ final state, to which my work has been mainly devoted. It is characterised by

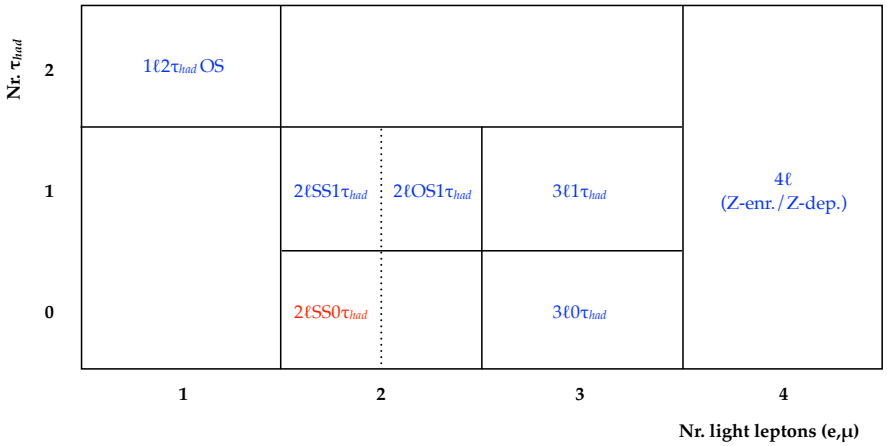


Figure 42: Categories of the $t\bar{t}H \rightarrow \text{multi}-\text{leptons}$ search, defined in terms of reconstructed light lepton and τ_{had} multiplicity.

Category	H decay mode WW^*	$\tau\tau$	ZZ^*	$(A \times \varepsilon)_{t\bar{t}H}$ $(\times 10^{-4})$	$S/\sqrt{B}$
2ℓ SS $0\tau_{had}$	80%	17%	2%	23	1.94
3ℓ $0\tau_{had}$	79%	14%	6%	13	1.79
3ℓ $1\tau_{had}$	35%	61%	4%	9.2	0.95
2ℓ OS $1\tau_{had}$	38%	58%	1%	7.8	0.51
1ℓ $2\tau_{had}$ OS	2%	98%	$< 1\%$	2.3	0.51
2ℓ SS $1\tau_{had}$	39%	59%	1%	1.7	1.08
4ℓ (Z-enriched)	71%	13%	13%	0.6	0.71
4ℓ (Z-depleted)	77%	22%	1%	0.1	0.60

Table 6: Expected $t\bar{t}H$ signal fractions from different Higgs decay modes in the different analysis categories, their $S/\sqrt{B}$, and signal acceptances times efficiencies [125].

the presence of two leptons of same or different flavour, with same electric charge, missing energy from two neutrinos, and high jet multiplicity. A representation of a typical 2ℓ SS $0\tau_{had}$ $t\bar{t}H$ signal event is shown in Figure 43. In such event, the top quarks decay to $\bar{t} \to \bar{b}(W^- \to \ell^-\bar{\nu}_\ell)$ and $t \to b(W^+ \to qq')$ respectively, whereas the Higgs decays semileptonically to $H \to WW^* \to \ell^-\bar{\nu}_\ell qq'$. Six reconstructed jets are expected, without taking into account additional jets from initial or final state radiation, or missing jets because of limited acceptance and reconstruction inefficiency. Two of those jets originate from b quark hadronisation. Being the channel with the least number of leptons, it profits from higher signal statistics than the other channels owing to the largest branching ratio and acceptance (see Table 6). Indeed, as Table 6 shows, the 2ℓ SS $0\tau_{had}$ is the category with the highest sensitivity. This justifies the effort put into its analysis optimisation for Run 2.

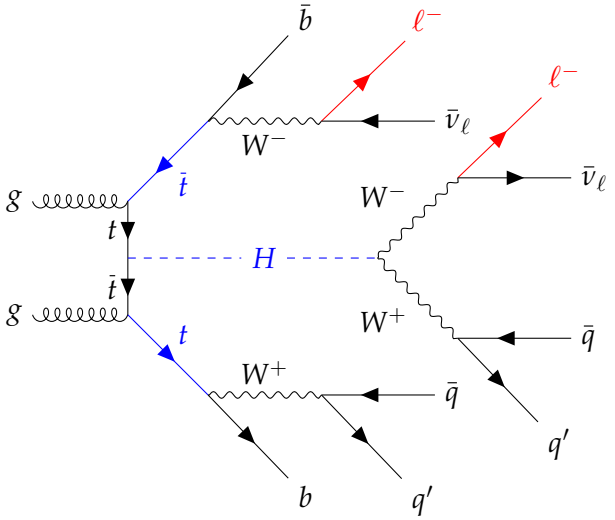


Figure 43: Leading order Feynman diagram of the most common $t\bar{t}H$ signal event in the 2ℓ SS $0\tau_{had}$ category, where the Higgs decays in a pair of W bosons. One lepton originates from the decay of the W^- (W^+) from the Higgs, the other lepton of same charge originates from the semileptonic decay of the $\bar{t}$ (t).

4.2.2 Backgrounds in the $t\bar{t}H$ to multileptons search

In the SM, processes giving rise to pairs of leptons of same electric charge are very rare. $t\bar{t}H$ represents one of such processes, e.g. when one lepton originates from the decay of the W^- (W^+) from the Higgs and another lepton of same charge originates from the semileptonic decay of the $\bar{t}$ (t). There-

fore, selecting events with same-sign lepton pairs as in the 2ℓ SS $0\tau_{had}$ category is a very effective handle against some of the most common SM background processes which possess a cross section orders of magnitude larger than $t\bar{t}H$, and could easily mimic its signature in an opposite-sign dileptonic final state. These include the $t\bar{t} \to (W^+ \to \ell^+\nu_\ell)b(W^- \to \ell^-\bar{\nu}_\ell)\bar{b}$ dileptonic process and Z + jets production.

Irreducible backgrounds

Nevertheless, other rare SM processes do provide a genuine same-sign dilepton final state too. These are mainly $t\bar{t}W$, $t\bar{t}(Z/\gamma^*)$ [4], and diboson production such as WZ, ZZ with additional jets from parton radiation. Other rare processes with this signature include $t\bar{t}t\bar{t}$ production, tZ, and tWZ. Figure 44 displays a few illustrative Feynman diagrams for the $t\bar{t}Z$, $t\bar{t}W$ and WZ production with a same-sign dileptonic final state, and their cross sections at $\sqrt{s}$ = 13 TeV calculated at NLO in perturbative QCD [126, 127]. Cross sections times branching ratios for such processes are of the same order of magnitude as the $t\bar{t}H$ production (see Tables 7,8). Because of their signal-like signature, events from these processes likely pass the analysis selection. Therefore, they are indicated as *irreducible* backgrounds.

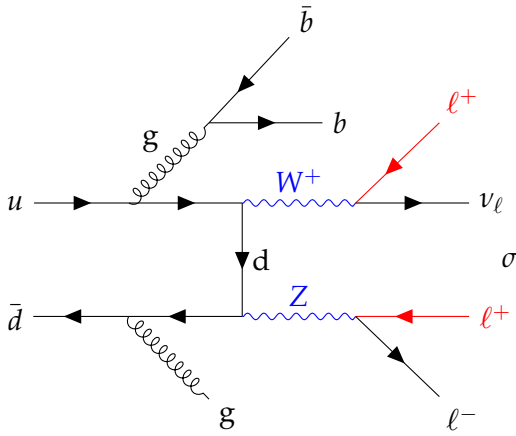
Reducible backgrounds

A final state with two same-sign light leptons can also arise from misreconstructing the charge of one lepton in an opposite-sign pair, by mistakenly identifying a particle as a lepton, or by the presence of spurious leptons not associated with the primary interaction. For the 2ℓ SS $0\tau_{had}$ category, mainly three effects are to be considered:

- *Charge flips*: a lepton - mostly an electron - radiates through hard bremsstrahlung via $e^\pm \to e^\pm\gamma \to e^\pm e^- e^+$, and the electron daughter track with opposite sign to the original electron is reconstructed and incorrectly associated to the EM calorimeter cluster. Another charge misidentification source can be due to wrong measurement of the track curvature, which is more likely for slightly bent, high p_T trajectories. Again, this is seen to affect electrons, and is found to be negligible for muons owing to the longer lever arm of the muon tracks.

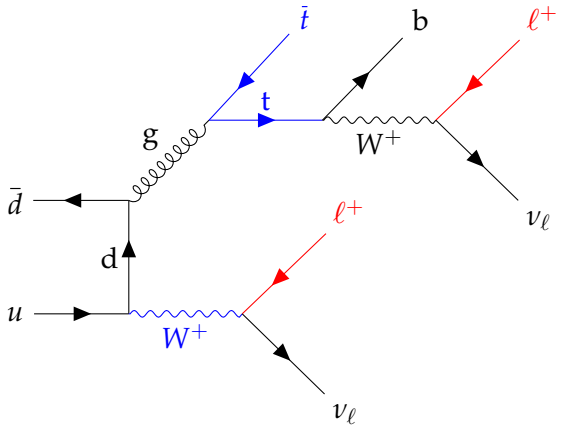
- *Electrons from photon conversions*: a real photon emitted as initial or final state radiation, or in the parton showering, can interact with the

4 For brevity, this process is also referred to as $t\bar{t}Z$.



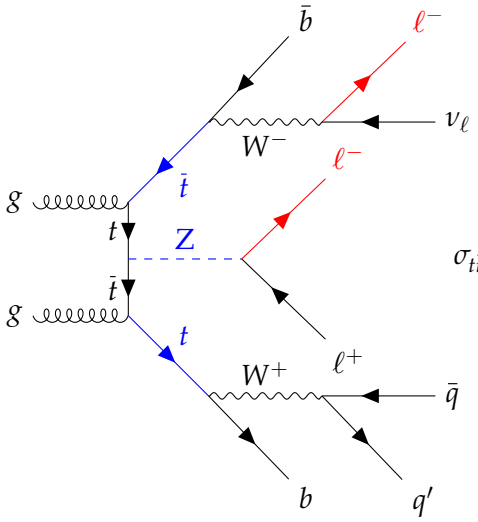
$\sigma_{WZ} \times \mathrm{BR}(WZ \to 3\ell) = 11.88$ [pb]

(a) WZ



$\sigma_{t\bar{t}W} = 0.601$ [pb]

(b) $t\bar{t}W$



$\sigma_{t\bar{t}(Z/\gamma^*)} \times \mathrm{BR}(Z \to \ell^+\ell^-) = 0.12$ [pb]

(c) $t\bar{t}(Z/\gamma^*)$

Figure 44: Examples of LO Feynman diagrams for the $WZ \to 3\ell$, $t\bar{t}(W \to \ell\,^\circ)$ and $t\bar{t}(Z/\gamma* \to \ell^+\ell^-)$ irreducible backgrounds with a 2ℓ SS $0\tau_{had}$ final state, and their cross sections (times branching ratio of the decay, for $t\bar{t}Z$ and WZ) at $\sqrt{s}$ = 13 TeV calculated at NLO in perturbative QCD [128, 126]. In the WZ and $t\bar{t}Z$ cases, the third lepton can be missed when falling outside the detector acceptance, failing the object selection, or because of reconstruction inefficiency.

detector material as early as in the beam pipe, and produce additional electron pairs.

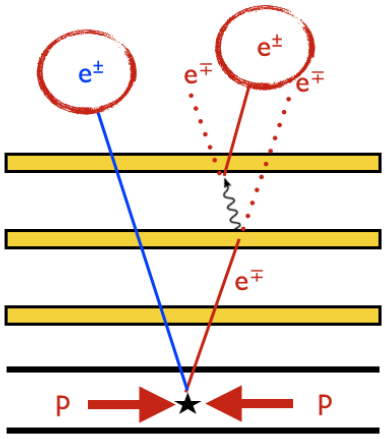
- *Non-prompt leptons*: the most relevant source of fake electrons and muons are the leptonic decays of bottom-flavoured or charm-flavoured hadrons produced in association of b-c-originated jets.

Such leptons are generally referred to as *fake* leptons. A pictorial representation of the main types of fakes in this analysis is presented in Figure 45.

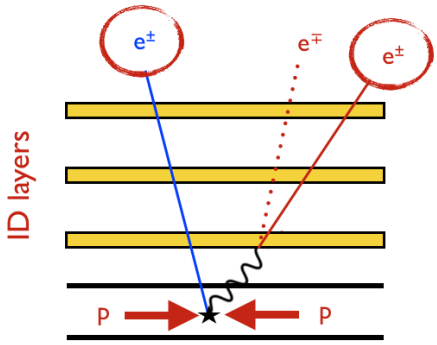
The main process that can contaminate the 2ℓ SS $0\tau_{had}$ category by the presence of fake leptons is largely $t\bar{t}$, followed by single top production, tW and to a lesser extent, $W+$jets and $Z+$jets . This is described in Figure 46, where the fake background process composition is displayed as obtained from pure simulation in the inclusive 2ℓ SS $0\tau_{had}$ category, and for the subsets split by lepton flavour composition. In the following, we will always consider $t\bar{t}$ as the reference process for fakes estimation.

A typical diagram for the $t\bar{t} \to (W^+ \to \ell^+\nu_\ell)b(W^- \to qq')\bar{b}$ semileptonic process is shown in Figure 47. The additional same-sign lepton can be either a mis-identified physics object, or a lepton originating from outside the primary interaction. Generally, the impact of these backgrounds can be strongly reduced by requiring strict quality criteria on the reconstructed lepton candidates, to ensure they are most likely real leptons from the primary interaction. These requirements usually include passing some level of particle identification, small lepton track impact parameters, and low activity around the lepton. Since these processes are compatible with the signal final state only via the presence of a fake lepton, they are known as *reducible* backgrounds.

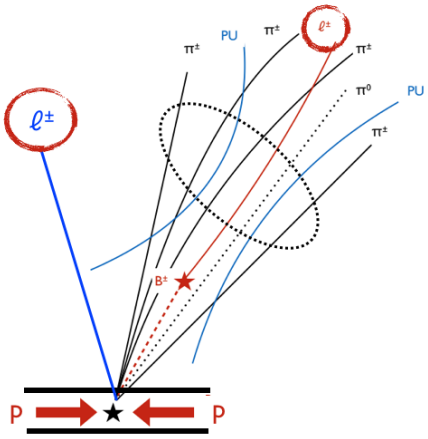
Yet, selection algorithms intrinsically have some degree of inefficiency at rejecting fakes. Moreover, these processes are characterised by cross sections which are typically orders of magnitude higher than the $t\bar{t}H$ one. This implies a non negligible amount of events with fake leptons populate the 2ℓ SS $0\tau_{had}$ category. The estimation of the rate of fakes and of their final state kinematics is a delicate task. Many fake leptons are produced in busy environments, or through secondary particle interactions with the material of the detector, whose rate may not be precisely modelled by simulation. Furthermore, the sample of simulated events with fakes is usually small, since those events only take up for a small fraction of the total generated events. Therefore, data-driven methods are used to estimate the reducible backgrounds.



(a) Charge flip electron



(b) Electron from γ conversion



(c) Non-prompt lepton in B hadron decay

Figure 45: Major sources of fake leptons for the 2ℓ SS $0\tau_{had}$ analysis. The star indicates the primary hard-scattering vertex of the p-p collision. 45a: in an OS event, an electron is reconstructed with flipped charge because of hard bremsstrahlung. 45b: in a single lepton event, an additional same-sign electron can be reconstructed as a result of photon conversion in the detector material. 45c: a $B^{\pm} \to \ell^{\pm}\nu_{\ell} + X$ decay in a b-originated jet in a single lepton event can lead to a same-sign dileptonic signature. The secondary decay vertex of the $B^{\pm}$ is indicated with a red star. These sketches have a purely illustrative purpose, and are not to scale.

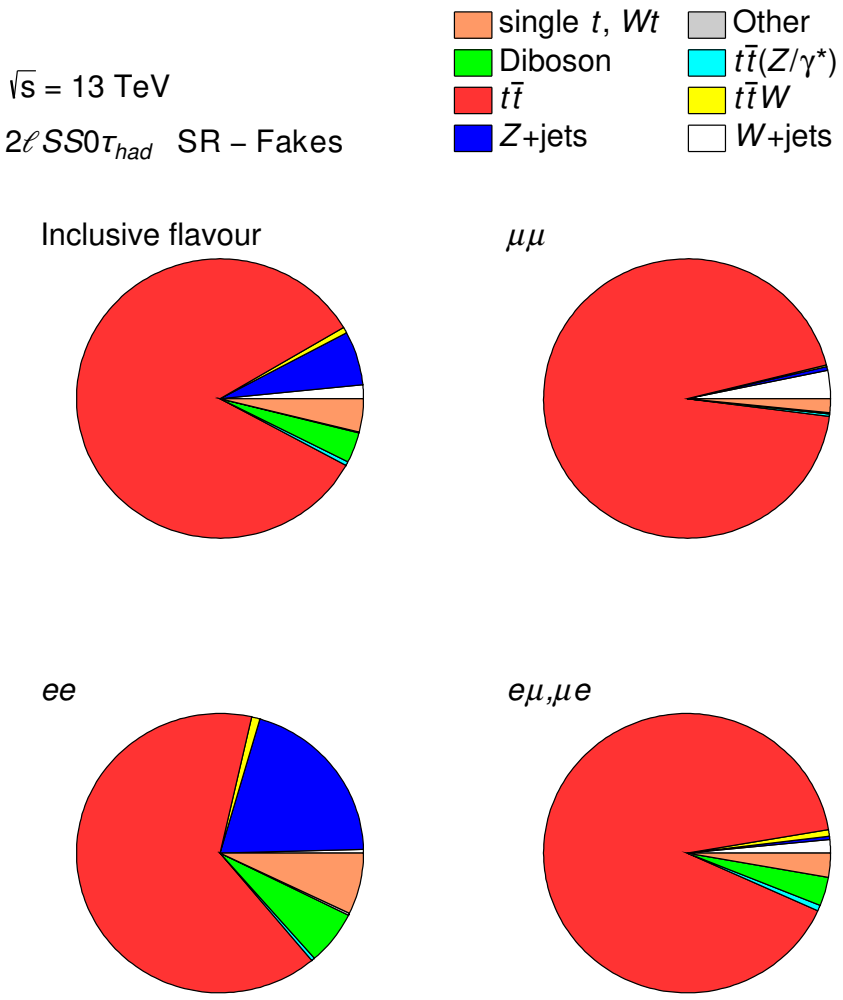


Figure 46: Pie charts showing the fractional contributions to the total fake leptons background in the 2ℓ SS $0\tau_{had}$ category, based on simulation.

A large part of this thesis work has been focused on developing and validating the estimation of fake backgrounds in the context of this search. A full description of the topic will be provided later in the chapter.

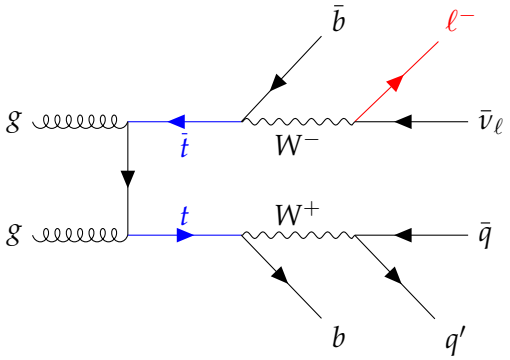


Figure 47: LO Feynman diagram for the reducible $t\bar{t} \to (W^+ \to \ell^+\nu_\ell)b(W^- \to qq')\bar{b}$ background for the $t\bar{t}H$ to multi-leptons search. The source of the second same-sign light lepton can be either a mis-identified particle, or a lepton produced outside the primary interaction.

4.3 COLLISION DATASET AND SIMULATED SAMPLES

This analysis considers the Run 2 p-p collision data collected during the years 2015 and 2016 at a centre-of-mass energy of $\sqrt{s} = 13$ TeV. The total size corresponds to an integrated luminosity of 36.1 fb^{-1}, as obtained after applying data quality requirements to ensure all the detector subsystems were fully operative. The fractions of integrated luminosity split for 2015 and 2016 is illustrated in Figure 17. The presence of at least one primary vertex - reconstructed from at least two well measured tracks with $p_T > 400$ MeV - is required. The hard-scattering vertex of the collision is then identified as the one with maximal $\sum p_T^2$ of all associated tracks. Corrupted data events and events incompletely reconstructed are excluded, as well as events affected by significant noise in the calorimeters.

4.3.1 Simulation

The full ATLAS event simulation workflow has been described previously in section 2.4. For all samples used in this analysis, generated events are passed through a full simulation of the ATLAS detector. Minimum bias pile-up events are modeled with the PYTHIA 8.1 [129] generator, using the

Process	σ ($\times$ BR) [fb]	Simulated $\int Ldt$ [fb^{-1}]
$t\bar{t}H$	$507^{+5.8\%+3.6\%}_{-9.2\%-3.6\%}$	–
$(t\bar{t} \to 2\ell)H$	53	9.24×10^4
$(t\bar{t} \to 1\ell)H$	222	2.21×10^4
$(t\bar{t} \to 0\ell)H$	231	1.95×10^6

Table 7: Production cross section of the SM $t\bar{t}H$ signal process calculated at NLO precision in QCD and EWK couplings. The per-cent uncertainties reported are relative to the QCD factorisation and renormalisation scales (left), and the choice of PDF and α_s (right). The bottom part of the table shows the cross section times branching ratio of the simulated events filtered by lepton multiplicity, and their integrated luminosity.

MSTW2008LO [116] PDF set with the A2 tune for the underlying event generator [130]. Photon emission is simulated to leading-logarithm accuracy either by the parton showering generator used for each sample, or by the PHOTOS software [131].

Based on the recommendations provided by the LHC Higgs Cross Section Working Group [46], the $t\bar{t}H$ signal events are simulated to NLO accuracy in both QCD and QED using the MADGRAPH5_AMC@NLO [128] generator, interfaced to PYTHIA 8.1 with the A14 tune [132] to simulate parton showering and hadronisation. The PDF set used in the cross section calculation is the NNPDF 3.0 [133] set, whereas NNPDF 2.3 [23] is used in the parton shower generator. The Higgs boson mass in the calculation is set to $m_H = 125$ GeV, and the top quark mass to $m_t = 172.5$ GeV. The uncertainty associated to the baseline NNPDF 3.0 PDF choice is assessed by making a comparison with results obtained from other PDF sets - each with a different value of α_s - following the procedure recommended by the PDF4LHC collaboration [134], and their size is 3.6%. The central value for renormalisation and factorisation scales is set to a fixed scale of $\mu_0 = \mu_R = \mu_F = m_t + m_H/2$, and the scale uncertainty is estimated by varying independently both scales in the range $\mu_0/2 \leq \mu_R, \mu_F \leq 2\mu_0$, eventually considering the largest variation of the two: the size of this uncertainty is about $+6\%, -9\%$. In order to enhance the available event statistics for the multi-leptonic final states, three separate $t\bar{t}H$ samples are generated, split by the number of prompt leptons in the decay mode of the top-antitop quark pair. The total $t\bar{t}H$ cross section with its associated uncertainties is summarised in Table 7. The table also displays the cross section times branching ratio for each of the three samples, and the simulated integrated luminosity.

Process	σ ($\times$ BR) [fb]	Simulated $\int Ldt$ [fb^{-1}]
$t\bar{t}W$	$600.8^{+12.9\%+3.4\%}_{-11.5\%-3.4\%}$	1.25×10^4
$t\bar{t}Z$	$839.3^{+9.6\%\ +4\%}_{-11.3\%-4\%}$	–
$t\bar{t}(Z \to \ell\ell)$	$124.0^{+9.6\%\ +4\%}_{-11.3\%-4\%}$	3.35×10^4
$WZ(\to 3\ell)$+jets	$118.8 \times 10^{2+5\%+2\%}_{-5\%-2\%}$	3.45×10^2

Table 8: Production cross sections of the main irreducible background processes, and their simulated integrated luminosity. The reported value is the product of the total cross section times the branching ratio for leptonic decays of the vector bosons except for $t\bar{t}W$, where the inclusive final state is considered. The per-cent uncertainties reported are relative to the QCD factorisation and renormalisation scales (left), and the choice of PDF and α_s (right).

The $t\bar{t}W$ and $t\bar{t}Z$ processes are also simulated at NLO precision in perturbative QCD based on the procedure described in [128], using MadGraph5_aMC@NLO, with NNPDF 3.0 PDFs. For $t\bar{t}W$, the top quark and W boson masses are fixed to 172.5 GeV and 80.3 GeV, respectively. The matrix element calculation includes NLO electroweak corrections as well [135]. The choice for the central value of the renormalisation and factorisation scales is $\mu_0 = \mu_R = \mu_F = m_t + m_W/2$. The systematic uncertainties from scale variations and PDFs are estimated using the same procedure as described previously for the $t\bar{t}H$ sample: the former are found to be of order 12%, the latter 3.4%. The cross section with its uncertainties is shown in Table 8. For $t\bar{t}Z$, the matrix element accounts for both on-shell Z and off-shell Z/γ^* contributions, with the requirement of $m_{\ell\ell} > 5$ GeV. A dedicated Z/γ^* sample for the low dilepton invariant mass phase space is generated as well, to account for events with highly asymmetric p_T of the leptons. Size of the systematic uncertainties from scale variations and PDFs is very similar to the $t\bar{t}W$ case.

Diboson events - including $WZ(\to 3\ell)$ - are generated at NLO accuracy in QCD with radiation of up to three extra partons in the final state using Sherpa v2.1.1 [136], which is also used to model the parton showering. The PDFs employed are the CT10 [137] set. Samples are classified by the number of leptons in the final state. The estimated uncertainties on the cross section calculated inclusively in the number of additional radiated partons are of order 5% for scale variations, 2% for the PDF choice [126].

Although fakes are modelled directly from data, $t\bar{t}$ simulation is still exploited for background validation. Events are generated with Powheg-BOX v2.0 [138] and interfaced with Pythia 8 for the parton showering and frag-

mentation. The A14 tune is used for the underlying event modelling. The same set of programs is also used to model other top backgrounds such as single top t-channel and s-channel and Wt. The heavy flavour hadron decays are modelled by EVTGEN 1.2.0 [139]. Additional (fake) leptons can enter in the final state through material interactions of photons produced as initial/final state emission, or radiation within the parton showering process. These two effects are simulated inclusively in the $t\bar{t}$ sample. An additional sample has been generated to specifically target $t\bar{t}\gamma$ production, where a real, prompt photon is radiated by any of the top quarks as part of the ME calculation. The ME generator used in this case is MADGRAPH5_AMC@NLO; similarly to the $t\bar{t}$ sample, PYTHIA 8 models the parton showering and fragmentation. In order to avoid double-counting of events between the $t\bar{t}$ and $t\bar{t}\gamma$ samples, an overlap removal procedure is implemented. As a result, events in which the photon originates from the ME calculation are kept in the $t\bar{t}\gamma$ sample, and events in which the photon is radiated in the parton shower or as initial/final state radiation are kept in $t\bar{t}$ instead.

Table 9 presents a summary of all Monte-Carlo generators and configurations for the samples used in this analysis. This includes also rare SM backgrounds - namely triboson production, $t\bar{t}t\bar{t}$, $t(Z/\gamma^*)$, $tW(Z/\gamma^*)$, $t\bar{t}W^+W^-$ - and processes populating the control regions, such as $Z \to \ell\ell$ + jets and $W \to \ell\nu$ + jets. The tH process is simulated as well, and it is treated as a background.

4.4 PHYSICS OBJECTS SELECTION

The event categorisation of the $t\bar{t}H \to \text{multi} - \text{leptons}$ analysis as sketched in Figure 42 relies on a common definition of physics objects. This is optimised based on studies on simulation, and follows the general guidelines of the central ATLAS combined performance (CP) groups. After the categorisation, a dedicated fake suppression selection for electrons and muons is used.

4.4.1 Jets and b-tagged jets

Jets are reconstructed and calibrated according to the procedure described in subsection 2.3.4. Since the $t\bar{t}H$ final state is very heavy, particles produced in the top and Higgs decays are expected to be produced in the central part of the detector. Consequently, the jet pseudorapidity range is restricted to

Process	ME Generator	Parton Shower	PDF	UE/PS Tune
$t\bar{t}H$	MG5_aMC	Pythia 8	NNPDF 3.0 NLO	A14
$t\bar{t}W$	MG5_aMC	Pythia 8	NNPDF 3.0 NLO/2.3 LO	A14
$t\bar{t}(Z/\gamma^*)$	MG5_aMC	Pythia 8	NNPDF 3.0 NLO/2.3 LO	A14
VV	Sherpa 2.1.1	Sherpa	CT10	Sherpa
$t\bar{t}$	Powheg-BOX	Pythia 8	CT10/CTEQ6L1	A14
$t\bar{t}\gamma$	MG5_aMC	Pythia 8	NNPDF 2.3 LO	A14
$t(Z/\gamma^*)$	MG5_aMC	Pythia 6	CTEQ6L1	Perugia2012
$tW(Z/\gamma^*)$	MG5_aMC	Pythia 8	NNPDF 2.3 LO	A14
$t\bar{t}t\bar{t}$	MG5_aMC	Pythia 8	NNPDF 2.3 LO	A14
$t\bar{t}W^+W^-$	MG5_aMC	Pythia 8	NNPDF 2.3 LO	A14
VVV	Sherpa 2.1.1	Sherpa	CT10	Sherpa
$tHqb$	MG5_aMC	Pythia 8	CT10	A14
tHW	MG5_aMC	Herwig++	CT10/CTEQ6L1	UE-EE-5
Wt, single top	Powheg-BOX	Pythia 6	CT10/CTEQ6L1	Perugia2012
$Z \to \ell^+\ell^- +$ jets	Sherpa 2.2.1	Sherpa	NNPDF 3.0 NNLO	Sherpa
$W \to \ell\nu +$ jets	Sherpa 2.2.1	Sherpa	NNPDF 3.0 NNLO	Sherpa

Table 9: Configurations used for the simulation of signal and background processes relevant to this analysis. The bottom part of the table shows simulated samples which are relevant to control regions only. If only one parton distribution function (PDF) is shown, the same one is used for both the matrix element (ME) and parton shower generators; if two are shown, the first is used for the matrix element calculation and the second for the parton shower. "V" refers to production of an electroweak boson (W or Z/γ^*). "MG5_aMC" refers to MadGraph5_aMC@NLO 2.2.1; "Pythia 6" refers to version 6.427; "Pythia 8" refers to version 8.2; "Herwig++" refers to version 2.7 [140].

the ID acceptance of $|\eta| < 2.5$, where it is possible to exploit tracking information to reject jets from pile-up and to identify jets from b quarks. Jets must satisfy $p_T > 25$ GeV, and for candidates with $p_T < 60$ GeV and $|\eta| < 2.4$ the JVT pile-up suppression algorithm is employed. Jets containing b-hadrons are tagged using the MV2c10 algorithm with the 70% b-jet efficiency working point.

4.4.2 Hadronic taus

The dileptonic same-sign final state enforces a veto for objetcs identified as hadronically decaying tau leptons (τ_{had}). The event categorisation therefore depends indirectly on the selection of tau objects.

The tau reconstruction was briefly described in subsection 2.3.5. The kinematic selection applied requires candidates to have $p_T > 25$ GeV and $|\eta| < 2.5$, excluding the poorly instrumented EM calorimeter region $1.37 < |\eta| < 1.52$. A boosted decision tree (BDT) discriminant using calorimeter and tracking-based variables is used to identify τ_{had} candidates, and reject quark-initiated jets. The chosen "Medium" working point provides a combined efficiency for reconstruction and identification of 55% (40%) for one-prong (three-prong) τ_{had} decays. Rejection of electrons reconstructed as one-prong τ_{had} candidates is provided by a BDT trained on $Z \to ee$ events with a reconstructed and identified τ_{had} . A working point that is 95% efficient for true τ_{had} candidates is used. The rejection factor for electrons is of order 30-100, depending on η and p_T. An angular separation of $\Delta R > 0.2$ between τ_{had} candidates and any reconstructed electron and muon with $p_T > 2$ GeV is also required. Candidates can also be faked by b-tagged jets. To reduce this, the τ_{had} is matched to the jet that is geometrically overlapping with it. If the overlapping jet is b-tagged with the chosen 70% WP, the τ_{had} candidate is rejected. Finally, in order to suppress tau fakes from pileup jets, the τ_{had} vertex is required to be the hard scattering vertex of the event.

4.4.3 Light leptons (e, μ)

The light lepton reconstruction has been previously outlined in subsection 2.3.2 for electrons, and in subsection 2.3.3 for muons.

The minimal transverse momentum requirement for electron candidates considered in the analysis is $p_T > 10$ GeV. The pseudorapidity coordinate of the reconstructed calorimetric EM cluster has to be within $|\eta_{\text{cluster}}| < 2.47$, excluding the badly instrumented region $1.37 < |\eta_{\text{cluster}}| < 1.52$. The

"Loose" working point of the electron ID likelihood-based algorithm is used to reject fake electrons from hadronic showers and converted photons in the ID material. In order to suppress fakes not belonging to the primary interaction, requirements on the electron candidates track impact parameters (IP) are applied. The track transverse impact parameter d_0 and longitudinal impact parameter z_0 are calculated with respect to the beam line, assuming a transversal size of the beamspot much smaller than the primary vertex position resolution. More effective discrimination is achieved by using the transverse impact parameter significance - defined as the unsigned impact parameter divided by its uncertainty, $|d_0|/\sigma_{d_0}$ - and the longitudinal impact parameter weighted by the sine of the track polar angle, $z_0 \sin\theta$. The latter definition helps to avoid rejecting good tracks in the forward region, where larger IP measurement errors are expected. For electrons, a selection of $|d_0|/\sigma_{d_0} < 5$ and $|z_0 \sin\theta| < 0.5$ mm is imposed.

Muon candidates are selected according to the "Loose" working point, which includes segment-tagged and calorimeter-tagged candidates, as well as combined muons. Candidates are required to satisfy $p_T > 10$ GeV and $|\eta| < 2.5$. The transverse impact parameter significance requirement for muon candidates is $|d_0|/\sigma_{d_0} < 3$, whereas the longitudinal impact parameter selection is the same as in the electron case.

4.4.4 Overlap removal

Despite satisfying the basic selection requirements outlined above, a single physics object could still be reconstructed as two different objects. In order to avoid double counting, an overlap removal procedure is employed to resolve ambiguities in the order described henceforth. Any electron candidate within $\Delta R = 0.1$ of another electron with higher p_T is removed, and the same happens if it is found within $\Delta R = 0.1$ of a muon candidate. Subsequently, any jet within $\Delta R = 0.3$ of an electron candidate is removed in favour of the electron. If a muon candidate and a jet lie within $\Delta R = \min(0.4, 0.04 + 10/p_T(\mu)\ [\text{GeV}])$ of each other, the jet is kept and the muon is removed; the cut value is optimised to maximise the acceptance of real muons at a fixed rejection factor for non-prompt muon candidates originating from hadron decays within the jet. Hadronic tau candidates within $\Delta R = 0.2$ of an electron or muon candidate are rejected in favour of the light leptons, and any jet within $\Delta R = 0.3$ of a τ_{had} candidate is removed in favour of the τ_{had} . The full overlap removal is schematised in Table 10.

Keep	Reject	ΔR cone size
e	e (low p_T)	0.1
μ	e	0.1
e	jet	0.3
jet	μ	$\min(0.4, 0.04 + 10/p_T(\mu)$ [GeV])
e	τ_{had}	0.2
μ	τ_{had}	0.2
τ_{had}	jet	0.3

Table 10: Summary of the overlap removal procedure between electrons, muons, hadronically decaying taus, and jets when more than one are found in a ΔR cone of the size indicated.

4.4.5 MVA-based selection of prompt electrons and muons

The baseline selection on light leptons and hadronic taus just presented encompasses the minimal set of requirements for the analysis categorisation to ensure orthogonality between all categories.

A tighter lepton selection is then applied in the 2ℓ SS $0\tau_{had}$ category, since the baseline lepton selection is too loose to effectively suppress the large fake backgrounds contaminating the 2ℓ SS $0\tau_{had}$ final state. Two independent multi-variate algorithms have been developed, specifically targeted at non-prompt lepton and charge flip suppression, respectively. A selection based on the output of such algorithms is eventually applied on top of the baseline one to define signal-like, *tight* lepton objects.

Non-prompt fakes rejection

In order to define tight leptons, both the Run1 and early Run 2 analyses rely on lepton *isolation* variables, which measure the activity around the reconstructed leptons. Two types of such variables can be defined, which differ only slightly between electrons and muons:

- *Calorimetric isolation*, based on the sum of the transverse energies of the calorimeter topological clusters - excluding the lepton candidate cluster itself - within a $\Delta R = 0.2$ cone around the lepton candidate.
- *Track isolation*, based on the sum of the transverse momenta of the tracks consistent with the primary vertex, excluding the lepton candidate track, within a ΔR cone whose size is typically inversely proportional to the lepton p_T for high momentum candidates. This is defined as $\Delta R = \min(X, 10/p_T$ [GeV]). The size X is 0.2 for electrons and 0.3 for muons.

These variables alone are not specifically designed to suppress non-prompt fakes within jets. In fact, even after the application of very tight isolation requirements, such fakes represent the dominant background in most categories of the Run 1 and early Run 2 analyses, and are the main limiting factor to the sensitivity [121, 124].

A key feature that can be exploited is the fact that heavy flavoured hadrons resulting from hadronisation of b, c quarks typically have longer lifetimes such that they decay away from the primary vertex, at a detectable distance. This represents the main idea behind the implementation of the `PromptLeptonIso` [141] algorithm used in the Run 2 analysis to reject non-prompt fakes. This algorithm determines the lifetime information by extracting features of all the tracks making up the jet candidate that matches the selected light lepton, instead of relying exclusively on the lepton object properties. Such an approach increases the precision of identifying the displaced decay vertex of bottom or charm hadrons which produced a non-prompt light lepton.

A set of jet and lepton-related information is combined into a boosted decision tree discriminant. In total, eight variables are employed to train the BDT on simulated $t\bar{t}$ events. Two of such variables use the relationship between the jet and lepton: the ratio of the lepton p_T with respect to the jet p_T as resulting from its tracks, and ΔR between the lepton and the jet axis. Three of the input variables are the output of the jet flavour tagging algorithms used in ATLAS to identify b and c decays. Finally, three variables describe the lepton isolation: the number of tracks within the jet, and the standard lepton track and calorimeter isolation variables mentioned earlier, divided by the lepton p_T.

Charge flip electrons rejection

In the 2ℓ SS $0\tau_{had}$ categories with electrons of the early Run 2 analysis, the charge flip background is as large as 30% of the total. A dedicated algorithm known as `QMisIDKiller` [142] has been developed to specifically suppress this background. It combines eight electron variables in a boosted decision tree: the p_T and η of the electron, the track curvature significance - defined as the ratio between the electric charge and the track momentum divided by the estimated uncertainty on the measurement - and its impact parameter times the electric charge, the cluster width along the azimuthal direction, and the quality of the matching between the track and the cluster, both in terms of energy/momentum and azimuthal position.

Electrons		
	Loose	*Tight*
$\lvert d_0^{sig} \rvert$	5	5
$\lvert z_0 \sin\theta \rvert$	0.5 mm	0.5 mm
Electron ID	Loose LH	Tight LH
Isolation	—	Loose WP
Non-prompt rejection	—	`PromptLeptonIso` < -0.5
Charge flip rejection	—	`QMisIDKiller` > 0.067

Muons		
	Loose	*Tight*
Quality	Loose	Loose
$\lvert d_0^{sig} \rvert$	3	3
$\lvert z_0 \sin\theta \rvert$	0.5 mm	0.5 mm
Isolation	—	Loose WP
Non-prompt rejection	—	`PromptLeptonIso` < -0.5

Table 11: Definition of baseline loose leptons used for the categorisation, and tight leptons used for the 2ℓ SS $0\tau_{had}$ signal region.

Tight lepton definition, and calibration in data

The tight leptons used in the signal region are defined by requirements on the `PromptLeptonIso` and `QMisIDKiller` scores, combined with tight particle ID (for electrons only) and a lepton isolation requirement designed to achieve 99% efficiency for prompt leptons in $Z \to \ell^+\ell^-$ events. The latter is referred to as "Loose" isolation working point, and it exploits the calorimeter and track-based isolation variables described above. The tight lepton definition, in comparison with the baseline loose selection used for the analysis categorisation, is summarised in Table 11.

The efficiency of the tight lepton selection is measured in control regions enriched in $Z \to \ell^+\ell^-$ events using the technique described in [87, 143]. Figure 48 illustrates the efficiency for tight muons and electrons as a function of p_T. The efficiency for muons (electrons) is of order 70% (60%) for $p_T \approx$ 10 GeV, and reaches a plateau of 96% (94%) at $p_T \approx 45$ GeV. The chosen `PromptLeptonIso` working point gives a rejection factor against non-prompt leptons from heavy flavoured hadron decays of about 20. With this selection, the signal acceptance is twice as large as the one reached in the early Run 2 analysis for a similar signal over background ratio. The rejection factor

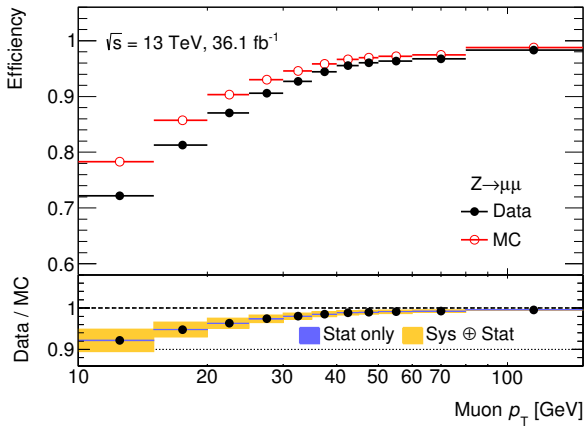
for electrons with a wrong charge assignment owing to the `QMisIDKiller` selected working point is ≈ 17.

The efficiency of the tight selection in simulation is matched to the one measured in data via scale factors parametrised as a function of the lepton p_T and η. These correction factors are approximately 0.92 for low transverse momentum muons, and average at 0.98-0.99 for higher p_T muons, with similar performance achieved for electrons. The larger corrections at low p_T are to be ascribed to a poor description of the vertex density (i.e., the number of vertices along beam direction per mm) as observed in 2016 data by the ATLAS simulation [144]. No sizeable dependency of the scale factors on other observables such as N_{jets}, $\Delta R(\ell, jet)$ is seen. A significant decrease of the scale factor as a function of the number of vertices per event is observed instead. This effect is covered by adding a systematic uncertainty. The total size of the systematic uncertainties on the scale factors ranges between $\approx 3\%$ at low transverse momentum to less than 1% at high p_T.

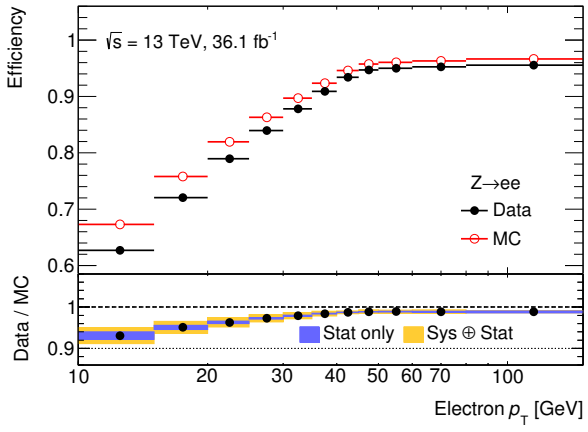
4.5 TRIGGER SELECTION

An event pre-selection based on trigger requirements common to all the $t\bar{t}H$ multi-leptonic categories is applied.

The triggers used in this analysis either require the presence of a single, isolated and well identified electron or muon (*single lepton triggers*, SLT), or a pair of light leptons with same or opposite flavour (*dilepton triggers*, DLT). The full set of triggers used is outlined in Table 12. The minimal online p_T thresholds of the trigger chains - as well as the online object selections - vary between 2015 and 2016, in order to cope with the changes in the LHC operating conditions as summarised in Table 4. For the single muon trigger, the harsher 2016 conditions required increasing the minimal p_T threshold from 20 to 26 GeV to keep the rates within the available bandwidth, as well as tightening the online isolation requirement from *loose* to *medium*. Similarly, for the single electron trigger the online p_T threshold is raised from 24 to 26 GeV, and the likelihood-based electron ID (subsection 2.3.2) is tightened from *medium* to *tight*. In 2016, a loose level of isolation at HLT is also required. For both single electrons and muons, the actual triggers are a logical `OR` combination of several chains with increasing p_T thresholds and looser selections applied, in order to preserve high trigger efficiency where the rate is lower and a strong fake lepton rejection is not needed. As for the dilepton



(a) μ



(b) e

Figure 48: The efficiency in data (black) and simulation (red) of the tight muon (48a) and electron (48b) selection as a function of p_T.

	Single lepton triggers	
	2015	2016
μ	($p_T > 20$ GeV AND loose isolation) OR ($p_T > 50$ GeV)	($p_T > 26$ GeV AND medium isolation) OR ($p_T > 50$ GeV)
e	($p_T > 24$ GeV AND medium ID) OR ($p_T > 60$ GeV AND medium ID) OR ($p_T > 120$ GeV AND loose ID)	($p_T > 26$ GeV AND tight ID AND loose isolation) OR ($p_T > 60$ GeV AND medium ID) OR ($p_T > 140$ GeV AND loose ID)
	Dilepton triggers	
	2015	2016
$\mu\mu$	($p_T > 18$ GeV (μ_1) AND $p_T > 8$ GeV (μ_2))	($p_T > 22$ GeV (μ_1) AND $p_T > 8$ GeV (μ_2))
ee	($p_T > 12$ GeV AND loose ID (e_1, e_2))	($p_T > 17$ GeV AND very loose ID (e_1, e_2))
$e\mu, \mu e$	($p_T > 17$ GeV AND loose ID (e_1) AND $p_T > 14$ GeV (μ_2))	($p_T > 17$ GeV AND loose ID (e_1) AND $p_T > 14$ GeV (μ_2))

Table 12: List of single lepton and dilepton triggers used during the 2015 and 2016 data taking. The trigger selection includes the online p_T threshold for the corresponding physics object, as well as lepton isolation and identification requirements, if any. The indexes "1, 2" indicate leading and sub-leading p_T leptons, respectively.

triggers, the changes between the 2015 and 2016 involve almost exclusively the online p_T cuts.

Events are selected by a logical OR combination of single lepton and dilepton triggers. This choice improves the signal acceptance by as much as 8% with respect to trigger strategy used in the previous $t\bar{t}H$ analyses, which exploited single lepton triggers only. This is illustrated in Table 13 for the 2ℓ SS $0\tau_{had}$ case. For the SLT OR DLT combination, the trigger efficiency for leptons passing the offline 2ℓ SS $0\tau_{had}$ event selection is found to be within 82 to 99 %, depending on the lepton flavour of the final state.

$t\bar{t}H$ acceptance ratio	$\mu\mu$	$e\mu, \mu e$	ee
DLT/SLT	1.04 ± 0.04	0.69 ± 0.02	0.74 ± 0.04
(DLT OR SLT)/SLT	1.08 ± 0.04	1.06 ± 0.03	1.05 ± 0.05

Table 13: Comparison of $t\bar{t}H$ signal acceptance (with respect to using single lepton triggers (SLT) only) for dilepton triggers (DLT), and a logical OR of single and dilepton triggers. Uncertainties are statistical only.

2ℓ SS $0\tau_{had}$ signal region	
N_ℓ	2, passing T (11)
$\sum Q_\ell$	± 2
$p_T(\ell)$	≥ 20 GeV
N_{jets}	≥ 4
N_{b-tags}	1 OR 2
$N_{\tau_{had}}$	0

Table 14: Definition of the 2ℓ SS $0\tau_{had}$ signal region.

4.6 SIGNAL REGION DEFINITION AND EVENT CLASSIFICATION

In the 2ℓ SS $0\tau_{had}$ category, signal-like events are selected by requiring the presence of exactly two same-sign light leptons satisfying the tight selection, with a p_T above 20 GeV. A lower threshold on the lepton transverse momentum has shown no improvement in sensitivity, in consideration of the larger amount of fake background introduced. A minimum of 4 jets is demanded, of which one or two must be b-tagged according to the 70% efficiency working point of the MV2c10 algorithm. This selection takes into account limited acceptance and inefficiencies, given an expected number of 6 jets (with two b-jets). Finally, a veto on the presence of hadronic taus is also enforced.

This set of requirements defines the 2ℓ SS $0\tau_{had}$ *signal region*. A summary of the full selection is in Table 14.

Based on the previous description, both reducible and irreducible background events produce very similar signatures to the $t\bar{t}H$ process. However, differences in jet multiplicity and kinematic features of the final state with respect to the signal are expected.

Due to the limited statistics of the selected events sample in the previous analyses, a cut-and-count approach based on a sequence of binary selections on weakly separated observables was followed. Taking advantage of the increased Run 2 statistics, multivariate statistical analysis (MVA) algorithms

based on machine learning techniques can be exploited for classification purposes. For the 2ℓ SS $0\tau_{had}$ analysis, a set of two such algorithms is trained to classify signal region events to discriminate $t\bar{t}H$ against the fake lepton background and the main irreducible backgrounds, respectively.

The class of algorithm used is a gradient boosted decision tree (BDTG) [145], as implemented in the TMVA [146] framework. BDTGs consist of a set of binary tree-structured classifiers sequentially constructed from re-weighted events of the same training sample, eventually averaged with weights calculated according to each tree's accuracy. The choice of the weak classifiers used in the combination is based upon minimisation of the loss of the model via a gradient descent-like procedure. This technique - known as *gradient boosting* - has been shown to enhance the classification performance with respect to a single decision tree, stabilise it against statistical fluctuations in the training sample and be less prone to overtraining [147]. Furthermore, BDTGs can deal with correlations between input variables. This improves the classification even when relying on observables with low separation power - such as basic kinematic features of the events. Our analysis indeed represents a typical scenario for exploiting such algorithms, given the typical size of the training sample of only few tens of thousands of raw events (see Table 15). An alternative analysis approach has been developed as a cross-check to the BDTG-based analysis here presented, based on a simple cut-based sub-categorisation of the SR in terms of jet and b-jet multiplicity and lepton flavour composition in a similar fashion to the previous analyses. The final sensitivity obtained is about 30% poorer than the one achieved with the BDTGs.

A first BDTG algorithm ($BDTG_{t\bar{t}V}$) is trained to separate the $t\bar{t}H$ signal from the irreducible $t\bar{t}V$ processes (i.e., the combination of $t\bar{t}W$ and $t\bar{t}Z$), and a second independent classifier ($BDTG_{t\bar{t}}$) is trained to discriminate signal against the fake lepton background. The nominal $t\bar{t}H$, $t\bar{t}W$, $t\bar{t}Z$ simulated samples used for the analysis are also employed for the training of $BDTG_{t\bar{t}V}$. The training of $BDTG_{t\bar{t}}$ is based on the $t\bar{t}H$ sample and a fake-enriched sample of data events with leptons failing the tight selection. Such events with loose leptons are also used to estimate the fakes background of the analysis, as it will be described in more detail in the next section.

Both classifiers use nine input variables, computed in the laboratory frame:

1. Number of jets: N_{jets} (events with more than 7 jets are treated as if $N_{jets} = 7$).

2. Number of b-tagged jets: N_{b-tags}.

Nr. events	Raw	Weighted
$t\bar{t}H$	41879	42.6 ± 4.2
$t\bar{t}W$	50329	123 ± 18
$t\bar{t}Z$	41368	41.4 ± 5.6
Fakes	8784 (*)	233 ± 39

Table 15: Number of raw and weighted events for the signal and background samples used for the BDTG training. For $t\bar{t}H$, $t\bar{t}W$, $t\bar{t}Z$, weighted events correspond to the number of raw events rescaled by the product of correction factors on reconstruction, identification... efficiencies times the ratio of integrated luminosities of the data and generated samples. For fakes, (*) indicates the total number of selected collision data events where the two leptons are required to pass the baseline loose selection only - see subsection 4.8.2. The weighted number of fake events is computed also according to subsection 4.8.2. Half of each event set is used for training, and the other half is used for testing the performance. Cross-validation is done as well, using even/odd event classification.

3. Lepton flavour composition of the event (ee, $e\mu$ μe, $\mu\mu$, based on p_T ranking): $lFlav = 2 \times N_{\mu_0} + N_{\mu_1}$ ($N_{\mu_i} = 1$ if i-th lepton in the pair is a muon, otherwise zero).
4. Distance between leading lepton and its closest jet: $\min(\Delta R(\ell_0, \text{jet}))$.
5. Distance between sub-leading lepton and its closest jet: $\min(\Delta R(\ell_1, \text{jet}))$.
6. Maximum absolute pseudorapidity in the dilepton pair: $\max(|\eta_\ell|)$.
7. Sub-leading lepton transverse momentum: $p_T(\ell_1)$.
8. Missing transverse energy: E_T^{miss}.
9. Invariant mass of the dilepton pair: $m_{\ell_0\ell_1}$.

These variables were found to have the highest separation power. They are ranked from top to bottom in the list by counting how often each variable is used to split decision tree nodes, and by weighting each split occurrence by the separation gain-squared it has achieved and by the number of events in the node. The addition of further input variables is not seen to improve the performance of the classifiers. Despite the charge asymmetry of the $t\bar{t}W$ process at the LHC, the sum of lepton charges in the event $\sum Q_\ell$ has not been found to improve the final sensitivity, therefore it has not been included. For both classifiers, N_{jets} is the most powerful variable, which is expected given the higher jet multiplicity of the $t\bar{t}H$ final state with respect to $t\bar{t}V$ and $t\bar{t}$. Not surprisingly, the largest correlation is found between $p_T(\ell_1)$ and $m_{\ell_0\ell_1}$, and it is of order 60%. All the other correlations are below 20%.

The input event sample is divided into orthogonal subsets for training the algorithm and testing its performance, based on the parity of the `EventNumber` flag associated to each simulated event. For cross-validation purposes, for each of the two classifiers one BDTG is trained on even events and used to classify odd events, and the opposite is done for a second BDTG. The BDTG response function distribution of the $BDTG_{t\bar{t}V}$ and $BDTG_{t\bar{t}}$ classifiers is represented in Figure 49. A good agreement between the shape of training and test samples is observed, indicating absence of overtraining.

The discriminating power of $BDTG_{t\bar{t}V}$ and $BDTG_{t\bar{t}}$ is shown in Figure 49e, as the area under the curve representing the background rejection as a function of signal efficiency. The performance achieved for the fake background case is sizeably better than in the $t\bar{t}V$ case, as expected from the greater similarity of the $t\bar{t}V$ and $t\bar{t}H$ final states. Similar performance is achieved for even/odd event sets, proving the classification is robust.

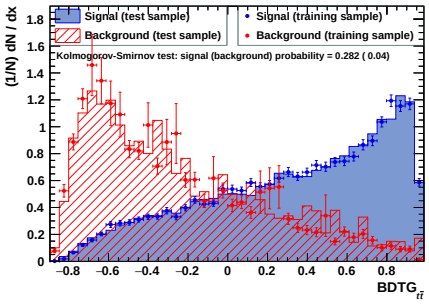
As it will be described in the next chapter, the outputs of the two BDTGs are eventually combined and used as a single discriminating variable in a template fit to the observed data to extract the $t\bar{t}H$ signal strength, $\mu_{t\bar{t}H} = \sigma/\sigma_{SM}$, that is, the $t\bar{t}H$ production cross section normalised to the SM value under the assumption of SM Higgs decay branching ratios.

4.7 IRREDUCIBLE BACKGROUND VALIDATION

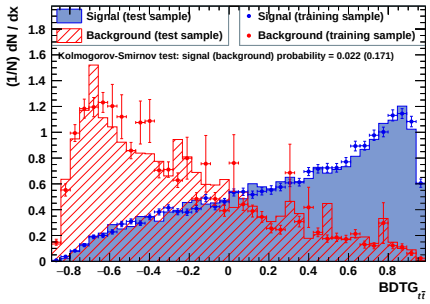
The irreducible backgrounds of this analysis - principally $t\bar{t}W$, $t\bar{t}Z$ and WZ - are directly estimated from simulation, similarly to the $t\bar{t}H$ signal. The validity of the background modelling is checked in validation regions, defined to be sufficiently close to the signal region event topology. All the processes contributing to the validation regions are based on simulation except for fakes, which are modelled via the matrix method as detailed in subsection 4.8.2.

4.7.1 Modelling of $t\bar{t}W$

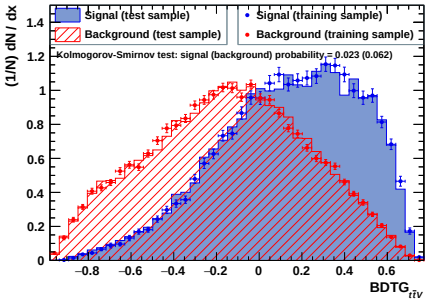
For the 2ℓ SS $0\tau_{had}$ category, $t\bar{t}W$ represents the dominant irreducible background. Only weak constraints on the cross-section measurement of this process - of order 30% - have been placed so far by ATLAS and CMS [148, 149]. It is therefore essential to ensure the prediction of its rate and kinematics from simulation is reliable.



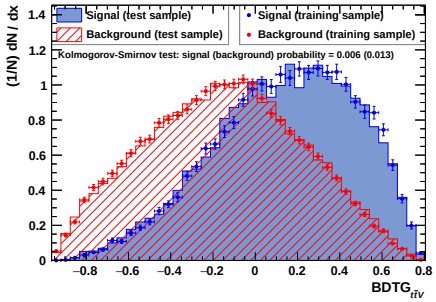
(a) $t\bar{t}H$ vs. DD fakes (even)



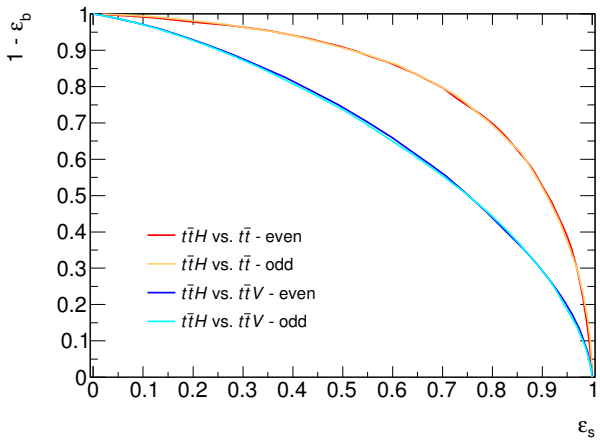
(b) $t\bar{t}H$ vs. DD fakes (odd)



(c) $t\bar{t}H$ vs. $t\bar{t}V$ (even)



(d) $t\bar{t}H$ vs. $t\bar{t}V$ (odd)



(e) ROC

Figure 49: Distribution of the BDTG response function for 49a, 49b $BDTG_{t\bar{t}}$ and 49c, 49d $BDTG_{t\bar{t}V}$ classifiers for training and test subsets (with even/odd events cross-validation), and 49e their classification performance as from the ROC (Receiver Operating Characteristic) curves showing background rejection versus signal efficiency.

Two validation regions are tested. One of them relies on a cut-based event selection, reaching a purity of 45%. An alternative validation region is defined in the two-dimensional $(BDTG_{t\bar{t}}, BDTG_{t\bar{t}V})$ plane by selecting events in a corner of phase space with maximal purity of $t\bar{t}W$ events. A similar value (42%) as for the cut-based region is obtained, but with four times more events. The two defintions are summarised in Figure 50. Despite that in both cases the selected events belong to the 2ℓ SS $0\tau_{had}$ signal region as well, the overall signal contamination is small, accounting for approximately 6% in both regions.

$t\bar{t}W$ VR selection
2ℓ passing T (11)
$p_T^{\ell} > 30$ GeV
$N_{jets} = 4$
$N_{b-tags} = 2$
$H_T^{jets} > 220$ GeV (ee,OF only)
$m_{ee} \notin [75,105]$ GeV
$E_T^{miss} > 50$ GeV (ee only)
$\sum Q_{\ell} = +2$

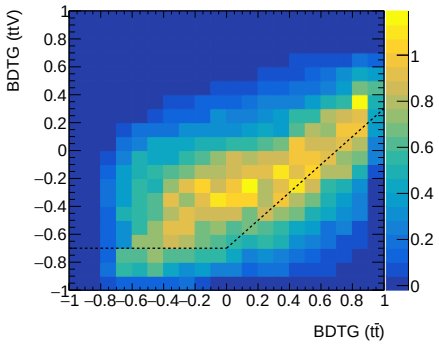


Figure 50: Definition of the validation regions (VR) for the $t\bar{t}W$ process. The H_T^{jets} variable corresponds to the scalar sum of the transverse momenta of all jets in the event and "OF" stands for "opposite-flavour'. The plot shows the distribution of $t\bar{t}W$ events in the two-dimensional $(BDTG_{t\bar{t}}, BDTG_{t\bar{t}V})$ plane: the area below the black dashed line - defined as $(BDTG_{t\bar{t}V} - BDTG_{t\bar{t}} < -0.7$ OR $BDTG_{t\bar{t}V} < -0.7)$ - represents the $t\bar{t}W$ validation region.

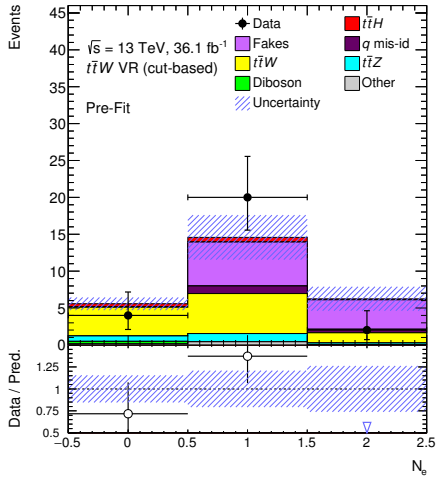
Distributions of some of the input variables of the BDT discriminators, namely the event lepton flavour composition (parametrised as the number of electrons per event), $p_T(\ell_1)$, $E_T^{miss}$and $m_{\ell\ell}$ are displayed in Figure 51. In the plots, the $t\bar{t}H$ yield is fixed to the SM expectation. A good agreement between data and the model is seen in all cases. In both regions, the total expected event yield is fitted to the observed data to extract the $t\bar{t}W$ strength $\mu_{t\bar{t}W}$, assuming the rates of all the other processes fixed to the SM prediction. The fit procedure is identical to the one employed for measuring the $t\bar{t}H$ signal strength ($\mu_{t\bar{t}H} = \sigma/\sigma_{SM}$) as described in the next chapter (see section 5.1), except for having the $t\bar{t}W$ cross section modifier $\mu_{t\bar{t}W}$ as the parameter of interest of the fit. A best-fit value of $\mu_{t\bar{t}W} = 1.28 \pm 0.58$ is obtained in the cut-based validation region, and a value of $\mu_{t\bar{t}W} = 1.35 \pm 0.43$ in the MVA-based region. The quoted uncertainties are statistical plus systematic (for the latter, refer to section 4.9). In both cases, the result is compatible with the SM expectation of 1 within 1σ, strengthening our confidence in the mod-

elling of the $t\bar{t}W$ process provided by simulation. In the final fit for $\mu_{t\bar{t}H}$, no correction is applied to the normalisation of the $t\bar{t}W$ simulated sample.

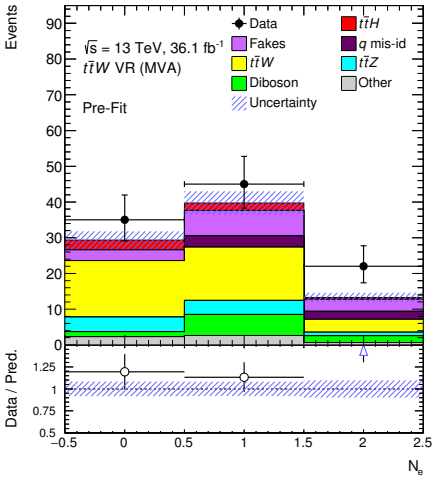
4.7.2 Modelling of $t\bar{t}Z$ and WZ

To test the modelling of the $t\bar{t}Z$ process, a trileptonic validation region is defined, requiring at least 4 jets in the event, of which more than one must be b-tagged, and for the presence of at least one pair of same-flavour, opposite-sign leptons with an invariant mass within a 10 GeV-wide window around the Z pole. As for the lepton composition and classification, the absolute value of the total charge must be equal to one; ℓ_0 indicates the lepton with opposite-sign charge to the other two in the triplet. Of the remaining same-charge leptons, ℓ_1 indicates the one closest in ΔR to ℓ_0, and ℓ_2 the furthest one. The ℓ_0 has very low probability of being a fake from $t\bar{t}$, therefore only loose isolation is applied on it, whereas both the SS leptons must pass the tight selection of Table 11. An invariant mass of $m_{\ell\ell} > 12$ GeV for any same-flavour (SF), opposite-sign (OS) pair is required to suppress heavy flavour resonances decays. The definition of the validation region is summarised in Table 16. The purity achieved for $t\bar{t}Z$ is around 85%. Figure 52 illustrates the distributions of N_e, N_{jets}, and the invariant mass of ℓ_0, ℓ_1. An adequate agreement is found, therefore no additional corrections on the pure simulation-based prediction are applied.

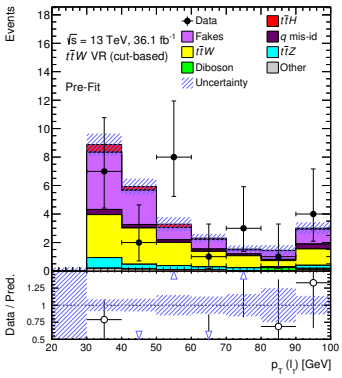
The diboson background for the 2ℓ SS $0\tau_{had}$ final state is associated mostly to the WZ process with associated radiation of gluons splitting into b-quarks, or c-jets mis-tagged as b. Although this background is sub-dominant with respect to $t\bar{t}W$ and $t\bar{t}Z$ in the 2ℓ SS $0\tau_{had}$ signal region, it yields a non-negligible contribution in the regions employed in the fake background estimation (section 4.8). Currently, no measurement of the cross section of this specific process is available. A cut-based trileptonic validation region for this process is also summarised in Table 16. The definition follows the $t\bar{t}Z$ one, except for loosening the requirement on the total jet multiplicity to allow for events with as little as two jets, and the presence of exactly one b-tagged jet. The agreement with data is good, as Figure 52 shows, even though the WZ purity achieved is only around 30%. Hence, a conservative uncertainty of 50% on its cross section is assigned.



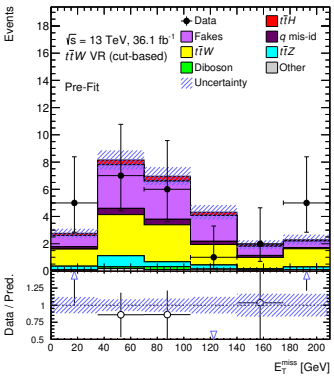
(a) Cut-based - N_e



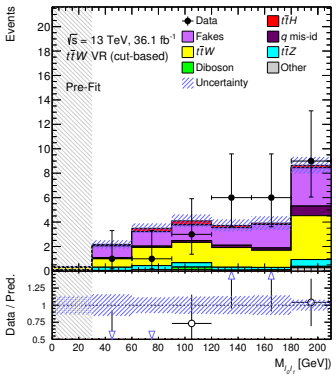
(b) MVA-based - N_e



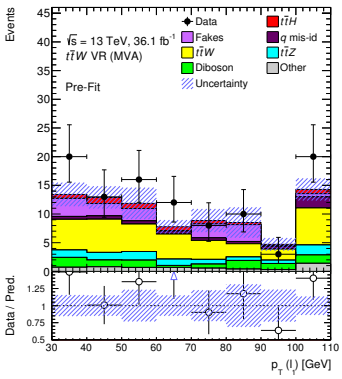
(c) Cut-based - $p_T(\ell_1)$



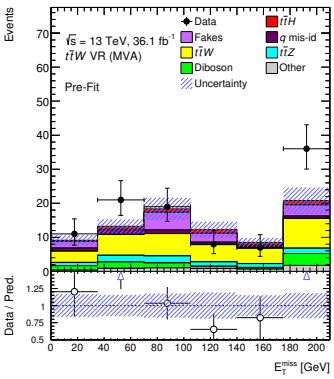
(d) Cut-based - E_T^{miss}



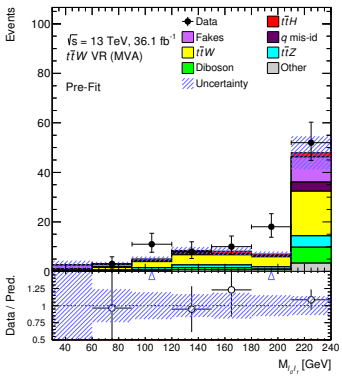
(e) Cut-based - $m_{\ell\ell}$



(f) MVA-based - $p_T(\ell_1)$

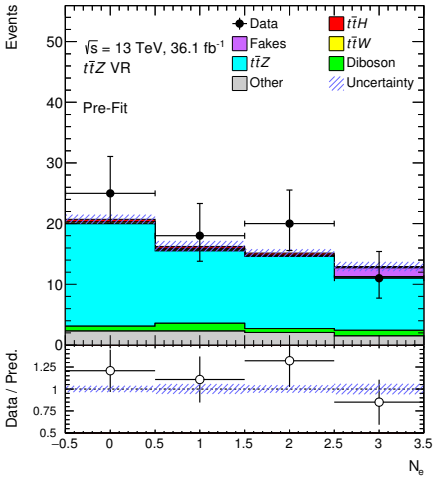


(g) MVA-based - E_T^{miss}

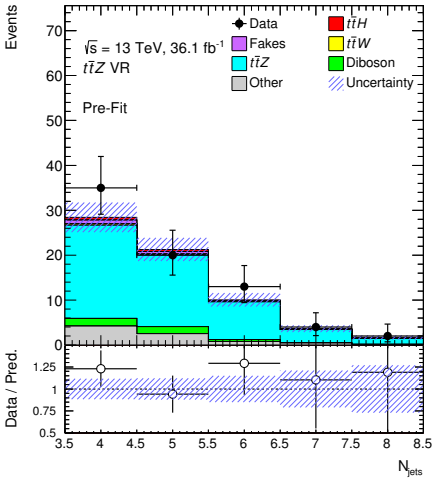


(h) MVA-based - $m_{\ell\ell}$

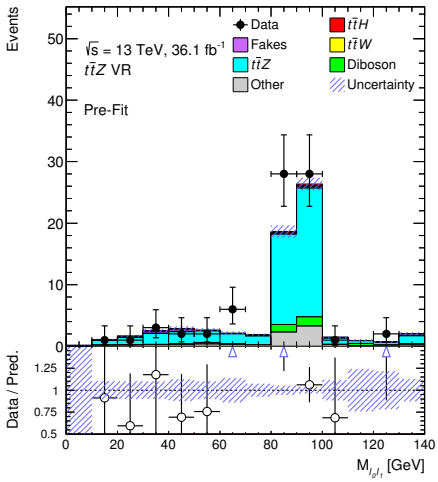
Figure 51: Distributions of the number of electrons, $p_T(\ell_1)$, $E_T^{miss}$and $m_{\ell\ell}$ for the cut-based and MVA-based $t\bar{t}W$ validation regions. Errors include statistical and systematic uncertainties (see section 4.9).



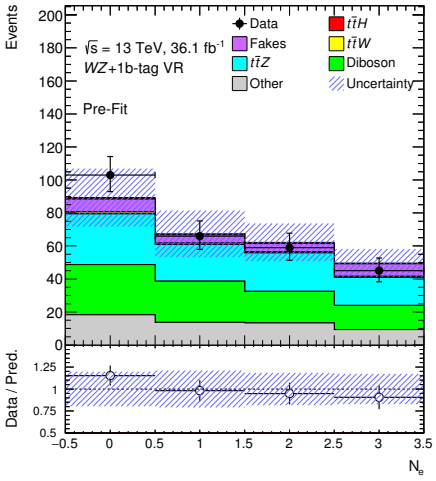
(a) $t\bar{t}Z$ VR - N_e



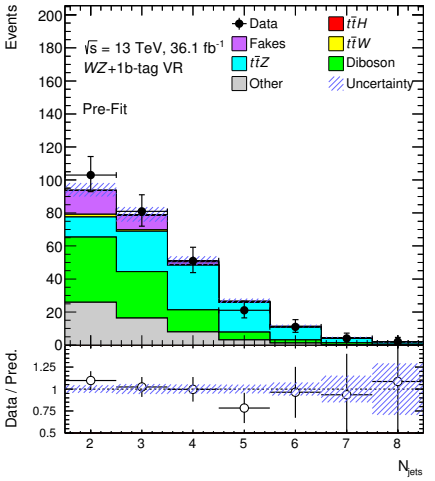
(b) $t\bar{t}Z$ VR - N_{jets}



(c) $t\bar{t}Z$ VR - $m_{\ell_0\ell_1}$



(d) WZ+ 1 b-tag - N_e



(e) WZ+ 1 b-tag - N_{jets}

Figure 52: Distributions of the number of electrons and jets for the cut-based $t\bar{t}Z$ and WZ+ 1 b-tag validation regions (VR). For the $t\bar{t}Z$ region, $m_{\ell_0\ell_1}$ is also displayed. Errors include statistical and systematic uncertainties (see section 4.9).

VR selection	
$t\bar{t}Z$	WZ+ 1 b-tag
$\sum Q_\ell \pm 1$	
OS ℓ passing Loose WP isolation	
SS ℓs passing T (11)	
$m_{\ell\ell} > 12$ GeV $\forall$ SF, OS pairs	
At least 1 SF, OS pair with $\lvert m_{\ell\ell} - 91.2 \rvert < 10$ GeV	
$N_{jets} \geq 4$	$N_{jets} \geq 2$
$N_{b-tags} \geq 2$	$N_{b-tags} = 1$

Table 16: Definition of the validation regions (VR) for the $t\bar{t}Z$ and WZ+ 1 b-tag processes.

4.8 REDUCIBLE BACKGROUND ESTIMATION

4.8.1 Estimation of the charge flip electrons background

When a true opposite-sign event such as $t\bar{t} \to (W^+ \to \ell^+\nu_\ell)b(W^- \to \ell^-\bar{\nu}_\ell)\bar{b}$ is produced, there are three possibilities for it to be reconstructed, assuming a finite probability ϵ of mis-identifying the charge of an electron[5].

- $e^+e^- + X$: opposite-sign, without any charge mis-identification, with a probability of $(1-\epsilon)^2$,
- $e^+e^- + X$: opposite-sign, where both electrons had their charge wrongly reconstructed , with a probability of ϵ^2,
- $e^\pm e^\pm + X$: same-sign, with only one of the two electrons charge mis-identified, with a probability of $2\epsilon(1-\epsilon)$.

Therefore, given N true opposite-sign events, the number of reconstructed opposite-sign and same-sign events will be:

$$\begin{aligned} N^{OS} &= \left(1 - 2\epsilon + 2\epsilon^2\right) N \\ N^{SS} &= 2\epsilon\left(1-\epsilon\right) N \end{aligned} \tag{24}$$

Knowing the charge mis-identification rate ϵ, it is possible to compute the estimated number of same-sign events N^{SS} satisfying a given selection from the measured number of opposite-sign events N^{OS} passing the same selection requirements, for the ee and $e\mu, \mu e$ (OF) categories, respectively:

5 We assume the probability of charge flip for muons is negligible

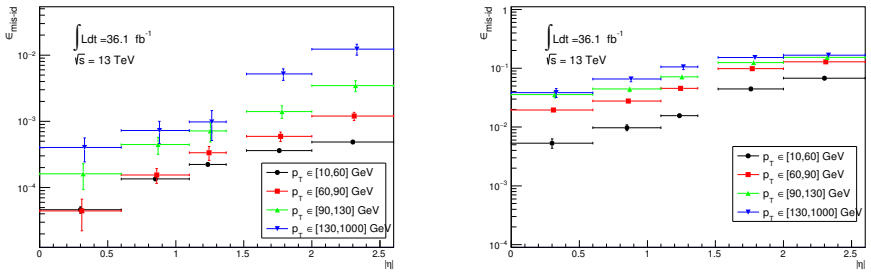
$$
\begin{aligned}
N_{ee}^{SS} &= \frac{\epsilon_i + \epsilon_j - 2\epsilon_i\epsilon_j}{1 - \epsilon_i - \epsilon_j + 2\epsilon_i\epsilon_j} N_{ee}^{OS} \\
&= w_{ee}^{QMisID} \times N_{ee}^{OS} \\
N_{OF}^{SS} &= \frac{\epsilon}{1-\epsilon} N_{OF}^{OS} \\
&= w_{OF}^{QMisID} \times N_{OF}^{OS}
\end{aligned}
\tag{25}
$$

where ϵ_i and ϵ_j are the charge mis-identification rates for each electron, assuming they depend on properties of each single object.

The rate of an electron being reconstructed with incorrect charge is measured using a likelihood method in $Z \to ee$ events around the Z peak [121]. The charge mis-identification rate is estimated in six bins in η and four bins in p_T, and two bins in electron quality, that is, estimating rates independently for electrons passing or failing the tight selection. The parametrisation in η is adopted since the probability of bremsstrahlung radiation emission increases with the amount of detector material crossed.

The bins are determined based on the available statistics, and the rate dependence on η and p_T. The background of events not consistent with the $Z \to ee$ process is subtracted using a sideband method.

The estimation is validated with a closure test in $Z \to ee$ simulated events, by comparing the charge flip rates from the Monte-Carlo truth with the ones obtained by the likelihood method. The maximum observed relative deviation is found to be 15%. An additional closure test is performed in data by comparing the number of measured same-sign charge events to the number of estimated same-sign events from reweighting opposite-sign events. The results are found to agree within uncertainties. Additional uncertainties include the statistical uncertainty from the data and the variation in the rates when the Z peak range definition is changed. The total uncertainty on the charge flip rate is at most 20% with the dominant contribution at low p_T from the closure tests and at high p_T from the statistical uncertainty. The charge flip rate as measured in data is shown in Figure 53 as a function of η for the four separate p_T bins for electrons passing the tight selection, and for electrons failing it. The error bars include both the statistical and systematic uncertainties.



(a) Electrons passing tight selection **(b)** Electrons failing tight selection

Figure 53: The probability for an electron to be reconstructed with the wrong charge as a function of η for four separate p_T bins, for 53a electrons passing tight selection and 53b electrons failing it, respectively.

4.8.2 The matrix method for non-prompt leptons and conversion fakes estimation

The estimation of the non-prompt fake leptons background and of the fake background from photon conversions is performed by means of the *matrix method* (MM). At hadron collider experiments, this is the most general data-driven technique for estimating the contamination of fake physics objects, given some level of selection. In ATLAS, it has been employed previously in several top physics analyses [150, 151]; for the $t\bar{t}H$ search, it represents a completely novel approach.

The underlying concept of the matrix method can be effectively outlined in a simplified scenario where only one lepton is taken into account. Firstly, consider a set of lepton candidates satisfying some baseline *loose* selection (L); a subset of this consists of leptons which also pass a *tight* selection (T) corresponding to the strict requirements for the signal region definition as in Table 11. All the loose leptons which fail the tight selection are referred to as *anti-tight* ($\bar{T}$); the "loose" set then corresponds to the union of "anti-tight" and "tight" subsets. The number of events with a tight lepton, and the one with an anti-tight lepton can be expressed in terms of efficiencies and inefficiencies for loose *real* - or *prompt* - (r) and *fake* (f) leptons to pass the tight selection via a system of two linear equations expressed in matrix form:

$$\begin{pmatrix} N^T \\ N^{\bar{T}} \end{pmatrix} = \begin{pmatrix} \varepsilon_r & \varepsilon_f \\ \bar{\varepsilon}_r & \bar{\varepsilon}_f \end{pmatrix} \begin{pmatrix} N^r \\ N^f \end{pmatrix}, \tag{26}$$

where ε_r (ε_f) represents the efficiency for a real (fake) lepton to pass tight selection, and $\not{\varepsilon}_r \equiv (1-\varepsilon_r)$ ($\not{\varepsilon}_f \equiv (1-\varepsilon_f)$) represents the probability for a real (fake) lepton to fail the tight but still pass the baseline selection.

By inverting this matrix equation, one can relate the unknown number of real and fake leptons in a region of interest to a set of observable quantities, i.e., the number of tight and anti-tight leptons and the efficiencies to pass the tight selection. These can be measured directly in dedicated control regions using data.

A simplified version of the method - known as the *fake factor* method - was used in the previous $t\bar{t}H$ multi-leptonic analyses. This method relies on the hypothesis that prompt leptons are correctly modelled in simulation, and leads to additional systematic uncertainties to cover up for this assumption. Since in the matrix method both real and fake leptons are directly modelled from data, a sizeable reduction of the systematic uncertainties is thus foreseen.

In our analysis case, there are two lepton candidates which could be fakes. Therefore, a two-dimensional matrix method formalism is employed. Unless differently stated, in the following leptons are ranked in descending order of p_T by convention. Depending on whether or not each lepton passes the tight selection, each i-th event can be categorised as:

- TT_i : event with both leptons passing tight selection (Tot. events: N^{TT}).
- $T\not{T}_i$: event with leading lepton passing tight selection and subleading lepton failing tight selection (Tot. events: $N^{T\not{T}}$).
- $\not{T}T_i$: event with leading lepton failing tight selection and subleading lepton passing tight selection (Tot. events: $N^{\not{T}T}$).
- $\not{T}\not{T}_i$: event with both leptons failing tight selection (Tot. events: $N^{\not{T}\not{T}}$).

The first set corresponds to the signal region; the remaining three orthogonal regions are referred to as *sidebands*. A 4×4 efficiency matrix can be defined to map the total number of such events into the total number of events in four dileptonic regions characterised by different real and fake lepton composition, namely:

- rr_i : event with both leptons being real (Tot. events: N^{rr}).
- rf_i : event with leading lepton being real and subleading lepton being fake (Tot. events: N^{rf}).
- fr_i : event with leading lepton being fake and subleading lepton being real (Tot. events: N^{fr}).

- ff_i : event with both leptons being fake (Tot. events: N^{ff}).

The 4×4 matrix equation can be written as:

$$\begin{pmatrix} N^{TT} \\ N^{T\not{T}} \\ N^{\not{T}T} \\ N^{\not{T}\not{T}} \end{pmatrix} = \begin{pmatrix} \varepsilon_{r,1}\varepsilon_{r,2} & \varepsilon_{r,1}\varepsilon_{f,2} & \varepsilon_{f,1}\varepsilon_{r,2} & \varepsilon_{f,1}\varepsilon_{f,2} \\ \varepsilon_{r,1}\not{\varepsilon}_{r,2} & \varepsilon_{r,1}\not{\varepsilon}_{f,2} & \varepsilon_{f,1}\not{\varepsilon}_{r,2} & \varepsilon_{f,1}\not{\varepsilon}_{f,2} \\ \not{\varepsilon}_{r,1}\varepsilon_{r,2} & \not{\varepsilon}_{r,1}\varepsilon_{f,2} & \not{\varepsilon}_{f,1}\varepsilon_{r,2} & \not{\varepsilon}_{f,1}\varepsilon_{f,2} \\ \not{\varepsilon}_{r,1}\not{\varepsilon}_{r,2} & \not{\varepsilon}_{r,1}\not{\varepsilon}_{f,2} & \not{\varepsilon}_{f,1}\not{\varepsilon}_{r,2} & \not{\varepsilon}_{f,1}\not{\varepsilon}_{f,2} \end{pmatrix} \begin{pmatrix} N^{rr} \\ N^{rf} \\ N^{fr} \\ N^{ff} \end{pmatrix}, \tag{27}$$

The indexes for ε_r and ε_f refer to leading (1) and subleading (2) lepton p_T.

To obtain the number of fakes in the signal region, the 4×4 matrix is inverted (provided that $\varepsilon_r \neq \varepsilon_f$):

$$\begin{pmatrix} N^{rr} \\ N^{rf} \\ N^{fr} \\ N^{ff} \end{pmatrix} = \begin{pmatrix} \varepsilon_{r,1}\varepsilon_{r,2} & \varepsilon_{r,1}\varepsilon_{f,2} & \varepsilon_{f,1}\varepsilon_{r,2} & \varepsilon_{f,1}\varepsilon_{f,2} \\ \varepsilon_{r,1}\not{\varepsilon}_{r,2} & \varepsilon_{r,1}\not{\varepsilon}_{f,2} & \varepsilon_{f,1}\not{\varepsilon}_{r,2} & \varepsilon_{f,1}\not{\varepsilon}_{f,2} \\ \not{\varepsilon}_{r,1}\varepsilon_{r,2} & \not{\varepsilon}_{r,1}\varepsilon_{f,2} & \not{\varepsilon}_{f,1}\varepsilon_{r,2} & \not{\varepsilon}_{f,1}\varepsilon_{f,2} \\ \not{\varepsilon}_{r,1}\not{\varepsilon}_{r,2} & \not{\varepsilon}_{r,1}\not{\varepsilon}_{f,2} & \not{\varepsilon}_{f,1}\not{\varepsilon}_{r,2} & \not{\varepsilon}_{f,1}\not{\varepsilon}_{f,2} \end{pmatrix}^{-1} \begin{pmatrix} N^{TT} \\ N^{T\not{T}} \\ N^{\not{T}T} \\ N^{\not{T}\not{T}} \end{pmatrix}. \tag{28}$$

The final number of events with fakes in the signal region N^f_{TT}, i.e. the total number of TT events with at least one fake lepton, can be obtained from the definition:

$$N^f_{TT} = N^{rf}_{TT} + N^{fr}_{TT} + N^{ff}_{TT} = \varepsilon_{r,1}\varepsilon_{f,2}N^{rf} + \varepsilon_{r,2}\varepsilon_{f,1}N^{fr} + \varepsilon_{f,1}\varepsilon_{f,2}N^{ff}. \tag{29}$$

By plugging the matrix equation (28) into this formula, and making the weighted sum over the events explicit, we finally obtain:

$$\begin{aligned} N^f_{TT} = & \\ &= \sum_i \left(\varepsilon_{r,1}\varepsilon_{f,2}\ rf\right)_i + \left(\varepsilon_{r,2}\varepsilon_{f,1}\ fr\right)_i + \left(\varepsilon_{f,1}\varepsilon_{f,2}\ ff\right)_i = \\ &= \sum_i^{\{TT\}} \left(w^{MM}_{TT}\ TT\right)_i + \sum_i^{\{T\not{T}\}} \left(w^{MM}_{T\not{T}}\ T\not{T}\right)_i + \sum_i^{\{\not{T}T\}} \left(w^{MM}_{\not{T}T}\ \not{T}T\right)_i + \sum_i^{\{\not{T}\not{T}\}} \left(w^{MM}_{\not{T}\not{T}}\ \not{T}\not{T}\right)_i \end{aligned} \tag{30}$$

where:

$$
\begin{aligned}
w_{TT}^{MM}{}_i &= \left(1 - \alpha\varepsilon_{r,1}\varepsilon_{r,2}\not{\varepsilon}_{f,1}\not{\varepsilon}_{f,2}\right)_i \\
w_{TT}^{MM}{}_i &= \left(\alpha\varepsilon_{r,1}\varepsilon_{r,2}\varepsilon_{f,2}\not{\varepsilon}_{f,1}\right)_i \\
w_{TT}^{MM}{}_i &= \left(\alpha\varepsilon_{r,1}\varepsilon_{r,2}\varepsilon_{f,1}a\not{\varepsilon}_{f,2}\right)_i \\
w_{TT}^{MM}{}_i &= -\left(\alpha\varepsilon_{r,1}\varepsilon_{r,2}\varepsilon_{f,1}\varepsilon_{f,2}\right)_i \\
\alpha_i &= \frac{1}{\left(\varepsilon_{r,1_i} - \varepsilon_{f,1_i}\right)\left(\varepsilon_{r,2_i} - \varepsilon_{f,2_i}\right)}
\end{aligned}
\tag{31}
$$

Each i-th event will contribute to only one of the four sums on the right side of Eq. (30).

Assumptions of the method

For the matrix method to work as previously illustrated, it is necessary that at least one of the following conditions is fulfilled:

- Fake leptons from different sources must have the same efficiency to pass the tight selection.
- The sources of fake leptons must have the same composition in the control regions where the fake rates are extracted, and in the target regions where the fake background is actually estimated.

As anticipated, several sources of fake leptons do coexist in the $t\bar{t}H$ multileptonic final state, notably non-prompt leptons from heavy hadron decays, electrons from converted photons, and charge flip electrons.

CHARGE FLIP BACKGROUND IN THE MATRIX METHOD Electrons with misidentified charge have a different efficiency of fulfilling the tight selection requirements than electrons with correctly assigned charge. Furthermore, their relative fraction is found to be larger in the low jet multiplicity regions where the fake rates are measured (defined in the next paragraphs) than in the 2ℓ SS $0\tau_{had}$, high jet multiplicity signal region. As a consequence, regions contaminated by such background - namely ee and $e\mu$ - must be treated differently than $\mu\mu$ in the context of the matrix method. As from section subsection 4.8.1, charge flip background events N_{QMisID} are estimated with an independent data-driven technique by reweighting opposite-sign (OS) data events by a factor w^{QMisID} which depends on the probability of single electrons to be reconstructed with the wrong charge, according to

Eq. (25). These events can be subtracted from the region used in the matrix method by taking advantage of the linearity of Eq. (30). This effectively decouples the estimation of this background from the estimation of the other fakes. In fact, the above equation can be rewritten as:

$$\begin{aligned} N^{f}_{TT} &= N^{rf}_{TT} + N^{fr}_{TT} + N^{ff}_{TT} \\ &= \varepsilon_{r,1}\varepsilon_{f,2}N^{rf} + \varepsilon_{r,2}\varepsilon_{f,1}N^{fr} + \varepsilon_{f,1}\varepsilon_{f,2}N^{ff} = \\ &= w^{MM}_{TT}\, N^{TT} + w^{MM}_{T\bar{T}}\, N^{T\bar{T}} + w^{MM}_{\bar{T}T}\, N^{\bar{T}T} + w^{MM}_{\bar{T}\bar{T}}\, N^{\bar{T}\bar{T}} \rightarrow \\ &\rightarrow w^{MM}_{TT}\, \left(N^{TT} - N^{TT}_{QMisID}\right) + w^{MM}_{T\bar{T}}\, \left(N^{T\bar{T}} - N^{T\bar{T}}_{QMisID}\right) + \\ &+ w^{MM}_{\bar{T}T}\, \left(N^{\bar{T}T} - N^{\bar{T}T}_{QMisID}\right) + w^{MM}_{\bar{T}\bar{T}}\, \left(N^{\bar{T}\bar{T}} - N^{\bar{T}\bar{T}}_{QMisID}\right) \end{aligned} \tag{32}$$

Once making the weighted sums explicit, it eventually reads:

$$\begin{aligned} N^{f}_{TT} = & \\ &= \sum_{i}^{\{TT,SS\}} \left(w^{MM}_{TT}\, TT\right)_i - \sum_{i}^{\{TT,OS\}} \left(w^{MM}_{TT}\, w^{QMisID}_{TT}\, TT\right)_i + \\ &+ \sum_{i}^{\{T\bar{T},SS\}} \left(w^{MM}_{T\bar{T}}\, T\bar{T}\right)_i - \sum_{i}^{\{T\bar{T},OS\}} \left(w^{MM}_{T\bar{T}}\, w^{QMisID}_{T\bar{T}}\, T\bar{T}\right)_i + \\ &+ \sum_{i}^{\{\bar{T}T,SS\}} \left(w^{MM}_{\bar{T}T}\, \bar{T}T\right)_i - \sum_{i}^{\{\bar{T}T,OS\}} \left(w^{MM}_{\bar{T}T}\, w^{QMisID}_{\bar{T}T}\, \bar{T}T\right)_i + \\ &+ \sum_{i}^{\{\bar{T}\bar{T},SS\}} \left(w^{MM}_{\bar{T}\bar{T}}\, \bar{T}\bar{T}\right)_i - \sum_{i}^{\{\bar{T}\bar{T},OS\}} \left(w^{MM}_{\bar{T}\bar{T}}\, w^{QMisID}_{\bar{T}\bar{T}}\, \bar{T}\bar{T}\right)_i \end{aligned} \tag{33}$$

From simulation, the fake electron origin is known to be in more than 90% of the cases the same in signal region dielectron events regardless of the relative lepton charge. This allows exploiting a unique measurement of ε_f - as will be described in the following - for reweighting both SS and OS data events.

NON-PROMPT LEPTON AND PHOTON CONVERSION BACKGROUNDS IN THE MATRIX METHOD Ignoring charge flips, Figure 54 shows the fractions of fake leptons from different origins in $t\bar{t}$ simulated events - defined as the ratio of fake leptons coming from a specific source over the total number of fakes - in different 2ℓ SS $0\tau_{had}$ regions split by lepton flavour composition and jet multiplicity. All leptons are required to pass the tight selection. In the plot, "CR" indicates dileptonic regions with low jet multiplicity ($2 \leq N_{jets} \leq 3$), whereas "SR" represent the signal regions, characterised by

a higher number of jets ($N_{jets} \geq 4$). The control regions with low jet multiplicity are used to measure the real and fake efficiencies, as detailed in the next paragraph.

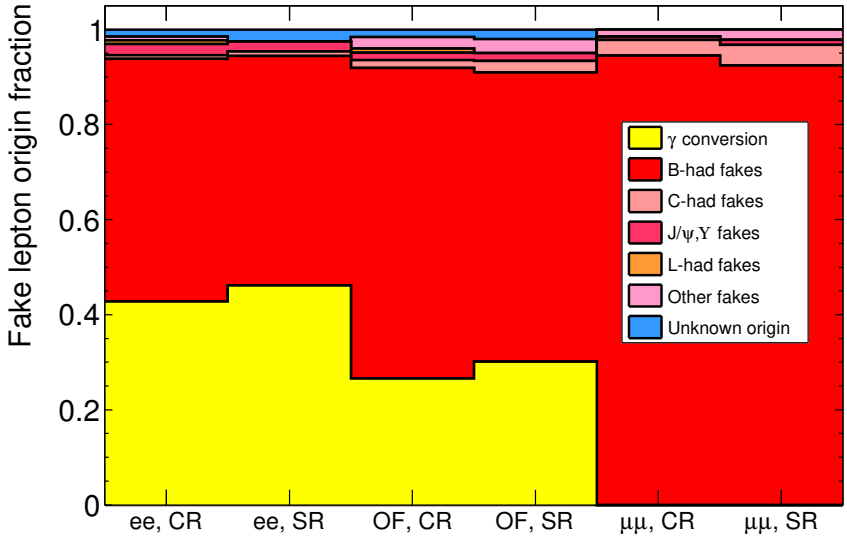


Figure 54: The composition of fake lepton sources in events with tight leptons in the 2ℓ SS $0\tau_{had}$ signal and control regions as estimated in $t\bar{t}$ simulation.

For muons, the majority of fakes consists of *non-prompt* leptons from b-flavoured hadron decays. Other types of fakes include muons from J/ψ resonance decays, charm and light flavoured mesons ($\pi, K \ldots$) decays, non-prompt τ lepton decays and mis-identified jets. All these sources have fractions that are either stable across the regions or negligible.

In the electron case, the picture is more complex. In general, the fraction of non-prompt electron fakes dominates, but a large fraction of electrons from photon conversions is also present. This is a consequence of the usage of the `PromptLeptonIso` algorithm, which is specifically targeted at suppressing the lepton fakes contribution from heavy flavour decays. The fractions of fakes which are neither non-prompt, nor conversions are small ($\lesssim 5\%$) and fairly similar in all regions, thus they are neglected.

On the contrary, the relative electron non-prompt and conversion fakes fractions are not stable. Also, the efficiency for non-prompt electrons is much smaller than the one for conversions by approximately a factor 6 in simulation, meaning that those two processes cannot be estimated properly by the matrix method described in Eq. (33). This would call for an independent fake estimation strategy for non-prompt electrons and photon conversions,

respectively, using separated control regions enriched in either type of fakes to measure their efficiency.

Since we could not define such independent regions with sufficiently high purity given the available data statistics, an alternative strategy has been devised. The approach is to account for the two contributions simultaneously by measuring an *effective* fake efficiency in a control region with both types of fakes, and correct this efficiency for the different proportion of conversions in the signal region using rescaling factors extracted from simulation, where the efficiencies for each fake type are known. Starting from the effective efficiency for non-prompt and conversions $\langle \varepsilon_{CR} \rangle$ as measured in the electron fake-enriched control region, the efficiency used to estimate fakes in a region of interest i is obtained by multiplying $\langle \varepsilon_{CR} \rangle$ by a rescaling factor α_i to account for the relative proportions of conversion and non-prompt fakes between the control region and the region itself [6]:

$$\begin{aligned} \langle \varepsilon_i \rangle &= (1 + \alpha_i) \cdot \langle \varepsilon_{CR} \rangle \\ \alpha_i &= \left[\frac{\left(1 - f_\gamma^{X_i}\right) \cdot \varepsilon_{HF} + f_\gamma^{X_i} \cdot \varepsilon_\gamma}{\left(1 - f_\gamma^{\mathrm{CR}}\right) \cdot \varepsilon_{HF} + f_\gamma^{\mathrm{CR}} \cdot \varepsilon_\gamma} \right] - 1, \end{aligned} \tag{34}$$

where the $f_\gamma^{X_i}$ indicate the conversion fractions, and $\varepsilon_{HF}(\varepsilon_\gamma)$ are the non-prompt (conversion) fake rates as predicted from simulation.

Real and fake lepton efficiency measurement

A set of two control regions - referred to as *real and fake control regions* (CRs) in the following - is defined to measure the efficiency for real and fake leptons to pass the tight selection requirements. In the following, these are sometimes referred to as *real efficiency* and *fake efficiency* for simplicity.

Such control regions are designed to be representative of the 2ℓ SS $0\tau_{had}$ signal region in term of kinematics and background composition, although retaining sufficiently large statistics. Orthogonality is enforced by means of a requirement of 2 or 3 jets only in the event. The requirement for at least one b-tagged jet is kept to avoid potentially large changes in the fakes composition.

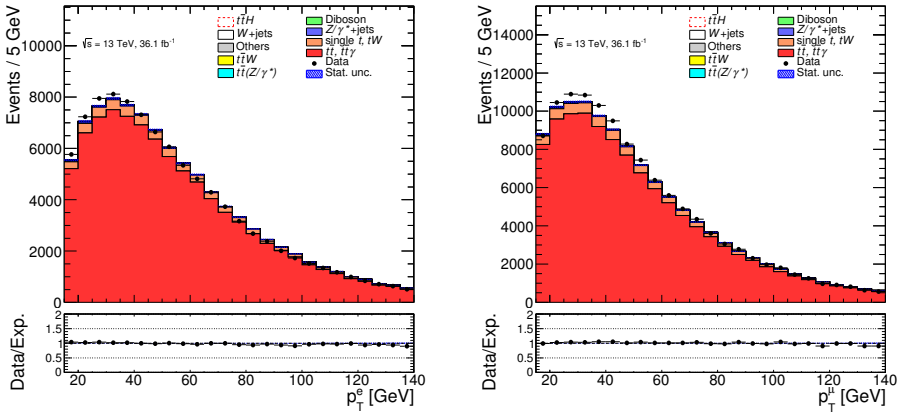
REAL CONTROL REGION DEFINITION AND ε_r MEASUREMENT The real control region is designed to be enriched in prompt leptons from dileptonic $t\bar{t}$ decays by requiring the presence of two leptons with opposite-sign charge and

6 Here we assume electron fakes are either non-prompt or photon conversions.

	Real lepton-enriched CR
N_{jets}	2,3
$N_{\tau_{had}}$	0
N_{b-tags}	≥ 1 (`MV2c10`, 70% eff.)
N_{ℓ}	2
Lepton charge	OS
Lepton flavour	$e\mu$,μe
Lepton selection	L (11)
p_T^{ℓ}	≥ 20 GeV
Trigger	Logical `OR` of SLTs (e, μ) as from Tab. 12
Nr. trigger-matched ℓ	≥ 1

Table 17: Definition of the control region used for measuring the real lepton efficiency. The same region is used to measure both $\varepsilon_r(e)$ and $\varepsilon_r(\mu)$.

opposite flavour. The event topology is very close to the signal region, where we expect prompt leptons to be often found close to b-jets. The prompt lepton purity achieved is very high, as can be observed in Figure 55. The full definition of the control region is summarised in Table 17.



(a) *Tight e* **(b)** *Tight μ*

Figure 55: Observed and expected events in the real control region with *tight* electrons and muons.

A *tag and probe* method is used to measure the real efficiency for both electrons and muons. Each event in the control region is tagged by requiring at least one of the leptons to pass the tight offline selection and be matched to a lepton reconstructed at trigger level. Then the other lepton - which is unbiased by both the offline and trigger selection - is picked as the probe candidate for the efficiency measurement. In case two tag candidates are

found, both leptons are considered as valid probes and used for the measurement. This procedure follows the one used to measure the muon and electron efficiencies for calibrating the tight selection working point.

The efficiency for prompt leptons is defined on a per-event basis as the ratio of "numerator" events where the probe passes the tight selection, and "denominator" events where the probe passes no additional selection other than the baseline loose one:

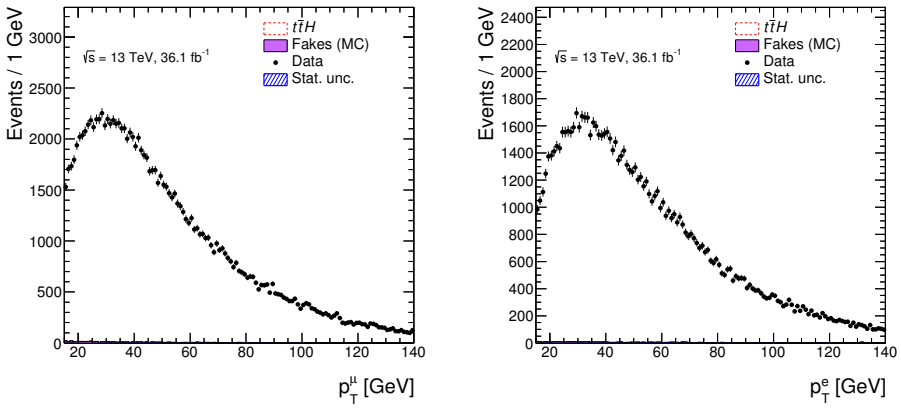
$$\varepsilon_r(e,\mu)_i = \frac{N_i^{T(e,\mu)} - N_{bkg}^{T(e,\mu)}{}_i}{N_i^{L(e,\mu)} - N_{bkg}^{L(e,\mu)}{}_i}, \tag{35}$$

where the background from events with fake leptons is subtracted. Given the small amount of such fake events, these are estimated directly from $t\bar{t}$ and single top simulation. The index i in the definition accounts for the p_T binning chosen for parametrising the efficiency.

Since the tag lepton must always be tight, the numerator events are associated to tight-tight (TT) dilepton pairs, whereas denominator events are associated to tight-loose ($TL = TT + T\not{T}$), loose-tight ($LT = TT + \not{T}T$) dilepton pairs. Figure 56 shows the probe lepton p_T distributions in data and background, for numerator and denominator events. After subtracting the background from the data, their ratio eventually gives the real efficiencies displayed in Figure 57. The error bars include systematic uncertainties, which will be discussed later in subsection 4.9.1.

FAKE CONTROL REGION DEFINITION AND ε_f MEASUREMENT Unlike the real control region, which represents a pure sample of prompt leptons once the small fake background contamination has been subtracted, it is not possible to select a purely fake-enriched region. In fact, only a semileptonic $t\bar{t}$ control region with one real lepton and one fake can be defined. Because of this ambiguity, a slightly different approach than the standard tag-and-probe method is followed to measure the fake efficiency for electrons and muons.

Measurement of $\varepsilon_f(e)$ When considering opposite-flavour, same-sign dilepton events where a tight muon firing the single muon trigger is found, the muon itself has a high probability of being the prompt lepton of the pair, leaving the electron as a suitable unbiased probe candidate to measure the fake efficiency. An additional advantage of using only opposite-flavour events to measure the electron fake efficiency is the small amount of charge flip events to be subtracted compared to same-sign ee events. Table 18 sum-



(a) *Tight* probe μ (N)

(b) *Tight* probe e (N)

(c) *Loose* probe μ (D)

(d) *Loose* probe e (D)

Figure 56: Distributions of p_T for probe electrons and muons in the real control region (numerator and denominator events) for data events, and for background events to be subtracted to data.

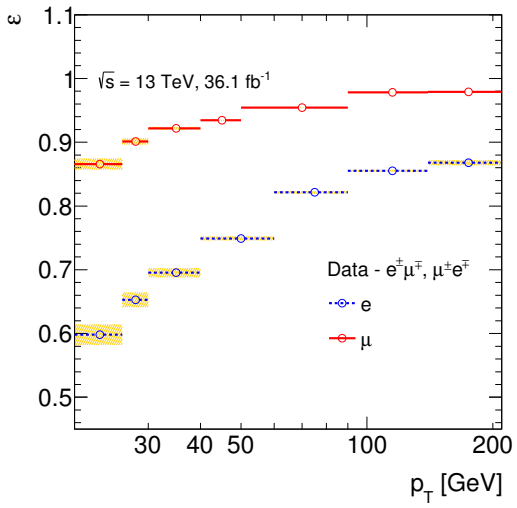


Figure 57: Electron and muon real efficiencies as measured in data. The bands represent the total systematic uncertainty (see subsection 4.9.1). The last bin includes overflows.

marises the definition of the control region used to measure the electron fake efficiency.

A large variation of the fake electron origin fractions depending on the number of b-tagged jets is observed in simulation: events with $N_{b-tags} = 2$ have a consistently higher fraction of fakes from photon conversions (approximately 60% of the total) compared to $N_{b-tags} = 1$ events ($\sim$ 20%). This is an effect of the electron-jet overlap removal procedure by which jets found within $\Delta R = 0.3$ of an electron are removed. This enhances the population of fakes not necessarily related to jet proximity - such as converted photons - in all the events where both the b-jets are kept. Therefore, a N_{b-tags}-dependent parametrisation of the electron fake rate has been adopted. This choice is particularly beneficial, as the $N_{b-tags} = 2$ corner of the phase space is the one where we expect to be most sensitive to the $t\bar{t}H$ signal. Alternative parametrisations such as (p_T, η), $(p_T, \Delta R(e, jet))$ were tested, but they did not improve the modelling of electron fakes. A two-dimensional (N_{b-tags}, p_T) parametrisation for the electron fake efficiency has also been tried. However, the p_T slices of the 2D efficiency map have been observed to be approximately flat within the statistical uncertainties across the full transverse momentum spectrum, hence a one dimensional parametrisation on N_{b-tags} only was considered appropriate. The choice made represents the best compromise between providing a good modelling of fakes and minimising the statistical uncertainty.

	Fake electron-enriched CR
N_{jets}	2,3
$N_{\tau_{had}}$	0
N_{b-tags}	≥ 1 (MV2c10, 70% eff.)
N_{ℓ}	2
Lepton charge	SS
Lepton flavour	$e\mu$,μe
Lepton offline sel.	L (11)
p_T^{ℓ}	≥ 20 GeV
Tag lepton sel.	T (11) μ and matched to any SLT (μ) (12)

Table 18: Definition of the control region used for measuring the electron fake efficiency. Note that the muon is required to tag the event, with the remaining electron being used as selection-unbiased probe for the efficiency measurement.

The efficiency for fake electrons, similarly to the real electron case, is defined on a per-event basis as:

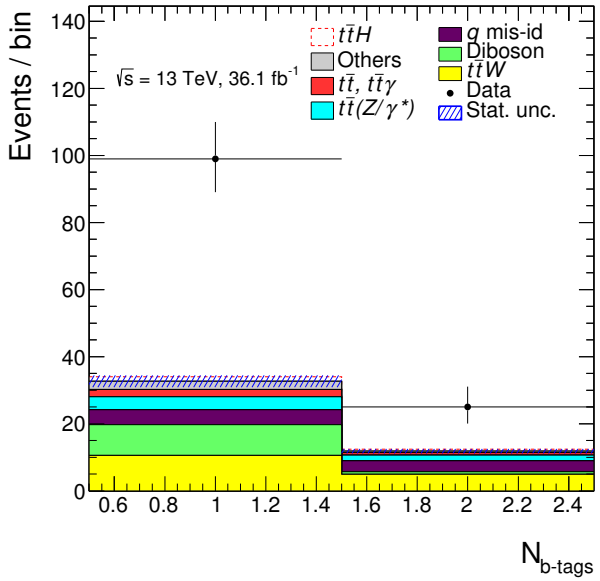
$$\varepsilon_f(e)_i = \frac{N_i^{T(e)} - N_{bkg\ i}^{T(e)}}{N_i^{L(e)} - N_{bkg\ i}^{L(e)}}. \tag{36}$$

The index i refers to the N_{b-tags} bin in question. The "background" to be subtracted accounts for events with genuine prompt same-sign lepton pairs such as $t\bar{t}W$, $t\bar{t}Z$ and diboson, charge flip events, and semileptonic $t\bar{t}$ events with mis-assigned probe electron. Except for charge flips, these backgrounds are all estimated from simulation.

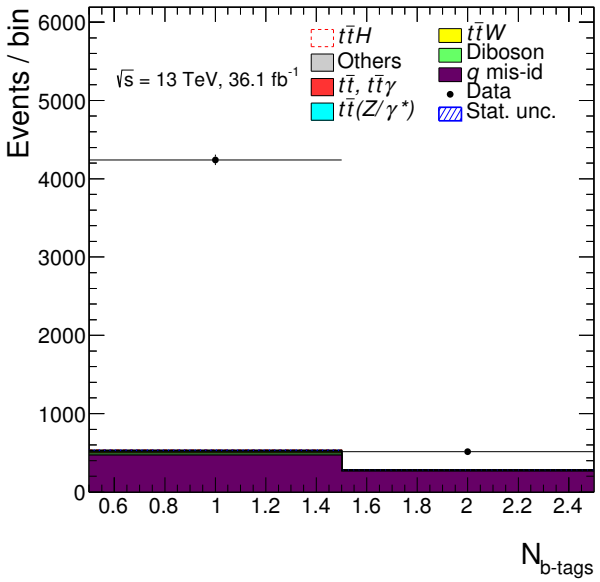
The overall fake lepton purity achieved is manageable, as shown in Table 19. The background to be subtracted is approximately 17% of the observed data yield for denominator events, and approximately 35% for numerator events. The amount of $t\bar{t}$ events where the probe electron happens to be prompt is small, accounting for only $\approx$ 6% of the total background events to be subtracted at the numerator.

Figure 58 shows the N_{b-tags} distribution for probe electrons in data for numerator and denominator events, whose ratio gives the efficiency displayed in Figure 59 after subtracting the backgrounds.

Measurement of $\varepsilon_f(\mu)$ In the muon case, using a similar approach as for electrons by selecting opposite-flavour SS events, tagging on a tightly selected, trigger-matched electron and using the muon as probe, is not opti-



(a) *Tight* probe e (N), N_{b-tags}



(b) *Loose* probe e (D), N_{b-tags}

Figure 58: Distributions of N_{b-tags} for probe electrons in the same-sign, opposite-flavour fake e control region (numerator and denominator events) for data events, and for background processes with prompt probe electrons that have to be subtracted to data.

	Loose probe e (D)		*Tight* probe e (N)	
	Yield	Fraction of Tot. Sub. [%]	Yield	Fraction of Tot. Sub. [%]
Observed	4755		124	
Tot. Sub.	814.9 ± 9.0		44.5 ± 1.9	
$t\bar{t}Z$	4.5 ± 0.2	0.5	3.0 ± 0.1	6.6
Diboson	24.7 ± 6.1	3.0	10.0 ± 1.6	22.4
Others	4.2 ± 0.2	0.5	2.9 ± 0.2	6.4
Charge flip	738 ± 6	90.5	7.7 ± 0.1	17.4
$t\bar{t}W$	22.7 ± 0.4	2.8	15.5 ± 0.4	34.8
$t\bar{t}$	11.1 ± 2.2	1.4	2.9 ± 0.9	6.5
$t\bar{t}(\gamma^*\ell\ell)$	10.1 ± 0.3	1.2	2.6 ± 0.2	5.9

Table 19: Observed and expected event yield for *numerator* (N) events (i.e., where the probe passes the *tight* selection) and *denominator* (D) events of the electron fake CR. "Tot. Sub." represents the total yield from the listed background events with a prompt probe electron to be subtracted. Uncertainties are only statistical.

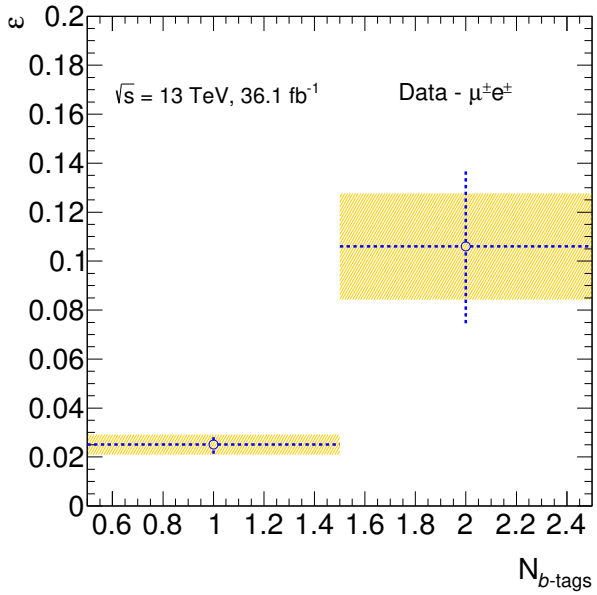


Figure 59: Electron fake efficiency as measured in data, as a function of N_{b-tags}. The bands represent systematic uncertainties, as discussed in subsection 4.9.1.

	Fake muon-enriched CR
N_{jets}	2,3
$N_{\tau_{had}}$	0
N_{b-tags}	≥ 1 (MV2c10, 70% eff.)
N_{ℓ}	2
Lepton charge	SS
Lepton flavour	$\mu\mu$
Lepton offline sel.	L (11)
p_T^{ℓ}	≥ 20 GeV
Tag lepton sel.	T (11) μ and matched to any SLT (μ) (12) OR (if both T-trigger-matched) leading-p_T μ

Table 20: Definition of the control region used for measuring the muon fake lepton efficiency. Note that when both muons are tight and trigger-matched, the subleading p_T muon is chosen as probe.

mal. In fact, for numerator events the probability of selecting an actual fake muon as the probe becomes very low when both the leptons are required to pass the tight selection. To overcome this, a tag and probe method in same flavour $\mu\mu$ events is used, assuming that in case both muons are tight and fired the trigger, the subleading one is more likely to be the fake lepton in the pair, hence it is chosen as the probe. Table 20 summarises the definition of the control region used to measure the muon fake efficiency.

A poor modelling of the low $\Delta R(\mu, jet)$ [7] corner of the signal region phase space is observed when using a one-dimensional p_T parametrisation of the muon fake rate. To account for the impact from nearby jets, a two-dimensional $(\Delta R(\mu, jet), p_T)$ parametrisation of the rate was found to be the best handle to improve the modelling. The choice of two bins for $\Delta R(\mu, jet)$ ($[0.0, 1.0, 5.0]$) and two bins for p_T ($[20, 50, \geq 210]$ GeV) represents the best compromise between the ability to model muon fakes properly and to cope with limited statistics in the control region.

The efficiency for fake muons is defined as:

$$\varepsilon_f(\mu)_i = \frac{N_{i,j}^{T(\mu)} - N_{bkg\ i,j}^{T(\mu)}}{N_{i,j}^{L(\mu)} - N_{bkg\ i,j}^{L(\mu)}}. \tag{37}$$

7 Henceforth $\Delta R(\ell, jet)$ indicates the distance in (η, ϕ) of a lepton to the *closest* jet.

	Loose probe μ (D)		*Tight* probe μ (N)	
	Yield	Fraction of Tot. Sub. [%]	Yield	Fraction of Tot. Sub. [%]
Observed	1680		159	
Tot. Sub.	80.7 ± 3.4		61.5 ± 2.8	
$t\bar{t}Z$	5.2 ± 0.2	6.4	4.5 ± 0.2	7.3
Diboson	21.6 ± 2.5	26.7	15.8 ± 2.2	25.7
Others	5.3 ± 0.2	6.6	4.7 ± 0.2	7.6
$t\bar{t}W$	27.4 ± 0.5	34.0	24.6 ± 0.5	40.0
$t\bar{t}$	11.6 ± 2.2	14.4	8.6 ± 1.6	14.0
$t\bar{t}(\gamma^* \ell\ell)$	9.6 ± 0.3	11.9	3.3 ± 0.2	5.3

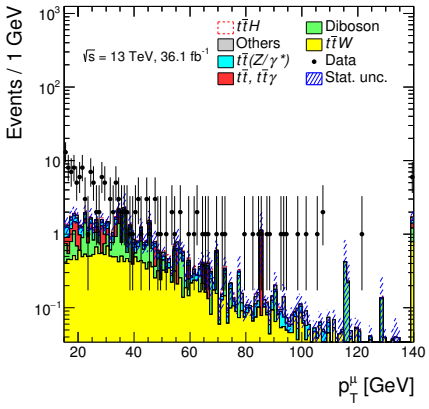
Table 21: Observed and expected event yield for *numerator* (N) events (i.e., where the probe passes the *tight* selection) and *denominator* (D) events of the muon fake CR. "Tot. Sub." represents the total yield from the listed background events with a prompt probe muon to be subtracted. Uncertainties are only statistical.

The double index (i, j) refers to the $(\Delta R(\mu, jet), p_T)$ bin in question. In the muon case, the "background" to be subtracted does not account for charge flip contamination. The assumption of using the p_T ranking as criterion to solve the ambiguous cases of both muons being tight and matched to the trigger - despite introducing a selection bias - does not have a significant negative impact on the muon fake estimate. This can be inferred from Table 21: the contamination of events where the probe was mis-assigned reaches at maximum 14% of the total yield to be subtracted. Decent fake lepton purity is achieved: the prompt event yield to be subtracted is approximately 5% of the observed data yield for denominator events, growing to approximately 38% for numerator events.

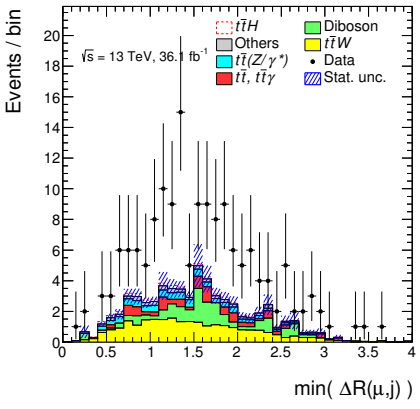
Probe muon p_T and $\Delta R(\mu, jet)$ projections of the $(\Delta R(\mu, jet), p_T^{\ell})$ distributions in data for numerator and denominator events are represented in Figure 60. Figure 61 displays the 2D $(\Delta R(\mu, jet), p_T)$ efficiency map projected along p_T for different $\Delta R(\mu, jet)$ regions.

Closure test in simulation

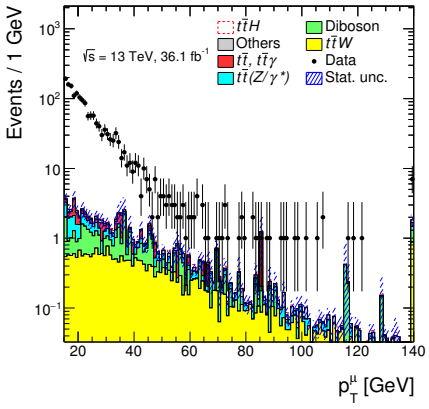
A *closure test* of the method has been performed in order to validate the assumptions of the method, as well as to ensure no additional method bias is present and the matrix method algebra itself is correct. The test consists of comparing the prediction from "out-of-the-box" simulation to the fakes estimation obtained from measurement of real and fake efficiencies as ex-



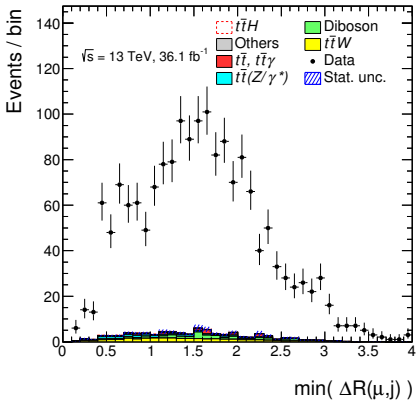
(a) *Tight* probe μ (N), p_T



(b) *Tight* probe μ (N), $\Delta R(\mu, jet)$



(c) *Loose* probe μ (D), p_T



(d) *Loose* probe μ (D), $\Delta R(\mu, jet)$

Figure 60: Distribution of p_T and $\Delta R(\mu, jet)$ for probe muons in the same-sign, same-flavour fake μ control region (numerator and denominator events) for data events, and for background processes with prompt probe muons that have to be subtracted to data.

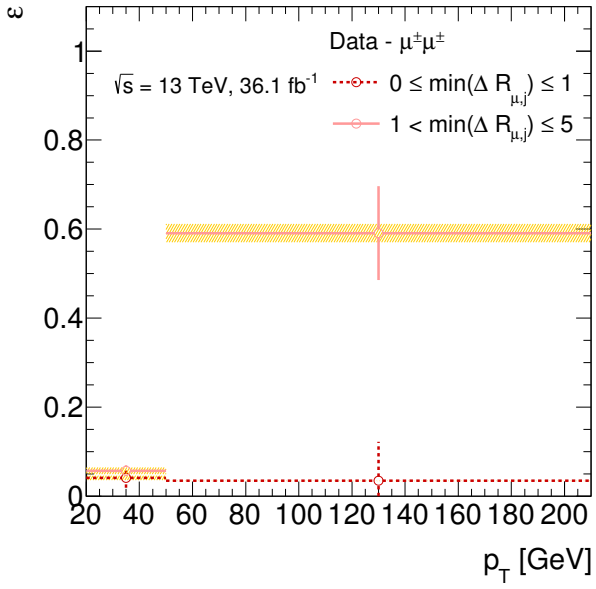


Figure 61: Muon fake efficiency as measured in data. The plot represents the fake rate projection over p_T for each $\Delta R(\mu, jet)$ slice of a 2-dimensional efficiency map. The bands represent systematic uncertainties, as discussed in subsection 4.9.1.

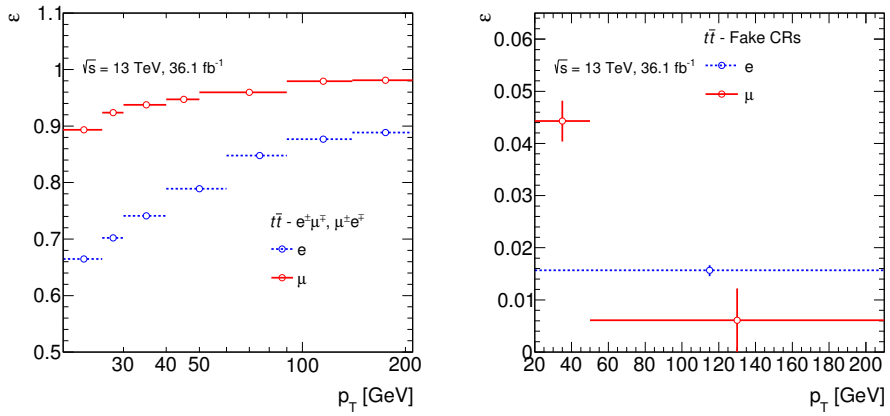
plained previously, but using only simulated events from the same process. The $t\bar{t}$ sample is employed for this test.

One obvious potential source of non-closure is the N_{jets}-dependent extrapolation between the low jet multiplicity fake control regions and the high jet multiplicity signal regions. To assess the robustness of the method and check whether any additional extrapolation effect other than the one just cited, the closure test is firstly performed in three low jet multiplicity regions split by lepton flavour composition - ee, $\mu\mu$ and OF. The different flavour combinations are tested separately to check any effect of flavour-dependent fake lepton origin fraction instability. This is particularly crucial to ensure the α rescaling procedure applied works as expected. Any non closure observed there will also account for missing non-trivial efficiency parametrisations and binning effects. Finally, the closure test is performed in the signal regions for the final state categories split by flavour: ee, $\mu\mu$ and OF.

The event selection is the same used for the main analysis signal region definition, except for the jet multiplicity cuts which can vary between $2 \leq N_{jets} \leq 3$ and $N_{jets} \geq 4$. The expected prediction is obtained from $t\bar{t}$ events, selecting a sample with two tight leptons, and requiring the presence of at least one truth-matched fake lepton - either coming from a photon conversion or from heavy flavour decay. On top of this selection, events with

truth charge flip leptons are vetoed as well. The same truth selection is also applied in all the four loose sidebands contributing to the matrix method estimate. Events are normalised to the $t\bar{t}$ cross section, and to an integrated luminosity of 36.1 fb^{-1}.

Figure 62 displays the real and fake efficiencies as measured in $t\bar{t}$ events according to the previously described methods, where a truth requirement of exactly two prompt leptons is applied in the real control region. In the fake case, the probe lepton is instead required not to be prompt, and truth charge flip events are vetoed. Such efficiencies are used to compute the matrix method weights for the same $t\bar{t}$ sample used to derive them.



(a) ε_r **(b)** ε_f

Figure 62: Electron, muon real (left) and fake (right) efficiencies as measured in $t\bar{t}$ simulation. In the muon fake case, the plots is an inclusive projection over p_T of the $(\Delta R(\mu, jet), p_T)$ efficiency map previously described.

The *non-closure* (ζ) can be directly obtained from the ratio of the pure simulation and matrix method predictions - which we refer to as *scale factor (SF)* in the following - and is defined as the relative bias between the two estimates:

$$\begin{aligned} SF &= \frac{N_{t\bar{t}}}{N_{MM}}, \\ \zeta &= 1 - SF = \frac{N_{MM} - N_{t\bar{t}}}{N_{MM}}, \\ \sigma_\zeta &= \sigma_{SF} = \sqrt{\frac{1}{N_{MM}^2} \cdot \sigma_{N_{t\bar{t}}}^2 + \frac{N_{t\bar{t}}^2}{N_{MM}^4} \cdot \sigma_{N_{MM}}^2}. \end{aligned} \tag{38}$$

The error on the matrix method estimate σ_{NMM} is a combination of the statistical uncertainty driven by the statistical size of the signal region and

	ee	$\mu\mu$	OF
$t\bar{t}$ MC	73.1 ± 4.5	84.4 ± 4.9	141.8 ± 6.3
MM fakes ($t\bar{t}$)	67.2 ± 4.5	81.0 ± 5.7	134.1 ± 7.1
Non-closure (ζ)	-8.9 ± 10.0 [%]	-4.2 ± 9.5 [%]	-5.8 ± 7.3 [%]

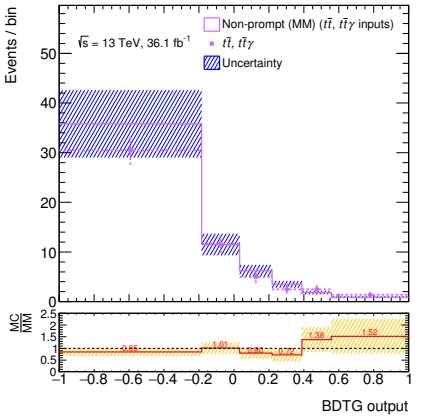
Table 22: Total yield of $t\bar{t}$ events with fakes based on simulation, and the yield of fake events estimated with the matrix method applied on the same simulated sample in the $2 \leq N_{jets} \leq 3$, $N_{b-tags} \geq 1$, and the non-closure ζ. Errors for the matrix method fakes arise from the SR and sidebands statistics, as well as from the statistics of the control regions.

of the MM sidebands, and the uncertainty on the measured efficiencies. In this test, the latter depends exclusively on the statistical size of the control regions. The above formula for the non-closure uncertainty holds in case $\sigma_{N_{MM}}$ and $\sigma_{N_{t\bar{t}}}$ are not correlated. This is true in first approximation, since the contribution of the TT region - which is the only one not independent between the pure simulated event set and the matrix method set - has a small contribution to the total fakes yield compared to the sidebands with anti-tight leptons.

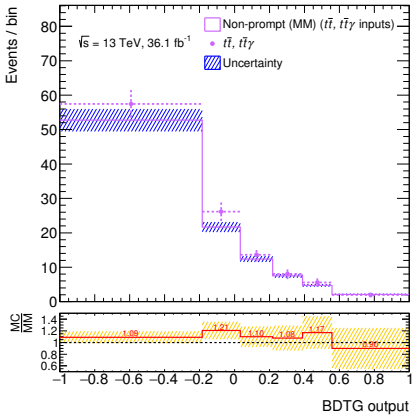
In the following, we consider the method as fully validated if the non-closure ζ is consistent with zero within $1\sigma_\zeta$. Still, if a small non-closure ($\lesssim 1.5\sigma_\zeta$) is found, we choose not to correct for the bias. However, the maximum between ζ and its statistical uncertainty is conservatively accounted for as the non-closure systematic uncertainty on the matrix method. The total event yield for matrix method fakes and pure simulation prediction in the low jet multiplicity region is summarised in Table 22. The non-closure in all channels is found to be consistent with zero within its uncertainty.

Figure 63 shows the shape of the final BDTG discriminator [8] in the signal region split by lepton flavour - for matrix method fakes and pure simulation prediction in the signal region. No significant trend within the statistical uncertainty for the non-closure is seen across the BDTG shape. The total event yield and the non-closure are summarised in Table 23. No significant non-closure is seen in the $\mu\mu$ channel, and in the ee, OF channel is of order $\sim 11 \pm 9\%$. Based upon our previous considerations, we therefore neglect such small bias. This implies the N_{jet} extrapolation uncertainty is small, and that the α rescaling procedure is effective in accounting for the uneven electron fake composition between the electron fake control region and the signal regions.

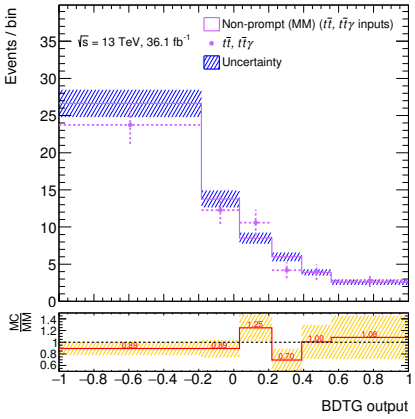
8 This is obtained by combining the output of the $BDTG_{t\bar{t}V}$, $BDTG_{t\bar{t}}$ classifiers, as described in the next chapter.



(a) *BDTG* output, *ee*



(b) *BDTG* output, *OF*



(c) *BDTG* output, $\mu\mu$

Figure 63: Distributions of the final BDTG discriminator output in the 2ℓ SS $0\tau_{had}$ signal region split by lepton flavour, for pure $t\dot{t}$ events (dots) and matrix method fakes with efficiencies extracted from $t\dot{t}$ (purple histogram). The dashed lines represent the sum in quadrature of the statistical error of the sidebands and the error obtained by propagating the uncertainty on the efficiencies to the matrix method event weights. The line in the bottom panel represents the scale factor in each bin of the distribution, with the dashed band representing its uncertainty.

	ee	$\mu\mu$	OF
$t\bar{t}$ MC	53.4 $\pm$ 3.8	57.5 $\pm$ 4.0	112.8 $\pm$ 5.6
MM fakes ($t\bar{t}$)	59.7 $\pm$ 3.8	61.4 $\pm$ 4.9	101.2 $\pm$ 5.9
Non-closure (ζ)	10.6 $\pm$ 8.5 [%]	6.3 $\pm$ 9.8 [%]	-11.5 $\pm$ 8.5 [%]
SF	1	1	1
ζ systematic uncertainty	11 [%]	10 [%]	11 [%]

Table 23: Total yield of $t\bar{t}$ events with fakes based on simulation, and the yield of fake events estimated with the matrix method applied on the same simulated sample in the 2ℓ SS $0\tau_{had}$, and the non-closure ζ. Errors for the matrix method fakes are statistical plus systematic. If the non-closure ζ is consistent with zero within $\lesssim 1.5\sigma$, a SF of 1 is quoted, meaning no significant method bias is found. The non-closure systematic is taken as $\max(\zeta, \sigma_\zeta)$.

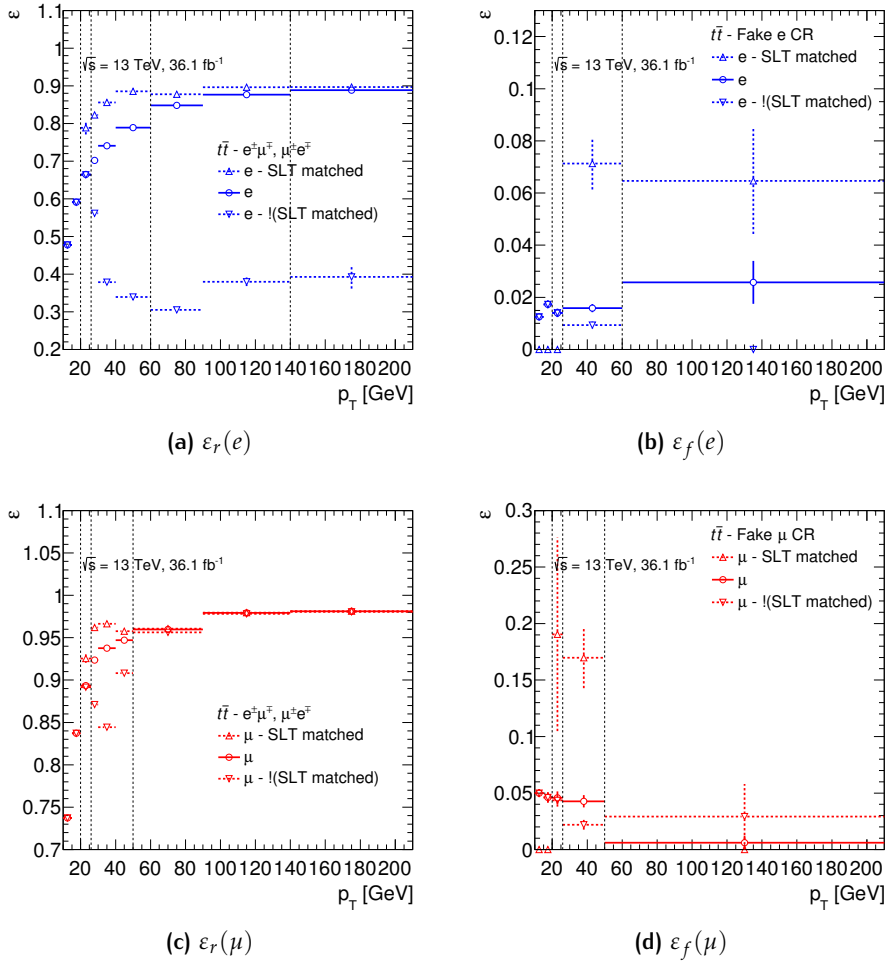
Study of trigger bias effects

In the adopted implementation of the matrix method, efficiencies are measured without any lepton trigger selection, which can be tighter than the offline loose selection. Still, such unbiased efficiencies are used to compute a MM weight for events where one (or both) lepton has fired the trigger, thus carrying its selection bias. This could potentially result in an incorrect fake prediction. Particular care must be taken when single lepton triggers are used to select events both in the signal region and in the MM sidebands, as it is in our case of a `SLT||DLT` trigger selection (see Table 12). In fact, SLTs require quite tight isolation and/or particle ID. On the contrary, dilepton triggers do not have any of these tight selections, therefore they are expected not to bias the offline efficiency measurement.

A study to assess the impact of such an effect has been carried out on $t\bar{t}$ simulated events by comparing the "nominal" matrix method closure test with the one obtained by applying trigger-dependent efficiencies in Eq. (33) selectively on each lepton, depending on whether it fired the single lepton trigger or not (see Table 12). Trigger-dependent efficiencies are measured with the tag-and-probe method previously described, additionally requiring the probe lepton to be either matched to the relevant SLT trigger object, or fail the matching. The p_T binning of the efficiencies has been modified to match with the online p_T thresholds of the SLTs used. Figure 64 illustrates the effect of the trigger bias on the efficiencies in comparison to the baseline measurement.

For p_T below the lowest trigger threshold, events are always being triggered by the tag lepton only, hence no change in the efficiency is seen. As

expected, for probe leptons with p_T above threshold, the probability of passing the offline tight selection is higher when the reconstructed lepton also fired the trigger. For muons, the isolation requirement is applied only for leptons with p_T up to 50 GeV. Above this threshold, no significant difference in the efficiency to pass the offline tight selection is found, regardless of the muon being trigger-matched or not.



(a) $\varepsilon_r(e)$ **(b)** $\varepsilon_f(e)$

(c) $\varepsilon_r(\mu)$ **(d)** $\varepsilon_f(\mu)$

Figure 64: Real and fake lepton efficiencies measured in $t\bar{t}$ simulated events. Dashed lines represent efficiencies for leptons which either passed the SLT trigger (upward triangle markers) or not (downward triangle markers). The vertical lines indicate the online p_T thresholds for the SLTs listed in Table 12.

The non-closure as extracted from the total estimated event yield is shown in Table 24. For the nominal setup, values are slightly different than in Table 23 because of the different binning adopted. Figure 65 displays a comparison of the sub-leading lepton p_T closure plots in the signal region, split by flavour composition.

	ee	$\mu\mu$	OF
	Nominal MM setup		
Non-closure (ζ)	10.7 $\pm$ 8.6 [%]	6.5 $\pm$ 9.8 [%]	-11.0 $\pm$ 8.5 [%]
	Trigger-dependent efficiencies		
Non-closure (ζ)	9.4 $\pm$ 9.3 [%]	9.2 $\pm$ 11.4 [%]	-7.9 $\pm$ 9.6 [%]

Table 24: Comparison of the closure test in the 2ℓ SS $0\tau_{had}$ signal region, split by lepton flavour, in the nominal MM setup and the one with trigger-dependent efficiencies used in the MM equations (see Figure 64). Errors for the matrix method fakes are statistical plus systematic.

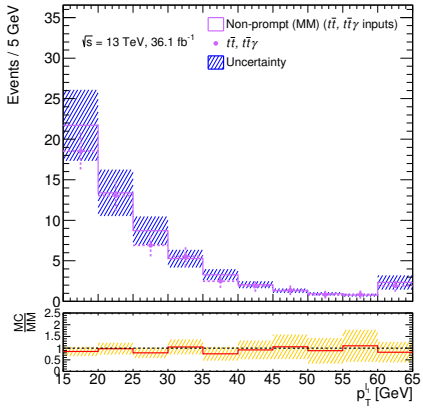
A slightly better overall closure is observed for the trigger-dependent setup, although with larger uncertainties due to the CR events being split depending on the probe matching the trigger. No significant change in the p_T distribution is seen. Given the small gain in using trigger-dependent efficiencies in the matrix method, it is therefore concluded that the usage of a `SLT||DLT` trigger selection is sufficiently robust against trigger bias effects.

Matrix method fakes modelling validation in data

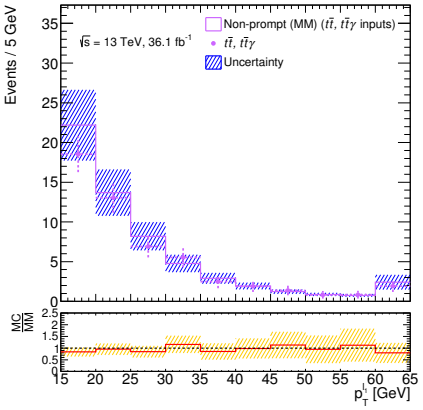
The full data-driven fakes estimation is validated by comparing data with the expected distributions from our background model in a set of validation regions (VRs), designed to be enriched in fakes events and similar to the signal region, albeit orthogonal to it. They are defined by reverting the jet multiplicity cut to include only $2 \leq N_{jets} \leq 3$ events, and are separated in terms of flavour of the leptons in the pair. In the $\mu\mu$ and OF cases, the VRs overlap with the "numerator" subset of the fake control region used for the efficiency measurement. A full set of control plots is displayed in Figures 67, 68, 69, and the normalisation is checked in Figure 66. In all figures, the $t\bar{t}H$ signal is fixed to the SM prediction.

The comparison of data to the total expectation reveals fair modelling of the fakes background as predicted by the matrix method within the uncertainties. In particular, good modelling of the N_{b-tags} distribution in ee can be observed, owing to the usage of the two-dimensional parametrisation for the electron fake rate. Also, the low $\Delta R(\mu, jet)$ tail of the distribution for muon fakes is well modelled.

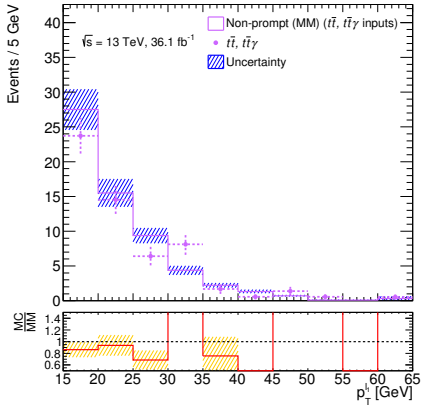
In the 2ℓ SS $0\tau_{had}$ SR, the ratio for the fake background yield estimated by the matrix method in data with respect to the predictions from $t\bar{t}$ simulation is found to be 2.0 ± 0.5 for ee, 1.5 ± 0.5 for $\mu\mu$ and 1.7 ± 0.4 for OF. Errors



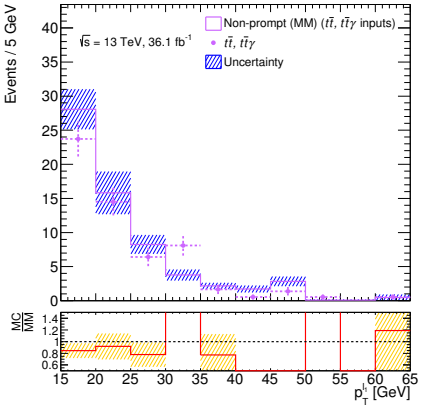
(a) ee, nominal $\varepsilon_r(e)$, $\varepsilon_f(e)$



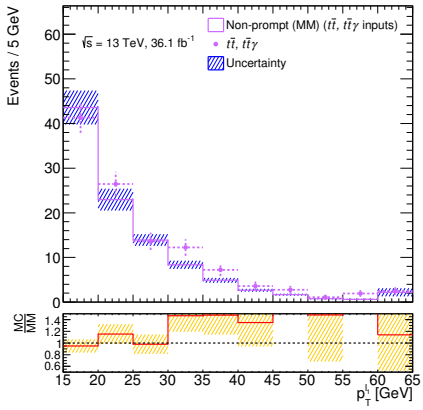
(b) ee, trigger-dependent $\varepsilon_r(e)$, $\varepsilon_f(e)$



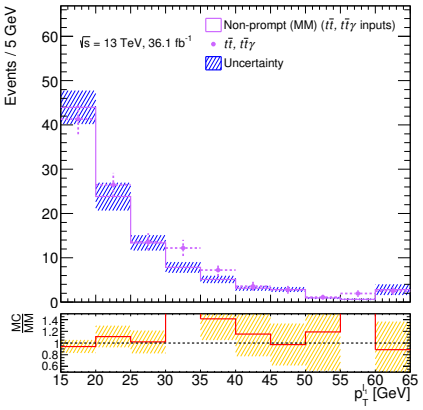
(c) $\mu\mu$, nominal $\varepsilon_r(\mu)$, $\varepsilon_f(\mu)$



(d) $\mu\mu$, trigger-dependent $\varepsilon_r(\mu)$, $\varepsilon_f(\mu)$



(e) OF, nominal $\varepsilon_r(e)$, $\varepsilon_f(e)$, $\varepsilon_r(\mu)$, $\varepsilon_f(\mu)$



(f) OF, trigger-dependent $\varepsilon_r(e,\mu)$, $\varepsilon_f(e,\mu)$

Figure 65: Sub-leading lepton p_T distribution for pure $t\bar{t}$ events (dots) and matrix method fakes with efficiencies extracted from $t\bar{t}$ (purple histogram), for ee, $\mu\mu$, OF SR events. In the left side plots, efficiencies are estimated without any trigger selection on the probe lepton (indicated as "nominal" MM setup). In the right side plots, trigger-dependent efficiencies are used instead in the matrix method equations.

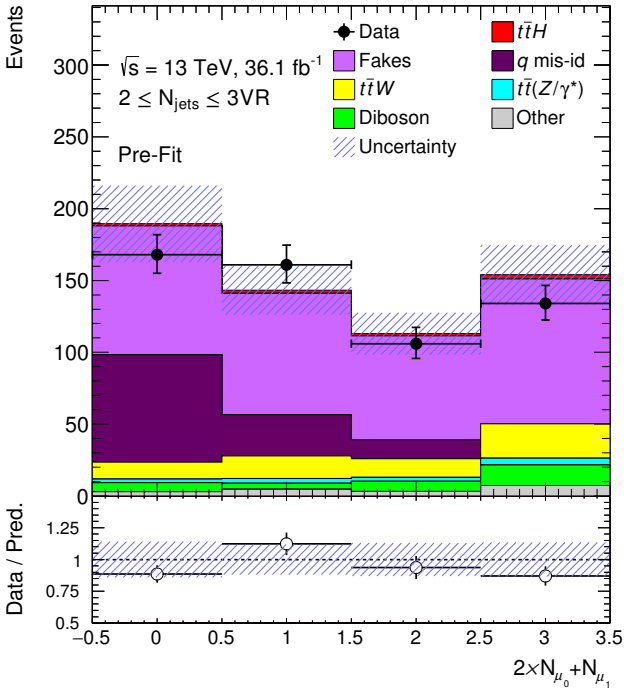


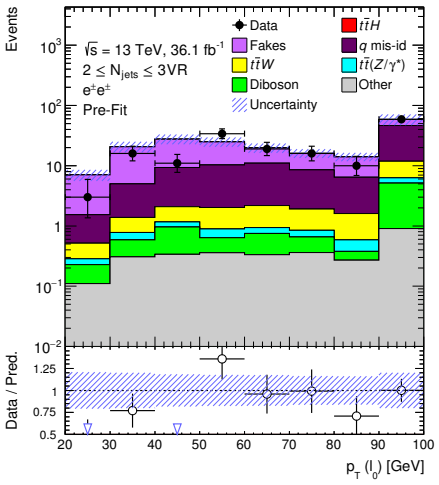
Figure 66: Total event yield in the $2 \leq N_{jets} \leq 3$ validation regions, split by lepton flavour.

include the total uncertainty on the matrix method prediction (see subsection 4.9.1) and the statistical uncertainty on the simulated yield.

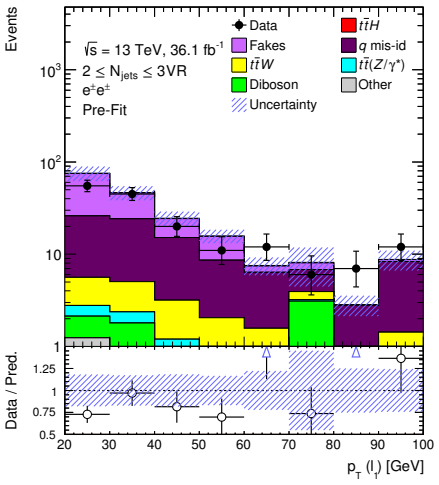
4.9 SYSTEMATIC UNCERTAINTIES

The numerous systematic uncertainties on signal and background modelling and acceptance, as well as on physics object reconstruction, calibration and identification efficiency can affect either the event yield in the signal region (*normalisation* systematics), the kinematic properties of each event (*shape* systematics), or both simultaneously. The effect of the systematics are estimated as a $\pm 1\sigma$ variations around the nominal value of a given quantity.

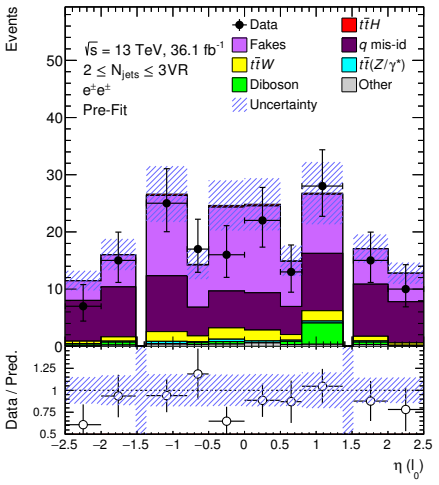
The following paragraphs summarise the evaluation of the detector-related experimental uncertainties, as well as of the systematics related to the signal and background modelling.


Events
Data / Pred.
√s = 13 TeV, 36.1 fb-1
2 ≤ Njets ≤ 3VR
e±e±
Pre-Fit
Data
Fakes
tt̄W
Diboson
Uncertainty
tt̄H
q mis-id
tt̄(Z/γ*)
Other
pT (l0) [GeV]

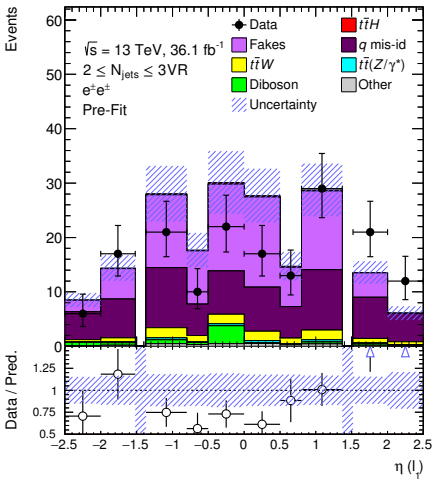
(a)


Events
Data / Pred.
√s = 13 TeV, 36.1 fb-1
2 ≤ Njets ≤ 3VR
e±e±
Pre-Fit
Data
Fakes
tt̄W
Diboson
Uncertainty
tt̄H
q mis-id
tt̄(Z/γ*)
Other
pT (l1) [GeV]

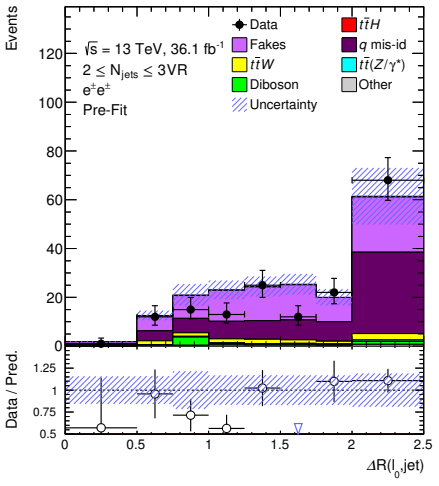
(b)


Events
Data / Pred.
√s = 13 TeV, 36.1 fb-1
2 ≤ Njets ≤ 3VR
e±e±
Pre-Fit
Data
Fakes
tt̄W
Diboson
Uncertainty
tt̄H
q mis-id
tt̄(Z/γ*)
Other
η (l0)

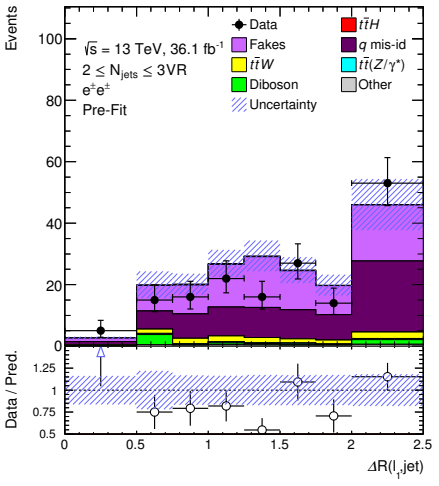
(c)


Events
Data / Pred.
√s = 13 TeV, 36.1 fb-1
2 ≤ Njets ≤ 3VR
e±e±
Pre-Fit
Data
Fakes
tt̄W
Diboson
Uncertainty
tt̄H
q mis-id
tt̄(Z/γ*)
Other
η (l1)

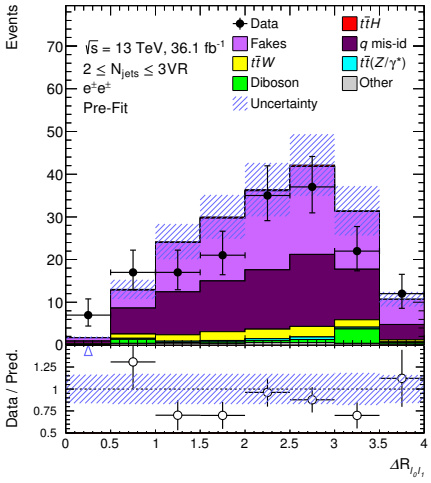
(d)


Events
Data / Pred.
√s = 13 TeV, 36.1 fb-1
2 ≤ Njets ≤ 3VR
e±e±
Pre-Fit
Data
Fakes
tt̄W
Diboson
Uncertainty
tt̄H
q mis-id
tt̄(Z/γ*)
Other
ΔR(l0,jet)

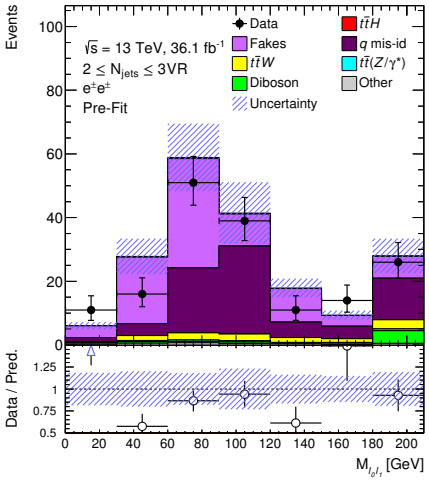
(e)


Events
Data / Pred.
√s = 13 TeV, 36.1 fb-1
2 ≤ Njets ≤ 3VR
e±e±
Pre-Fit
Data
Fakes
tt̄W
Diboson
Uncertainty
tt̄H
q mis-id
tt̄(Z/γ*)
Other
ΔR(l1,jet)

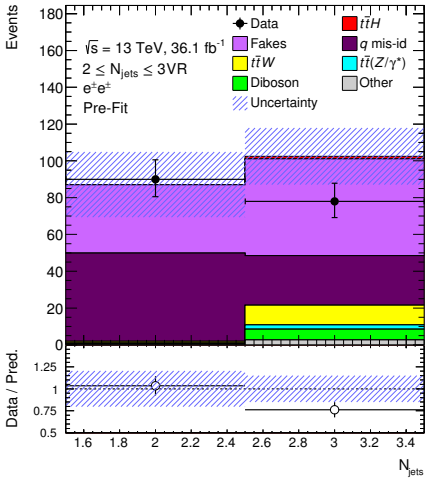
(f)



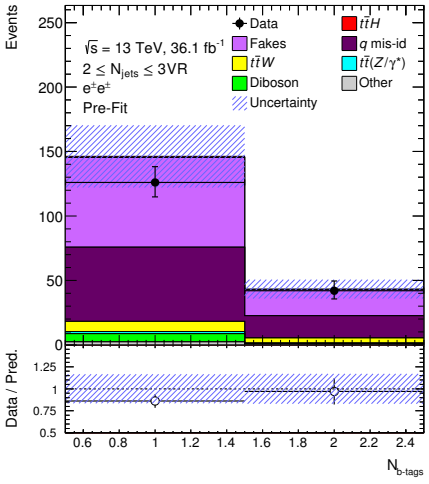
(g)



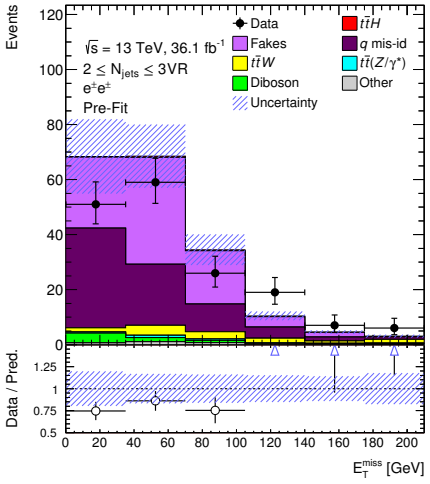
(h)



(i)

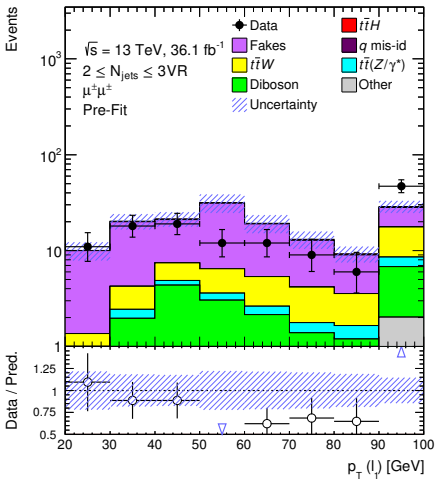


(j)

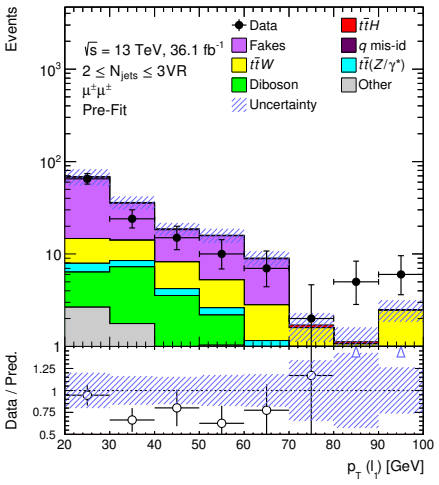


(k)

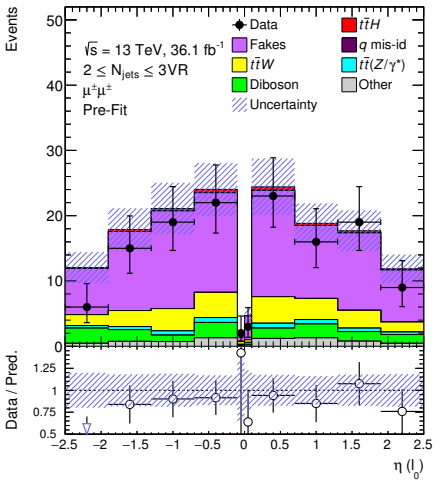
Figure 67: Observed and expected distributions in the *ee*, *TT* leptons, $2 \leq N_{jets} \leq 3$, $N_{b-jets} \geq 1$ validation region. The band in the lower panel represents the total uncertainty on the background model, statistical plus systematic.


Events
Data
Fakes
$t\bar{t}W$
Diboson
Uncertainty
$t\bar{t}H$
q mis-id
$t\bar{t}(Z/\gamma^*)$
Other
$\sqrt{s}$ = 13 TeV, 36.1 fb^{-1}
2 ≤ N$_{jets}$ ≤ 3VR
$\mu^\pm\mu^\pm$
Pre-Fit
Data / Pred.
p$_T$ (l$_1$) [GeV]

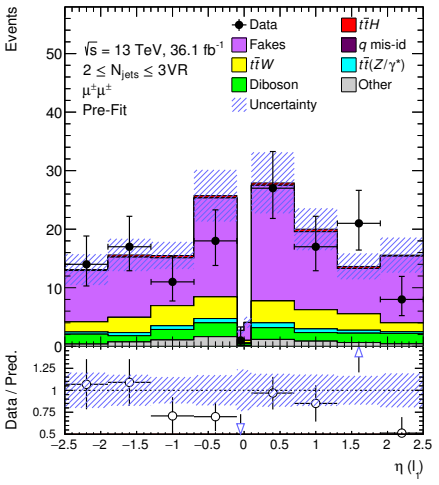
(a)


Events
Data
Fakes
$t\bar{t}W$
Diboson
Uncertainty
$t\bar{t}H$
q mis-id
$t\bar{t}(Z/\gamma^*)$
Other
$\sqrt{s}$ = 13 TeV, 36.1 fb^{-1}
2 ≤ N$_{jets}$ ≤ 3VR
$\mu^\pm\mu^\pm$
Pre-Fit
Data / Pred.
p$_T$ (l$_1$) [GeV]

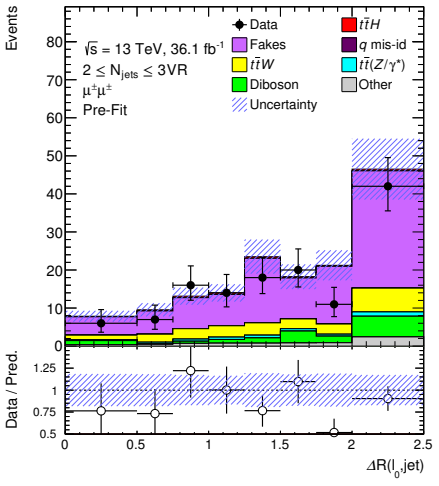
(b)


Events
Data
Fakes
$t\bar{t}W$
Diboson
Uncertainty
$t\bar{t}H$
q mis-id
$t\bar{t}(Z/\gamma^*)$
Other
$\sqrt{s}$ = 13 TeV, 36.1 fb^{-1}
2 ≤ N$_{jets}$ ≤ 3VR
$\mu^\pm\mu^\pm$
Pre-Fit
Data / Pred.
η (l$_0$)

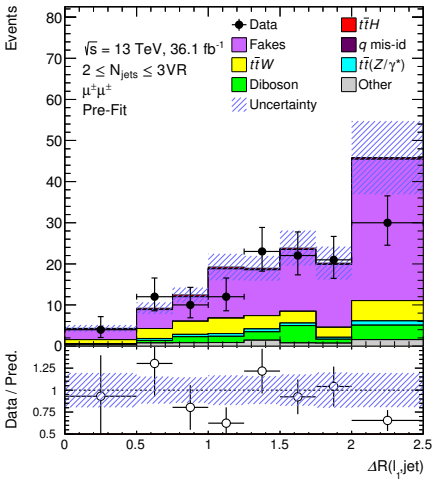
(c)


Events
Data
Fakes
$t\bar{t}W$
Diboson
Uncertainty
$t\bar{t}H$
q mis-id
$t\bar{t}(Z/\gamma^*)$
Other
$\sqrt{s}$ = 13 TeV, 36.1 fb^{-1}
2 ≤ N$_{jets}$ ≤ 3VR
$\mu^\pm\mu^\pm$
Pre-Fit
Data / Pred.
η (l$_1$)

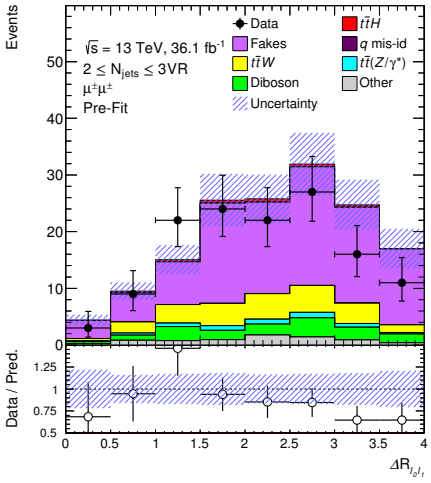
(d)


Events
Data
Fakes
$t\bar{t}W$
Diboson
Uncertainty
$t\bar{t}H$
q mis-id
$t\bar{t}(Z/\gamma^*)$
Other
$\sqrt{s}$ = 13 TeV, 36.1 fb^{-1}
2 ≤ N$_{jets}$ ≤ 3VR
$\mu^\pm\mu^\pm$
Pre-Fit
Data / Pred.
ΔR(l$_0$,jet)

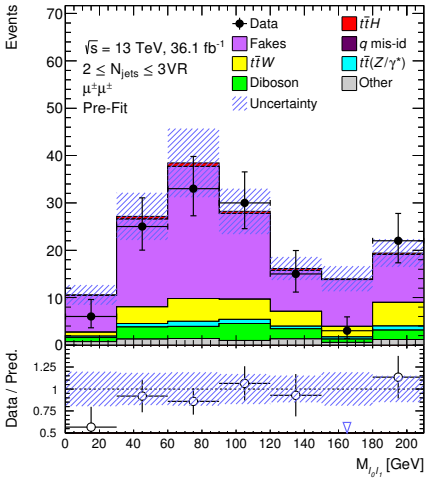
(e)


Events
Data
Fakes
$t\bar{t}W$
Diboson
Uncertainty
$t\bar{t}H$
q mis-id
$t\bar{t}(Z/\gamma^*)$
Other
$\sqrt{s}$ = 13 TeV, 36.1 fb^{-1}
2 ≤ N$_{jets}$ ≤ 3VR
$\mu^\pm\mu^\pm$
Pre-Fit
Data / Pred.
ΔR(l$_1$,jet)

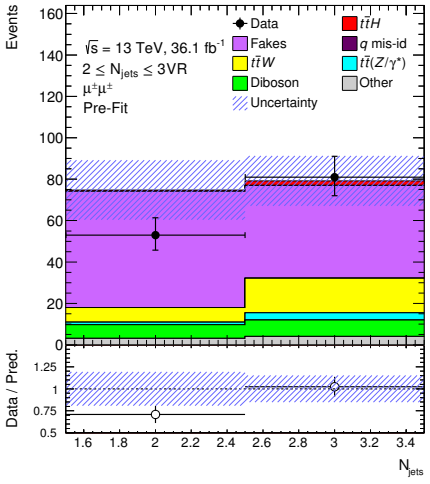
(f)



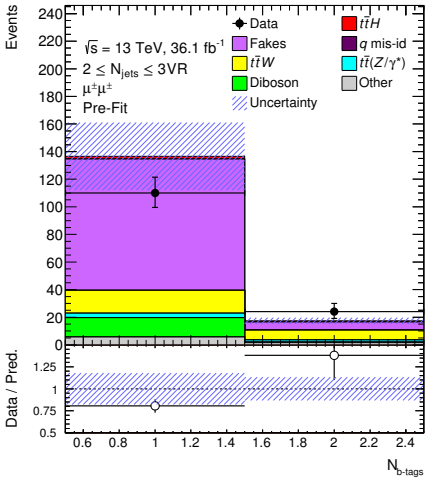
(g)



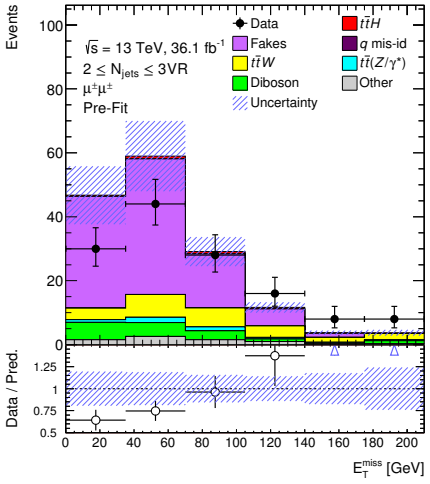
(h)



(i)

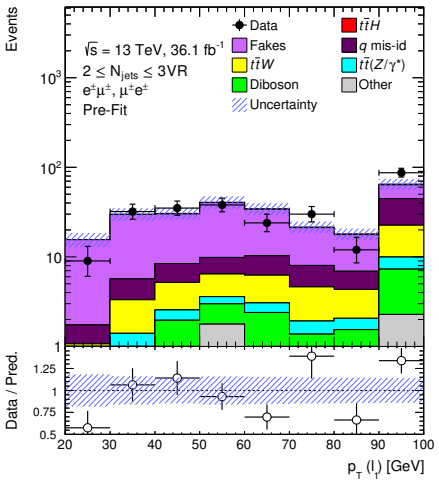


(j)

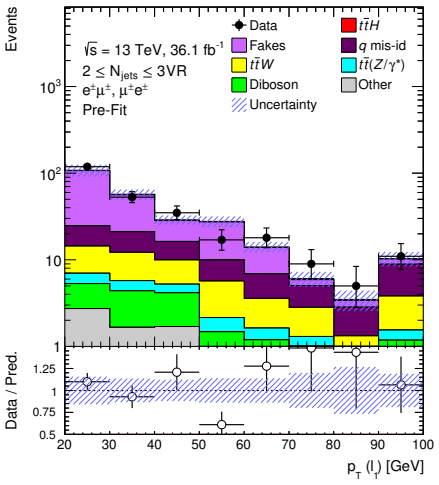


(k)

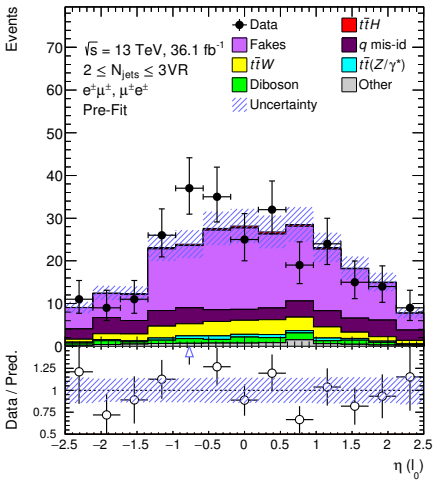
Figure 68: Observed and expected distributions in the $\mu\mu$, TT leptons, $2 \leq N_{jets} \leq 3$, $N_{b-jets} \geq 1$ validation region. The band in the lower panel represents the total uncertainty on the background model, statistical plus systematic.


Events
Data
Fakes
$t\bar{t}W$
Diboson
Uncertainty
$t\bar{t}H$
q mis-id
$t\bar{t}(Z/\gamma^*)$
Other
$\sqrt{s}$ = 13 TeV, 36.1 fb^{-1}
2 ≤ N$_{jets}$ ≤ 3VR
$e^{\pm}\mu^{\pm}$, $\mu^{\pm}e^{\pm}$
Pre-Fit
Data / Pred.
p$_T$ (l$_1$) [GeV]

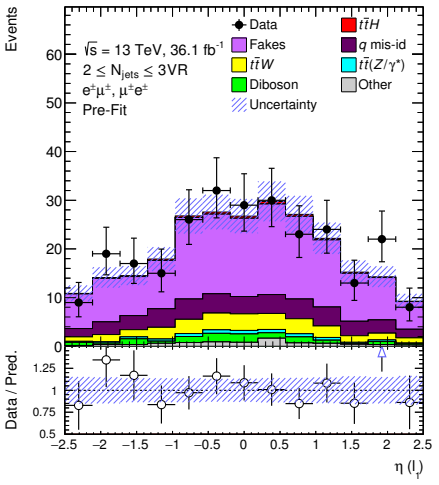
(a)


Events
Data
Fakes
$t\bar{t}W$
Diboson
Uncertainty
$t\bar{t}H$
q mis-id
$t\bar{t}(Z/\gamma^*)$
Other
$\sqrt{s}$ = 13 TeV, 36.1 fb^{-1}
2 ≤ N$_{jets}$ ≤ 3VR
$e^{\pm}\mu^{\pm}$, $\mu^{\pm}e^{\pm}$
Pre-Fit
Data / Pred.
p$_T$ (l$_1$) [GeV]

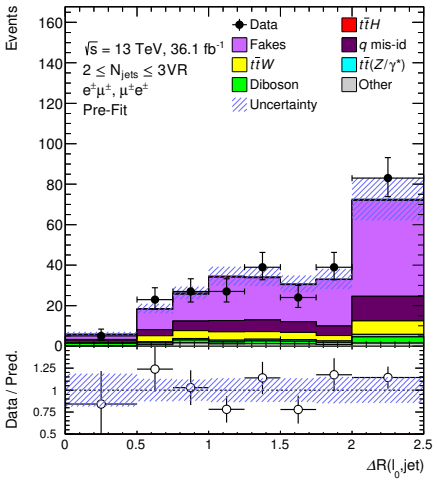
(b)


Events
Data
Fakes
$t\bar{t}W$
Diboson
Uncertainty
$t\bar{t}H$
q mis-id
$t\bar{t}(Z/\gamma^*)$
Other
$\sqrt{s}$ = 13 TeV, 36.1 fb^{-1}
2 ≤ N$_{jets}$ ≤ 3VR
$e^{\pm}\mu^{\pm}$, $\mu^{\pm}e^{\pm}$
Pre-Fit
Data / Pred.
η (l$_0$)

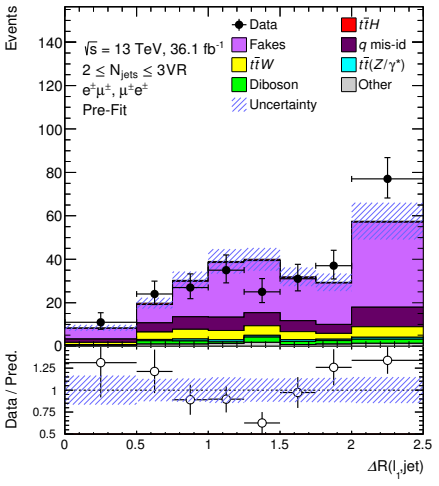
(c)


Events
Data
Fakes
$t\bar{t}W$
Diboson
Uncertainty
$t\bar{t}H$
q mis-id
$t\bar{t}(Z/\gamma^*)$
Other
$\sqrt{s}$ = 13 TeV, 36.1 fb^{-1}
2 ≤ N$_{jets}$ ≤ 3VR
$e^{\pm}\mu^{\pm}$, $\mu^{\pm}e^{\pm}$
Pre-Fit
Data / Pred.
η (l$_1$)

(d)


Events
Data
Fakes
$t\bar{t}W$
Diboson
Uncertainty
$t\bar{t}H$
q mis-id
$t\bar{t}(Z/\gamma^*)$
Other
$\sqrt{s}$ = 13 TeV, 36.1 fb^{-1}
2 ≤ N$_{jets}$ ≤ 3VR
$e^{\pm}\mu^{\pm}$, $\mu^{\pm}e^{\pm}$
Pre-Fit
Data / Pred.
ΔR(l$_0$,jet)

(e)


Events
Data
Fakes
$t\bar{t}W$
Diboson
Uncertainty
$t\bar{t}H$
q mis-id
$t\bar{t}(Z/\gamma^*)$
Other
$\sqrt{s}$ = 13 TeV, 36.1 fb^{-1}
2 ≤ N$_{jets}$ ≤ 3VR
$e^{\pm}\mu^{\pm}$, $\mu^{\pm}e^{\pm}$
Pre-Fit
Data / Pred.
ΔR(l$_1$,jet)

(f)

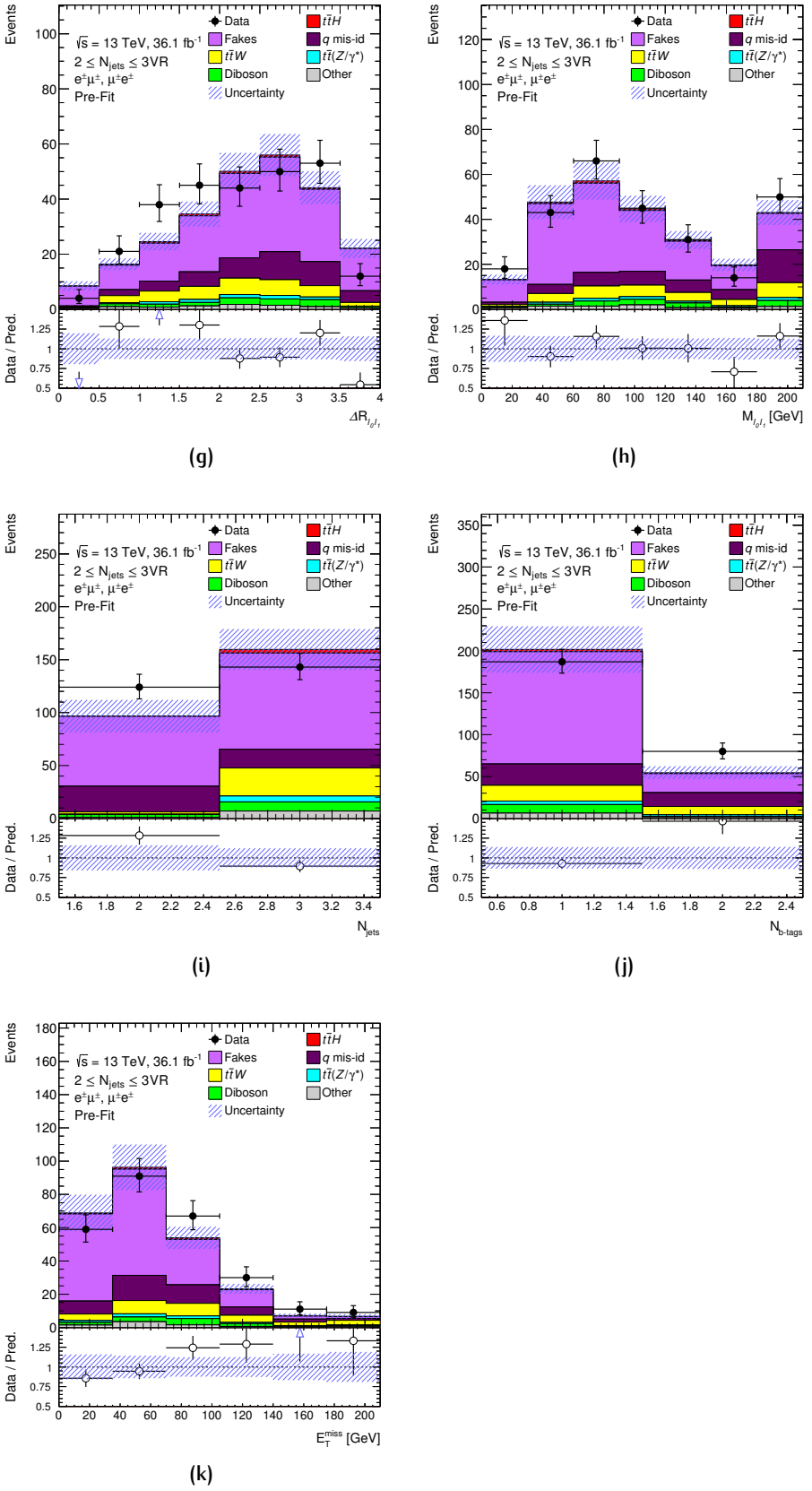


Figure 69: Observed and expected distributions in the *OF*, *TT* leptons, $2 \leq N_{jets} \leq 3$, $N_{b-jets} \geq 1$ validation region. The band in the lower panel represents the total uncertainty on the background model, statistical plus systematic.

4.9.1 Fake lepton background uncertainties

Systematic uncertainties on the data-driven, reducible fake lepton background derive from:

- The uncertainty on the normalisation of the background processes subtracted in the real/fake control regions.
- The statistical uncertainty of the control regions, i.e., the statistical uncertainty on the efficiencies.
- The statistical uncertainty of the SR and the loose sidebands regions.
- The uncertainty on the simulation-based α rescaling procedure for conversions.
- The uncertainty on the non-closure.

In the real control region, a conservative 30% uncertainty is assigned on the fake $t\bar{t}$ opposite-sign background normalisation. This is eventually shown to have no sizeable impact on the final result, therefore it is neglected. In the fake control regions, a theory uncertainty of 14% is assigned to the $t\bar{t}W$, $t\bar{t}Z$ normalisation to account for the production cross-section uncertainties from variations of QCD scales and PDF+α_s, following the prescription described in subsection 4.3.1. In accordance with the discussion in section 4.7, a 50% uncertainty is assigned on diboson processes, mostly WZ + 1 b-jet events. This is a conservative estimate based on the low purity achieved in our WZ validation region for constraining its normalisation. Finally, a 30% normalisation uncertainty is chosen for all the remaining simulated processes to be subtracted, namely $t\bar{t}$ events where the probe lepton was mis-assigned and rare SM processes. As for the charge flip subtraction, the total statistical plus systematic uncertainty on the data-driven estimate is accounted for as a normalisation uncertainty. The background theory uncertainties are treated as fully correlated across the efficiency bins, and between the numerator and denominator subsets of the control regions. In total, four systematic uncertainties account for variations of the background subtraction. All detector-related systematic uncertainties affecting simulation are neglected, as we assume they are sub-dominant compared to the size of the background normalisation uncertainties.

The uncertainties on the efficiencies are propagated to the event weights, Eq. (31). Each bin of the fake efficiency parametrisation has been varied independently from the others by its statistical uncertainty, leading to a different

systematic event weight for each bin variation. For real efficiency, no difference is seen when considering a fully correlated variation, likely due to the small size of this uncertainty. This approach is thus followed for simplicity. In total, 7 components are considered for the statistical uncertainty on the efficiencies.

Two α_i scaling factors (see Eq. (34)) are relevant to the 2ℓSS0τ signal region: α_{ee} and α_{OF}. A 40% systematic uncertainty on the $f_\gamma^{X_i}$ fractions is assigned. This is a conservative estimate, based on the linear sum of a 20% uncertainty estimated from studies of the Inner Detector material description [152], and a 20% uncertainty on the measured $t\bar{t}\gamma$ cross section [153]. This systematic is assumed to be fully correlated between the OF control region and all the regions to which the efficiency is extrapolated. A 50% uncertainty is then assigned on both ε_γ and ε_{HF}, based on the observed differences in the average fake rates between simulation and data. Under the assumption that all sources of uncertainties are independent, propagating them in Eq. (34) leads to a 42% relative uncertainty on each α_i. The absolute value of the uncertainty on α_i is then assigned as a relative uncertainty of that size on the final electron fakes normalisation in ee, OF: two uncertainties in total are accounted for. No shape uncertainty is considered, as the fraction of conversions is observed to be fairly stable across the bins of the final BDT discriminant to be fitted (Fig 70). The values of α_i's and their uncertainties are displayed in Table 25.

Finally, uncertainties arising from the method's non-closure are included in Table 23. Three systematic uncertainties are considered, of order 10% in all sub-categories and affecting the fake normalisation only, given the flat shape of the bias (see the bottom panel of Figure 63).

All the sources of systematic uncertainties on the fakes are summarised in Table 26. The purely statistical uncertainty arising from the limited sample of data events in the signal region and the three sidebands of the matrix method is also included in the table. A set of 6 independent uncertainties is taken, one for each bin of the final BDTG score distribution which is used in the fit. The size of this uncertainty is found to be about 5% in the ee and OF regions, up to 11% in $\mu\mu$. The total number of systematic uncertainties associated to the matrix method fake estimate is 22.

Overall, the total size of the uncertainty on fakes in the 2ℓ SS $0\tau_{had}$ SR is around $20-30\%$, depending on the lepton flavour composition, and is obtained by summing in quadrature each contribution. For all channels, the major source of uncertainty is the limited size of the fake control regions, followed by the non-closure. In the ee case, the uncertainty on the photon

	$f_\gamma^{X_i}$ [%]	ε_{HF}	ε_γ	α
		0.009 ± 0.005	0.057 ± 0.029	
OF, $2 \leq N_{jets} \leq 3$ (Electron Fake CR)	27 ± 11			—
ee, (2ℓ SR)	46 ± 18			$0.42\ ^{+0.06}_{-0.09}\ (f_\gamma)\ ^{+0.12}_{-0.09}\ (\varepsilon_{HF})\ ^{+0.07}_{-0.16}\ (\varepsilon_\gamma) = 0.42 \pm 0.18$
OF, (2ℓ SR)	30 ± 12			$0.07\ ^{+0.01}_{-0.01}\ (f_\gamma)\ ^{+0.02}_{-0.01}\ (\varepsilon_{HF})\ ^{+0.01}_{-0.02}\ (\varepsilon_\gamma) = 0.07 \pm 0.03$

Table 25: Fraction of fake electrons from photon conversion in the $2\ell SS0\tau$ signal region split by flavour composition (excluding $\mu\mu$) as illustrated in Figure 54, and the α extrapolation factors defined in Eq. (34), with their total uncertainty (including a breakdown of the different sources). The efficiencies for HF fakes and photon conversion fakes as extracted from $t\bar{t}$ simulation are shown as well.

conversion rescaling factor is also relevant. The uncertainties from the processes subtraction are below 10%.

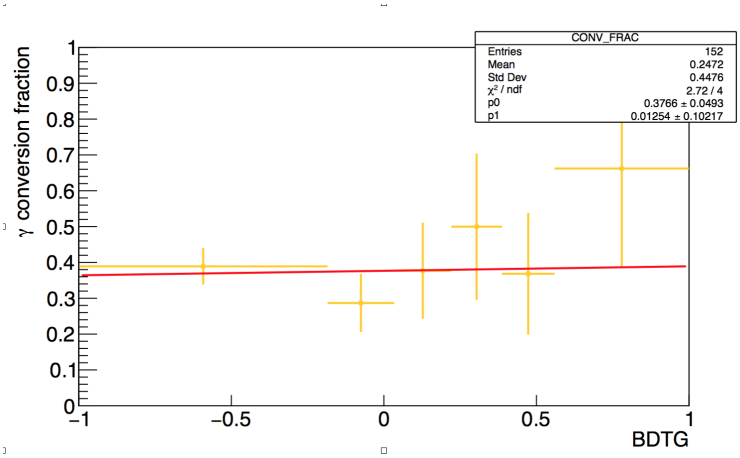


Figure 70: Fraction of fake electrons from photon conversions in the *ee* signal region as a function of the final BDTG discriminant. A linear fit of the histogram shows no significant slope.

4.9.2 Theory uncertainties

The systematic uncertainties associated with the generation of signal and irreducible background processes stem from the uncertainties in the predicted cross sections, and from the modelling of the acceptance for each process.

Source		Efficiency affected	Origin	Correlated across ε bins, N/D regions	Components	*ee* SR	*OF* SR	$\mu\mu$ SR
Theory (σ)	$t\bar{t}W$,$t\bar{t}Z$	$\varepsilon_f(e,\mu)$	14% on norm.	yes	1	5.4%	3.8%	5.5%
	Diboson	$\varepsilon_f(e,\mu)$	50% on norm.	yes	1	6.2%	4.9%	8.7%
	Other SS	$\varepsilon_f(e,\mu)$	30% on norm.	yes	1	3.0%	2.8%	4.8%
Charge flip		$\varepsilon_f(e)$	data-driven charge flip estimate	yes	1	5.2%	2.8%	—
CR stat. unc.		all ε's	stats. in ε bins	no	7	17%	16%	22%
α		$\varepsilon_f(e)$	knowledge of γ conversions in $t\bar{t}$ MC	yes	2	18%	3%	—
Non-closure		—	method bias and $t\bar{t}$ stats.	N/A	3	11%	11%	10%
Stat. unc.		—	SR and sidebands size	N/A	6	4.5%	6.1%	11.2%
Tot. uncertainty						30%	22%	29%

Table 26: Summary of the sources of uncertainty (systematic and statistical) on the matrix method fake estimate, and the associated number of uncertainty components. The three columns on the right display the relative impact of each uncertainty on the fake estimate in the SR, split by flavour composition; the total uncertainty is computed as the sum in quadrature of each contribution. The sources listed in the top part of the table directly affect the real or fake efficiencies, and are therefore propagated through the full matrix method weight computation. A shifted event weight for each source is eventually obtained; in the "CR size" case, one independent systematic variation for each fake efficiency bin is considered. The non-closure uncertainty in the bottom part of the table is considered to affect only the final event yield normalisation.

The estimation of the systematics on the production cross sections have been presented in subsection 4.3.1. They are accounted for as affecting only the total event yield. For all the rare SM background processes listed in Table 9 for which no validation regions could be defined, a 50% uncertainty on the cross section is assigned conservatively.

The SM Higgs boson branching ratio uncertainties for the decays to WW^*, ZZ^*, $\tau\tau$, as well as the one on the inclusive BR for all other SM decays, are considered as four other sources of systematic uncertainty. Much alike the cross section, they are normalisation-only uncertainties, and are listed in Table 2. These are assumed to affect also the tH samples.

The acceptance of a process can be affected by the choice of the QCD renormalisation and factorisation scales, since they have an impact on the shape of the differential cross section. The effect of the independent and simultaneous variations of both μ_R and μ_F up and down by a factor two is encoded as a set of three event weights in the $t\bar{t}H$, $t\bar{t}W$ and $t\bar{t}Z$ simulated samples. These are derived by re-weighting the nominal kinematic distributions at the parton level according to the procedure described in [154]. The choice and configuration of the parton showering simulation - as well as the UE and shower tune - can also alter the acceptance through modifications of the jet multiplicity and kinematics. The chosen A14 tune is the result of the optimisation of 10 parameters [132], grouped into three sets which are sensitive to underlying event effects, jet shapes and structure, and different aspects of extra jets production. The effect of the variation of each sets of tune parameters has been investigated and the latter set is observed to be the dominant one. Therefore, only one additional systematic event weight is taken into account for tune variation. For $t\bar{t}H$, Herwig++ has been utilised as an alternative simulation to Pythia 8 to estimate the uncertainty related to the modelling of parton showering. For $t\bar{t}W$ and $t\bar{t}Z$, the effect of a different choice of ME generator - namely, Sherpa - is considered as another systematic uncertainty on the acceptance, affecting both shape and normalisation. As for diboson processes, the Sherpa generator implementation of the ME-PS combination involves the presence of an additional merging scale - defined in terms of a jet resolution scale: any parton produced above threshold is generated with a corresponding higher-order matrix element and, conversely, any parton produced below is generated in the shower. The variation of μ_R, μ_F in the ME, of the parton shower starting scale, the ME-PS merging scale and the tune variation are embedded into a single event weight with up/down variations, which results in a shape and normalisation systematic uncertainty.

4.9.3 Experimental uncertainties

Several observables relevant to the analysis are subject to experimental systematic uncertainties. Such uncertainties are related to:

- The luminosity measurement and the description of the global event activity (i.e., the pile-up).
- The lepton momentum/energy calibration, and the measurement of lepton reconstruction, trigger and identification efficiency.
- The jet energy calibration and resolution, and the pile-up jet suppression.
- The measurement of the b-tagging efficiency and mis-tag rates [9] for b-jets and charm/light-flavoured jets, respectively.

Most of these uncertainties are directly estimated by the ATLAS combined performance groups, and are implemented either as a variation on an event weight or - in case of calibration uncertainties - through rescaling of the object energy or tranverse momentum.

Luminosity and event activity (pile-up modelling)

The uncertainty on the integrated luminosity is 2.1%. It is derived following a methodology similar to that detailed in [68], from a preliminary calibration of the luminosity scale using *x-y* beam separation scans performed in August 2015 and May 2016. As described in subsection 4.3.1, simulated events are weighted to correctly reproduce the distribution of the average number of p-p collisions per bunch crossing found in the data. The accuracy of the average correction factor is of order $4 - 6\%$. The pile-up reweighting uncertainty originate from the modelling of minimum bias events, including uncertainties in the p-p inelastic cross section. The final effect of such uncertainty is less than 1% variation of the total simulated event yield in the SR.

Lepton-related uncertainties

The measurement of the light lepton reconstruction, identification and trigger efficiency has been documented in subsection 2.3.2 and subsection 2.3.3. The calibration of these efficiencies in simulation to the collision data is performed with respect to the tight lepton selection based on the multi-variate

9 i.e., the efficiency for non-b jets to be tagged as b.

algorithms specifically designed for this analysis, as previously described in subsection 4.4.5. The uncertainty on the measured scale factors is generally small, around a few per-cent. Such uncertainties are considered to affect the event normalisation only. Their impact on the simulated event yield in the 2ℓ SS $0\tau_{had}$ signal region is found to be of order 1%. The muon and electron momentum (or energy) scale and resolution uncertainties modify the lepton kinematics, thus can affect the selection acceptance. Similarly to the efficiency calibration, these uncertainties are estimated exploiting data events at the Z resonance [85, 87]. They are found overall to have negligible impact.

Jet-related uncertainties

The methodology followed by ATLAS to calibrate the jet energy scale (JES) and resolution (JER) was summarised in subsection 2.3.4.

Several sources contribute to the jet energy scale systematic uncertainty [93, 155, 94]:

- The uncertainty on the *in situ* energy scale measurement in data with different methods.
- The uncertainty on the relative calibration across different η regions of the detector, known as *inter-calibration*.
- The uncertainty due to the difference in the calorimeter response to quark-initiated and gluon-initiated jets, as well as to the originating quark flavour.
- The uncertainty on the procedure to correct for the pile-up contributions to the jet energy.
- The uncertainty on the behaviour of high-p_T jets in the propagation of single hadron uncertainties to a jet.
- The uncertanity on the punch-through effect.

In this analysis, such sources are decomposed into a set of 21 uncorrelated components, treated as independent systematics terms. Out of this set, 8 systematic terms account for the in situ calibrations uncertainties, the other 13 describing the remaining sources. Due to the sizeably different quark flavour composition of the various processes, the JES uncertainties related to jet flavour composition are treated as uncorrelated systematic components affecting $t\bar{t}Z$, diboson, and the grouped $t\bar{t}W/t\bar{t}H$/other processes, respectively. This results into 3 separate components for this uncertainty source.

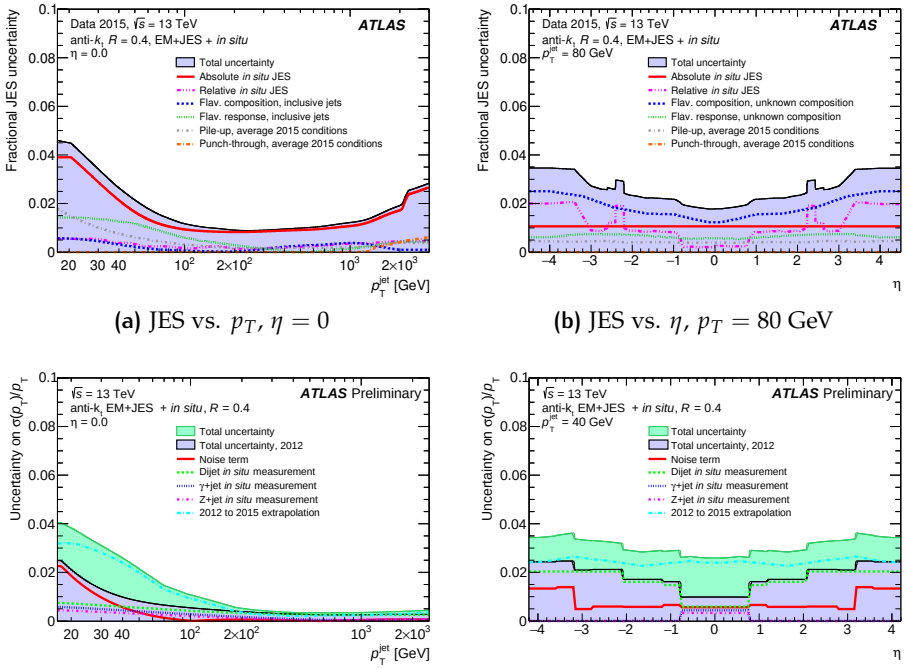
Figure 71 displays the total JES uncertainty as well as a breakdown into its main sources as a function of jet p_T and η. The total per-jet JES uncertainty is found to be within $1-6\%$, depending on the jet transverse momentum. The jet energy scale uncertainty has a sizeable impact on the analysis, resulting in a 5% uncertainty on the SR simulated event yield. This is understandable, given the large jet multiplicity of the signal region. Specifically, the largest contributions come from the uncertainty on the pile-up subtraction [156], the in situ calibration, and the different response to quark and gluon jets.

The jet energy resolution (JER) systematic uncertainty arise from differences between several in situ measurement methods, and variations in the noise thresholds defined for the topoclustering algorithm [157]. The individual sources for the JER are displayed in Figure 71. All systematic sources are found to have a similar shape, therefore they are combined into one systematic term to be propagated in a correlated way. Finally, the uncertainty on the JVT association efficiency is determined from the variation of the corresponding cut, and it induces an $\mathcal{O}(1\%)$ variation on the simulated SR event yield.

The MV2c10 b-tagging efficiency for b-jets and the mis-tag rate for c-jets and light flavour jets are measured in data with several independent techniques [98].The uncertainties in the b-tagging efficiencies and mis-tag rates correction factors account for multiple sources depending on each available method. Similarly to the JES, they are grouped into a set of 126 uncorrelated components. Their size is of order 2% for b-jets, 10% for c-jets and 30% for light jets. Due to the limited number of selected b-jets in the signal region - typically 1 or 2 - the impact of the b-tagging efficiency uncertainty is found to be rather small, of $\mathcal{O}(1.5\%)$ on the simulated event yield.

4.9.4 Summary of systematic uncertainties

A summary of all systematic uncertainties affecting the 2ℓ SS $0\tau_{had}$ signal region is presented in Table 27. The relative impact of each source of systematic uncertainty on the event yield of the process (or combination of processes) affected by the systematic in question is displayed. The uncertainties shown are symmetrised for simplicity.



(a) JES vs. p_T, $\eta = 0$

(b) JES vs. η, $p_T = 80$ GeV

(c) JER vs. p_T, $\eta = 0$

(d) JER vs. η, $p_T = 40$ GeV

Figure 71: Top row: combined uncertainty on the JES of fully calibrated jets as a function of 71a jet p_T at $\eta = 0$ and 71b η at p_T = 80 GeV. Systematic uncertainty components include pile-up, punch-through, and uncertainties propagated from in situ measurements and inter-calibration. The flavor composition and response uncertainties assume a quark and gluon composition taken from simulation [155]. Bottom row: jet energy resolution uncertainties estimated for 2015 data as a function of 71c jet p_T at $\eta = 0$ and 71d η at p_T = 40 GeV [157].

Systematic uncertainty	Type	Sample(s) affected	Size
Experimental			
Luminosity	N	All simulated	±2%
Pile-up modelling	SN	All simulated	$< 1\%$
e (ε's, scale, resolution)	SN	All simulated	$\approx \pm 1\%$
μ (ε's, scale, resolution)	SN	All simulated	$\approx \pm 1\%$
JVT	SN	All simulated	$< 1\%$
JES, JER	SN	All simulated	±5%
b-tagging	SN	All simulated	±2%
Reducible backgrounds			
SR and sidebands statistics	SN	Fakes	±4%
Background sub.	N	Fakes	±10%
Real/fake light lepton efficiencies	SN	Fakes	±16%
Matrix method non-closure	N	Fakes	±7%
α conversion fraction for conversions	N	Fakes	±6%
Electron charge mis-assignment	SN	Charge flip	±35%
Signal and irreducible backgrounds			
Cross section (QCD, PDF+α_s)	N	$t\bar{t}H$	±11%
Higgs branching ratios	N		±6%
QCD scales	S		±14%
Parton shower model	SN		±1%
Shower tune	SN		±4%
Cross section (QCD, PDF+α_s)	N	$t\bar{t}W$	±12%
QCD scales	S		±11%
Generator	SN		±2%
Parton shower tune	SN		±3%
Cross section (QCD, PDF+α_s)	N	$t\bar{t}Z$	±11%
QCD scales	S		±13%
Generator	SN		±13%
Parton shower tune	SN		±1%
Cross section	N	Diboson	±50%
QCD Scale, ME-PS, tune	SN		±36%
Cross section	N	Rare SM	±50%

Table 27: Summary of systematic uncertainties for the 2ℓ SS $0\tau_{had}$ channel of the $t\bar{t}H \rightarrow \text{multi} - \text{leptons}$ analysis. "N" indicates the uncertainty affects normalisation only, whereas "S" denotes systematics that are considered affecting only shapes. The combination "SN" implies the uncertainty affects both shape and normalisation. The rightmost columnn displays the overall effect of each systematic variation(s) on the event yield of the affected sample(s) in the SR. The up/down variations shown are symmetrised.

5 RESULTS OF THE $t\bar{t}H$ TO MULTI-LEPTONS ANALYSIS

5.1 THE STATISTICAL INTERPRETATION OF THE ANALYSIS

Observed collision data and the prediction for signal and backgrounds are eventually translated into a statistical model to determine the best-fit estimate for $\mu_{t\bar{t}H} = \sigma/\sigma_{SM}$, and to assess the relationship between the data and the prediction within the framework of frequentist *hypothesis testing*.

The model is implemented using the `TRExFitter` [158] package - which relies on the `HistFactory` [159] software of the `RooFit/RooStat` framework - and can handle:

- Signal and multiple background processes, encoded into histogram templates for a binned distribution of interest.
- Multiple signal regions of interest (channels).
- Unconstrained normalisation parameters.
- Normalisation and shape variations associated to systematic uncertainties, and constraints on them. The systematic uncertainties are encoded in templates of variations relative to the nominal histogram for each $\pm 1\sigma$ variation.
- Variations associated to the statistical uncertainty on the prediction.
- Correlations of systematic uncertainties between different processes, and across channels.

The description of the underlying key concepts of the statistical model is provided in the following sections.

5.1.1 Definition of the likelihood function

We start by considering a binned distribution of events in the signal region, i.e. the BDTG output histogram in Figure 78. The number of expected events in each bin i - $E[n_i]$ - is determined by the model described in Chapter 4, and is parametrised as [160]:

$$\begin{aligned} E[n_i] = \nu_i = \mu s_i + B_i = \mu s_i + \sum_{b \in \{bkgs\}} b_i \\ s_i = N_s \int_{bin_i} f_s(x, \boldsymbol{\theta_s}) dx \\ b_i = N_b(\boldsymbol{\theta}_b) \int_{bin_i} f_b(x, \boldsymbol{\theta_b}) dx, \end{aligned} \tag{39}$$

where:

- s_i and B_i indicate the mean signal and total background yield per bin, respectively. The latter is the sum of all the individual background processes b_i considered.
- The parameter μ quantifies the abundance of signal in units of SM prediction: a value of 1 represents the nominal SM prediction, whereas a value of 0 indicates absence of any signal.
- The normalisation factors N_s, $N_b(\boldsymbol{\theta}_b)$ are the overall signal and background events. The signal normalisation N_s is assumed to be fixed to the value predicted by the model.
- f_s, f_b are the signal and background probability density functions for a signal (background) event to be found in the i-th bin. The vectors $\boldsymbol{\theta}_s$, $\boldsymbol{\theta}_b$ represent a set of parameters which determine the signal and background shape through the pdfs, and the background normalisation through $N_b(\boldsymbol{\theta}_b)$.

The $t\bar{t}H$ signal strength $\mu_{t\bar{t}H} = \sigma/\sigma_{SM}$ (assuming a Higgs boson mass of $m_H = 125$ GeV) is chosen as the *parameter of interest* (PoI) of our test. It represents a free parameter of the model, meaning no prior constraints are placed on it. The vector $\boldsymbol{\theta} = \{\boldsymbol{\theta}_s, \boldsymbol{\theta}_b\}$ contains the so-called *nuisance parameters* (NPs), which encode all the uncertainties on quantities that can affect the model for signal and background. Unlike μ, the nuisance parameter probability distributions are constrained by the auxiliary measurements or theoretical predictions of the parameters ($\boldsymbol{\theta_0}$) previously illustrated in section 4.9. The

choice of the probability density function for the NP that defines the constraint depends on the type of the associated observable. It is important to notice that, since our PoI definition is based on the SM prediction for $\sigma_{t\bar{t}H}$, it depends on $\boldsymbol{\theta}_s^{\sigma}$, that is, the nuisance parameters associated to the predicted signal cross section uncertainties.

The number of observed events in a given bin (n) is assumed to be described by a Poissonian probability distribution with expectation value $\nu(\boldsymbol{\theta_0}) = \mu s(\boldsymbol{\theta_0}) + b(\boldsymbol{\theta_0})$:

$$\text{Pois}(n;\nu(\boldsymbol{\theta_0})) = \frac{(\nu(\boldsymbol{\theta_0}))^n}{n!}\mathrm{e}^{-(\nu(\boldsymbol{\theta_0}))}. \tag{40}$$

A nuisance parameter θ affecting the shape of signal and/or background can take both positive and negative values, and it is expected to be described by a Gaussian probability density:

$$\text{Gaus}(\theta;\theta_0,\sigma_{\theta_0}) = \frac{1}{\sqrt{2\pi}\sigma_{\theta_0}}\exp\left(-\frac{(\theta-\theta_0)^2}{2{\sigma_{\theta_0}}^2}\right), \tag{41}$$

where θ_0 and σ_{θ_0} indicate the prior estimate of the parameter and its uncertainty, respectively.

Nuisance parameters associated to quantities that affect the normalisation - such as the luminosity, efficiencies and cross sections - are described by a log-normal *pdf* of variance $\kappa_0(\theta_0,\sigma_{\theta_0})$, in order to avoid any unphysical negative yield estimate:

$$\begin{gathered}\text{LogN}(\theta;\theta_0,\kappa_0) = \frac{1}{\sqrt{2\pi}\ln(\kappa_0)}\frac{1}{\theta}\exp\left(-\frac{\ln^2(\theta/\theta_0)}{2\ln^2(\kappa_0)}\right) \\ \kappa_0 = \exp(2\theta_0+\sigma_{\theta_0}^2)\left[\exp(\sigma_{\theta_0}^2)-1\right]\end{gathered} \tag{42}$$

Additional NPs (γ) are associated to the rate in each bin of the total simulated background and the data-driven background. They are assumed to be described by a gamma probability density function - an extension of the Poisson distribution for real numbers - constrained by the statistical uncertainty γ_0 of the bin in question [161].

$$\text{Gamma}(\gamma;\gamma_0) = \frac{1}{\Gamma(\gamma_0+1)}\gamma^{\gamma_0}\mathrm{e}^{-\gamma} \tag{43}$$

The global likelihood function is constructed by taking the product of the blocks representing the likelihood terms for each set of model parameters described above:

$$\begin{aligned}\mathcal{L}(N_{obs}, \boldsymbol{\theta_0}|\mu, \boldsymbol{\theta}) = \mathcal{L}_{evt}(N_{obs}|\mu) \times \mathcal{L}_{aux}(\boldsymbol{\theta_0}|\boldsymbol{\theta}) = \\ \prod_{i}^{N_{bins}} \text{Pois}\left(n_i|\nu_i(\boldsymbol{\theta})\right) \times \\ \prod_{j\in\{\text{shape NPs}\}} \text{Gaus}\left(\theta_{0j}, \sigma_{\theta_{0j}}|\theta_j\right) \times \\ \prod_{k\in\{\text{norm. NPs}\}} \text{LogN}\left(\theta_{0k}, \kappa_{0k}|\theta_k\right) \times \\ \prod_{l}^{N_{bins}} \text{Gamma}\left(\gamma_{0l}|\gamma_l\right)\end{aligned} \tag{44}$$

Given the number of observed events N_{obs} and our prior estimation of $\boldsymbol{\theta}_0$, numerical maximisation [162] of this multi-dimensional likelihood function eventually leads to a best-fit estimate for μ and $\boldsymbol{\theta}$, that is, their maximum likelihood estimators (MLE) $\hat{\mu}$ and $\hat{\boldsymbol{\theta}}$. In practice, the equivalent minimisation of the negative log-likelihood (NLL) - $-2\ln(\mathcal{L})$ - is performed. This greatly simplifies the procedure by transforming the products into sums - so that constant terms can be neglected - and by removing the exponential terms.

It is assumed that any new signal can only increase ν_i beyond what is expected from background alone, that is, any *physical* estimator of the signal strength must satisfy $\mu \geq 0$. However, it should be noted that the MLE $\hat{\mu}$ can be negative, such in the case when a downward fluctuation in data results in less observed events than even background alone would predict. This is fine, so long as the expectation value for the number of events in each bin is non-negative.

5.1.2 Definition of test statistic for discovery

In order to establish a potential discovery of the $t\bar{t}H$ process, the compatibility of the observed data with the prediction under the background-only *null* hypothesis (namely, $\mu = 0$) must be assessed. A statistically significant comparison is found if the relationship between the two datasets represents an unlikely realisation of the null hypothesis according to a certain probability threshold. This is quantified by defining a suitable *test statistic*. In particle physics experiments, the optimal test statistic definition when nuisance parameters are taken into account is based on the profile likelihood ratio [163], given the likelihood function definition of Eq. (44):

$$\lambda(\mu) = \begin{cases} \frac{\mathcal{L}(\mu,\hat{\hat{\boldsymbol{\theta}}}(\mu))}{\mathcal{L}(\hat{\mu},\hat{\boldsymbol{\theta}})} & \text{for } \hat{\mu} \geq 0, \\ \frac{\mathcal{L}(\mu,\hat{\hat{\boldsymbol{\theta}}}(\mu))}{\mathcal{L}(0,\hat{\boldsymbol{\theta}}(0))} & \text{for } \hat{\mu} < 0. \end{cases} \tag{45}$$

The $\hat{\hat{\boldsymbol{\theta}}}$ at the numerator denotes the conditional maximum likelihood estimator of $\boldsymbol{\theta}$, given a specific value of μ. The denominator represents the unconditional maximum likelihood, given the MLEs for μ and $\boldsymbol{\theta}$. The definition of $\lambda(\mu)$ ensures that in case a negative $\hat{\mu}$ is found from the fit due to downward fluctuations, then the best agreement between model and data happens at $\mu = 0$. The presence of the nuisance parameters broadens the profile likelihood as a function of μ relative to what one would have if their values were fixed. This reflects the loss of information about μ due to the systematic uncertainties.

From Eq. (45) it holds $0 \leq \lambda \leq 1$. Given an hypothesised value of μ, it can be seen that $\lambda \approx 1$ implies good agreement between data and the hypothesis. It is convenient to define the test statistics as:

$$t_\mu = -2\ln\lambda(\mu) = \begin{cases} -2\ln\left(\frac{\mathcal{L}(\mu,\hat{\hat{\boldsymbol{\theta}}}(\mu))}{\mathcal{L}(\hat{\mu},\hat{\boldsymbol{\theta}})}\right) & \text{for } \hat{\mu} \geq 0, \\ -2\ln\left(\frac{\mathcal{L}(\mu,\hat{\hat{\boldsymbol{\theta}}}(\mu))}{\mathcal{L}(0,\hat{\boldsymbol{\theta}}(0))}\right) & \text{for } \hat{\mu} < 0, \end{cases} \tag{46}$$

so that higher values of t_μ correspond to increasing incompatibility between the data - as described by the MLE - and the hypothesis of having a signal with strength μ.

The level of disagreement is quantified by the *p-value*:

$$p_\mu = \int_{t_{\mu,obs}}^{\infty} f(t_\mu|\mu)dt_\mu \tag{47}$$

where $t_{\mu,obs}$ is the test statistics value obtained from the observed data, and $f(t_\mu|\mu)$ represents the pdf of the test statistic for a μ hypothesis.

When performing a statistical test for the discovery of a new signal with non-negative strength such as $t\dot{t}H$, the hypothesis under test to exclude is the background-only hypothesis, i.e. $\mu = 0$. This implies Eq. (46) for the test statistics becomes:

$$t_0 = \begin{cases} -2\ln\lambda(0) & \text{for } \hat{\mu} \geq 0, \\ 0 & \text{for } \hat{\mu} < 0, \end{cases} \tag{48}$$

and the p-value to quantify the level of disagreement with the null hypothesis is $p_0 = \int_{t_{0,obs}}^{\infty} f(t_0|0)dt_0$. In the particle physics community, it is customary to map p-values into the *significance Z*:

$$Z = \Phi^{-1}(1 - p), \tag{49}$$

where Φ^{-1} is the inverse of the cumulative distribution of a standard Gaussian probability density function. In particle physics, the rejection of the background-only hypothesis to claim for a discovery is conventionally achieved for a significance of $Z \geq 5$, corresponding to $p \leq 2.87 \times 10^{-7}$. This represents the well-known 5σ threshold to state the *observation* of a new phenomenon. *Evidence* can be claimed if a significance of $Z \geq 3$ is found.

In order to calculate Z, it is necessary to determine the sampling distribution for our test statistic, $f(t_0|0)$. This in general would require generating a large number of pseudo-experiments, which is computationally expensive. Nonetheless, for a test statistics based on the profile likelihood ratio and for its pdf as well, analytic approximations independent of the NPs can be derived in the asymptotic large sample limit, owing to a result due to Wilks and Wald [164, 165]. The validity of the asymptotic limit has been proven to hold true for even as few as $\mathcal{O}(10)$ events in a data sample [160]. The analytic function for Eq. (46) depends only on $\hat{\mu}$ and on its variance. The latter can be estimated from the so-called *Asimov dataset*, an artificial representative dataset defined as the one in which all observed quantities are set equal to their expectation values and all statistical fluctuations are suppressed. In the asymptotic limit (changing the notation $t \to q$), we find for the test statistics:

$$q_0 = \begin{cases} \hat{\mu}^2/\sigma^2 & \text{for } \hat{\mu} \geq 0, \\ 0 & \text{for } \hat{\mu} < 0, \end{cases} \tag{50}$$

where $\hat{\mu}$ is Gaussian-distributed with mean μ' and variance σ^2. For the discovery testing, $\mu' = 0$, and eventually we obtain a simple expression for the sensitivity:

$$Z_0 = \Phi^{-1}(1 - p_0) = \sqrt{q_0}. \tag{51}$$

The Asimov dataset can also be exploited to characterise the *expected* sensitivity of the search, that is, the median p-value for the rejection of the background-only hypothesis under a specific assumption μ' on the signal model. To test the sensitivity of the search to the presence of the SM $t\bar{t}H$ mechanism, it is therefore assumed $\mu' = 1$.

	Yield (pre-fit)
Fakes	233 ± 39
Charge flips	33 ± 11
$t\bar{t}W$	123 ± 18
$t\bar{t}Z$	41.4 ± 5.6
Diboson	25 ± 15
Others	28.4 ± 5.9
Tot. Background	484 ± 38
$t\bar{t}H$	42.6 ± 4.2
Observed	514

Table 28: Observed and expected event yield for data and background processes in the 2ℓ SS $0\tau_{had}$ signal region before the fit. Uncertainties are statistical plus systematic (see section 4.9).

5.2 EVENT YIELDS AND MODELLING OF BDTG INPUTS BEFORE THE FIT FOR THE 2ℓ SS $0\tau_{had}$ CHANNEL

The event counts in the 2ℓ SS $0\tau_{had}$ signal region before the fit[1] are given in Table 28 for all relevant processes, with errors including both statistical and systematic uncertainties. The dominant backgrounds are the fake lepton background and $t\bar{t}W$, accounting for about 48% and 30% of the total background, respectively.

Before any attempt at a statistical interpretation of the analysis result, it is essential to ensure the robustness of the modelling of all the relevant observables, namely the quantities used as input in the BDTGs and the output of the BDTGs themselves.

Figure 72 displays the pre-fit distributions of the input variables to the $BDTG_{t\bar{t}}$, $BDTG_{t\bar{t}V}$ discriminants in the signal region. Good agreement between data and prediction is seen in the N_{jets}, N_{b-tags}, lepton flavour, $\Delta R(\ell, jet)$ and $\max(|\eta_\ell|)$ distributions. A deficit of data events of order 1σ is found in the $20 < p_T(\ell_1) < 30$ GeV bin. As Figure 73 shows, this region is dominated by events with fakes, whose majority ($\approx 75\%$) are events with electrons. The mis-modelling can be mostly traced back to a similar one observed in the sub-leading lepton p_T spectrum in the fake-dominated ee low jet multiplicity VR shown in Figure 67. Different parametrisations and binning of both electron and muon fake rates have been tested, but no significant improvement of the modelling was found.

1 In the following, the "pre-fit" terminology is also used

Another notable feature in the $p_T(\ell_1)$ distribution is a $\approx 2\sigma$ excess of data at very high sub-leading lepton transverse momentum. This is correlated with the $\approx 1.5\sigma$ excess found at high $m_{\ell\ell}$. However, as Figure 74 reveals, no striking features appear when looking at several distributions for high $p_T(\ell_1)$ events. Note that unlike the low $p_T(\ell_1)$ corner of the final state, this represents a region where the fake background contribution is sub-dominant.

5.2.1 Combination of BDTG outputs

The pre-fit distributions of the output of $BDTG_{t\bar{t}V}$ and $BDTG_{t\bar{t}}$ are displayed in Figure 75.

Good modelling is observed overall, except for a $\approx 1\sigma$ data deficit in the $-1 < BDTG_{t\bar{t}} < -0.4$ region. Further investigation of this mis-modelling is shown in Figure 76, which displays several input distribution of the BDTGs for events falling in those bins. Such events are characterised by having mostly low $p_T(\ell_1)$ and a fake lepton background purity of about 65%. Specifically, they are dominated by events with fake electrons, accounting for 88% of total fake background. The discrepancy is therefore related to the one discussed previously about the $p_T(\ell_1)$ distribution. Otherwise, no striking features are observed.

Maximal sensitivity of the analysis can be achieved by taking advantage of the shape description of both BDTG outputs simultaneously.

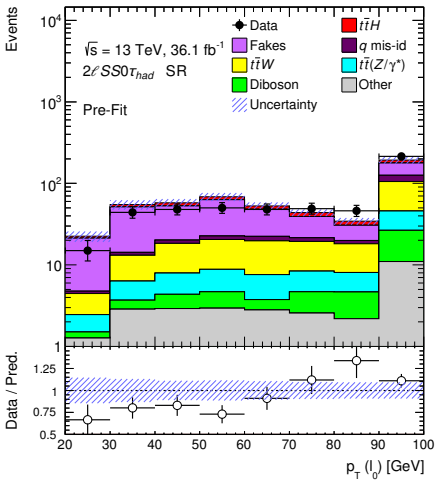
Several ways of combining $BDTG_{t\bar{t}V}$ and $BDTG_{t\bar{t}}$ were tried, such as selecting rectangular regions of the ($BDTG_{t\bar{t}}$,$BDTG_{t\bar{t}V}$) plane via a grid scan to try defining bins with high purity of $t\bar{t}H$, $t\bar{t}W$ and fake background events, respectively. The best sensitivity is eventually achieved by taking a linear combination of $BDTG_{t\bar{t}V}$ and $BDTG_{t\bar{t}}$, according to the following mapping:

$$BDTG = \frac{BDTG_{t\bar{t}} + a \times BDTG_{t\bar{t}V}}{1 + a}. \tag{52}$$

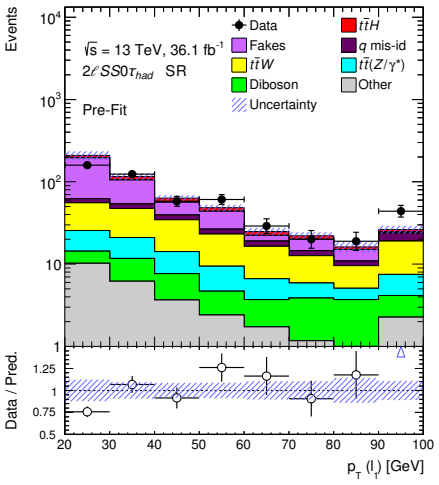
This quantity represents geometrically a plane surface, where the parameter $-\frac{1}{a}$ is the slope of the lines obtained by projecting the surface on the ($BDTG_{t\bar{t}}$,$BDTG_{t\bar{t}V}$) plane, as illustrated in Figure 77.

After an optimisation process [158], the best sensitivity is reached for $a = 1$. This corresponds to taking the simple arithmetic average of the two discriminators:

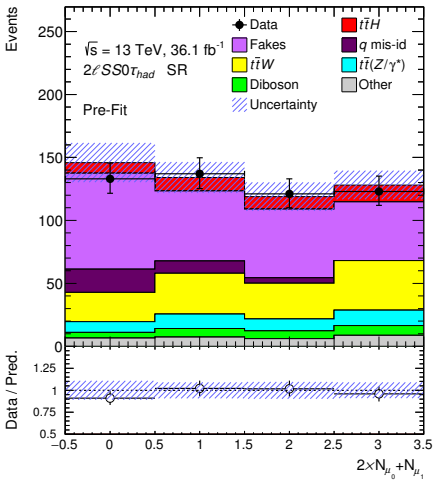
$$BDT = \frac{BDTG_{t\bar{t}} + BDTG_{t\bar{t}V}}{2}. \tag{53}$$


Events
Data / Pred.
√s = 13 TeV, 36.1 fb-1
2ℓSS0τhad SR
Pre-Fit
Data
Fakes
tt̄W
Diboson
Uncertainty
tt̄H
q mis-id
tt̄(Z/γ*)
Other
pT (l0) [GeV]

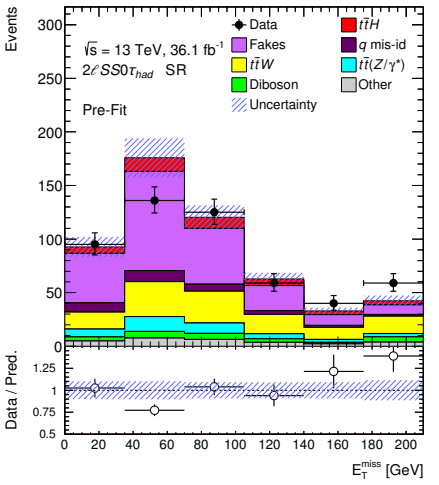
(a)


Events
Data / Pred.
√s = 13 TeV, 36.1 fb-1
2ℓSS0τhad SR
Pre-Fit
Data
Fakes
tt̄W
Diboson
Uncertainty
tt̄H
q mis-id
tt̄(Z/γ*)
Other
pT (l1) [GeV]

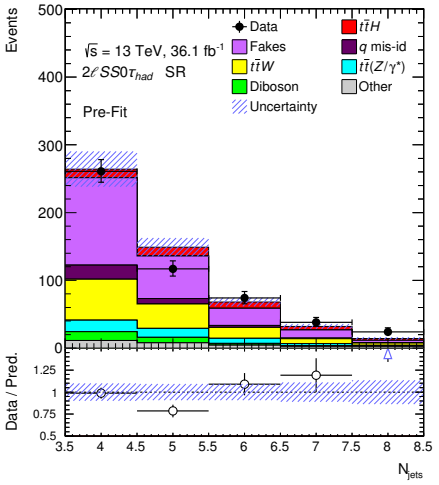
(b)


Events
Data / Pred.
√s = 13 TeV, 36.1 fb-1
2ℓSS0τhad SR
Pre-Fit
Data
Fakes
tt̄W
Diboson
Uncertainty
tt̄H
q mis-id
tt̄(Z/γ*)
Other
2×Nμ0 +Nμ1

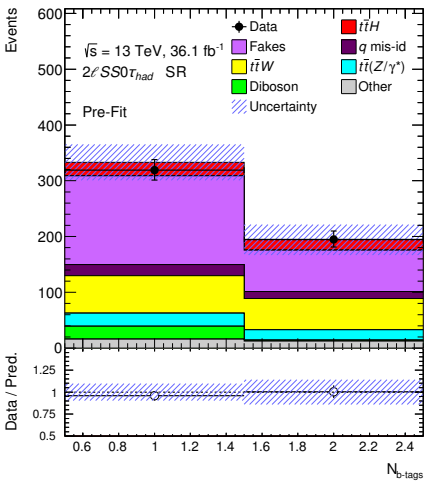
(c)


Events
Data / Pred.
√s = 13 TeV, 36.1 fb-1
2ℓSS0τhad SR
Pre-Fit
Data
Fakes
tt̄W
Diboson
Uncertainty
tt̄H
q mis-id
tt̄(Z/γ*)
Other
ETmiss [GeV]

(d)


Events
Data / Pred.
√s = 13 TeV, 36.1 fb-1
2ℓSS0τhad SR
Pre-Fit
Data
Fakes
tt̄W
Diboson
Uncertainty
tt̄H
q mis-id
tt̄(Z/γ*)
Other
Njets

(e)


Events
Data / Pred.
√s = 13 TeV, 36.1 fb-1
2ℓSS0τhad SR
Pre-Fit
Data
Fakes
tt̄W
Diboson
Uncertainty
tt̄H
q mis-id
tt̄(Z/γ*)
Other
Nb-tags

(f)

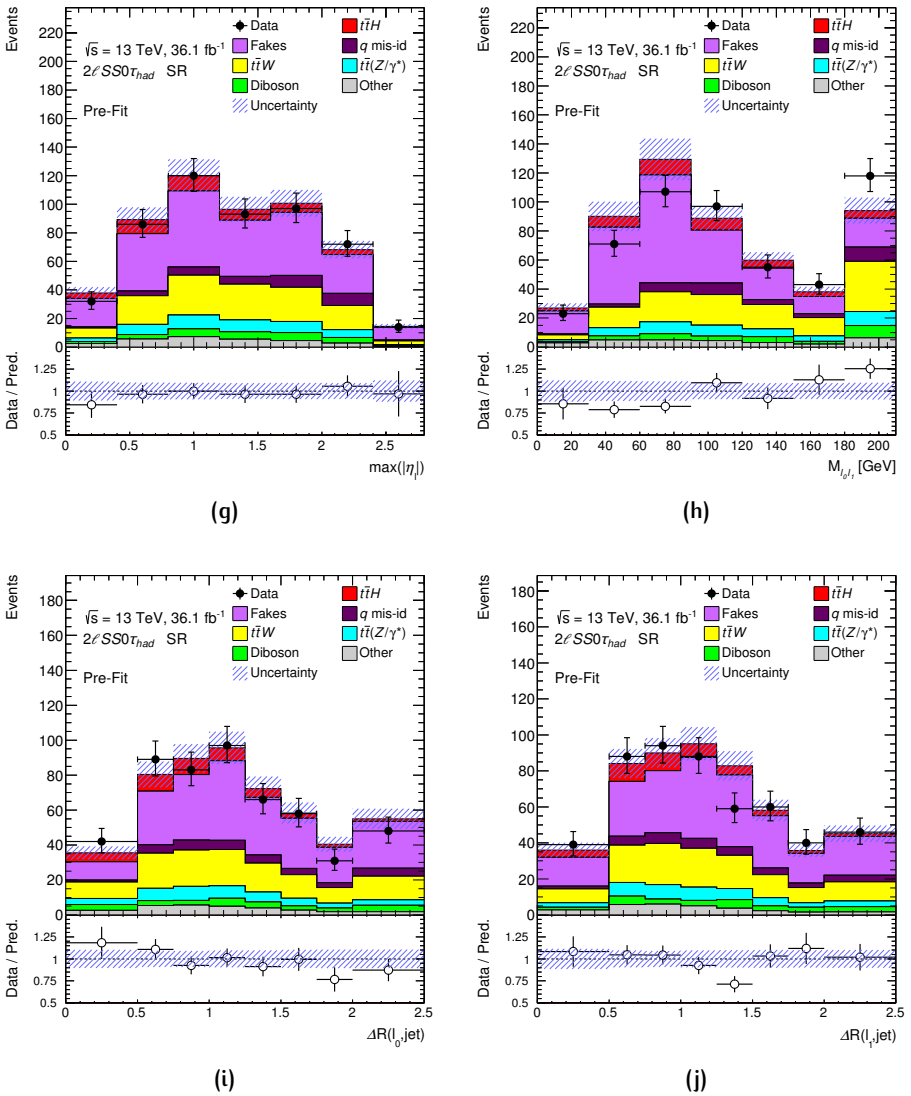


Figure 72: Pre-fit distributions of the 9 input variables to the BDTG algorithm in the 2ℓ SS $0\tau_{had}$ signal region (with the addition of $p_T(\ell_0)$). Events with fake leptons and charge flip electrons are estimated in a data-driven way, whereas irreducible background are simulated. Errors include the statistical uncertainty and all systematic uncertainties described in Table 27. The first and last bins include underflow and overflow events.

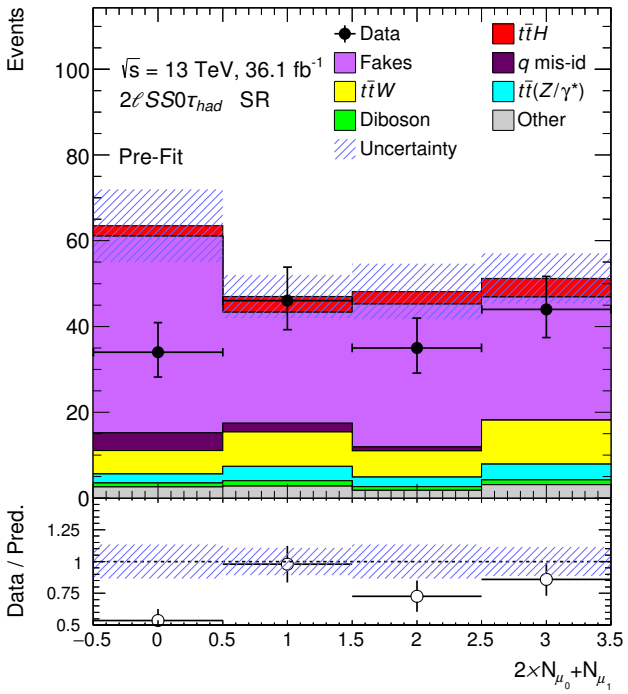


Figure 73: Total event yield in the 2ℓ SS $0\tau_{had}$ signal region for events with $20 < p_T(\ell_1) < 30$ GeV, split by lepton flavour.

Optimal sensitivity is then found by grouping the BDTG discriminant values into 6 bins of different size, imposing the signal yield to be constant across the bins. As a result, the lowest BDTG bins are more populated by events with fake leptons, the central bins by $t\bar{t}W$, $t\bar{t}Z$ and the rightmost bins have the highest purity of $t\bar{t}H$ signal events.

Given an observable y with probability density function $\hat{y}_{s,b}$ for signal and background, the *separation* is defined as:

$$\langle S^2 \rangle = \frac{1}{2} \int \frac{(\hat{y}_s(y) - \hat{y}_b(y))^2}{\hat{y}_s(y) + \hat{y}_b(y)} dy. \tag{54}$$

This integral is equal to zero for identical signal and background PDF shapes, and is equal to one for non-overlapping shapes.

Figure 78 shows the output of the final $BDTG$ discriminant. The achieved separation between the signal shape and the total background shape according to the previous definition is around 16%. The pre-fit modelling of the signal and background distribution is good within the uncertainties. A slight deficit of data characterises the first bin of the distribution, although still covered by the estimated uncertainty. This region is dominated by fake events, therefore the discrepancy is driven by the one previously discussed about the $BDTG_{t\bar{t}}$ shape, traced back in turn to the mis-modelling at low trans-

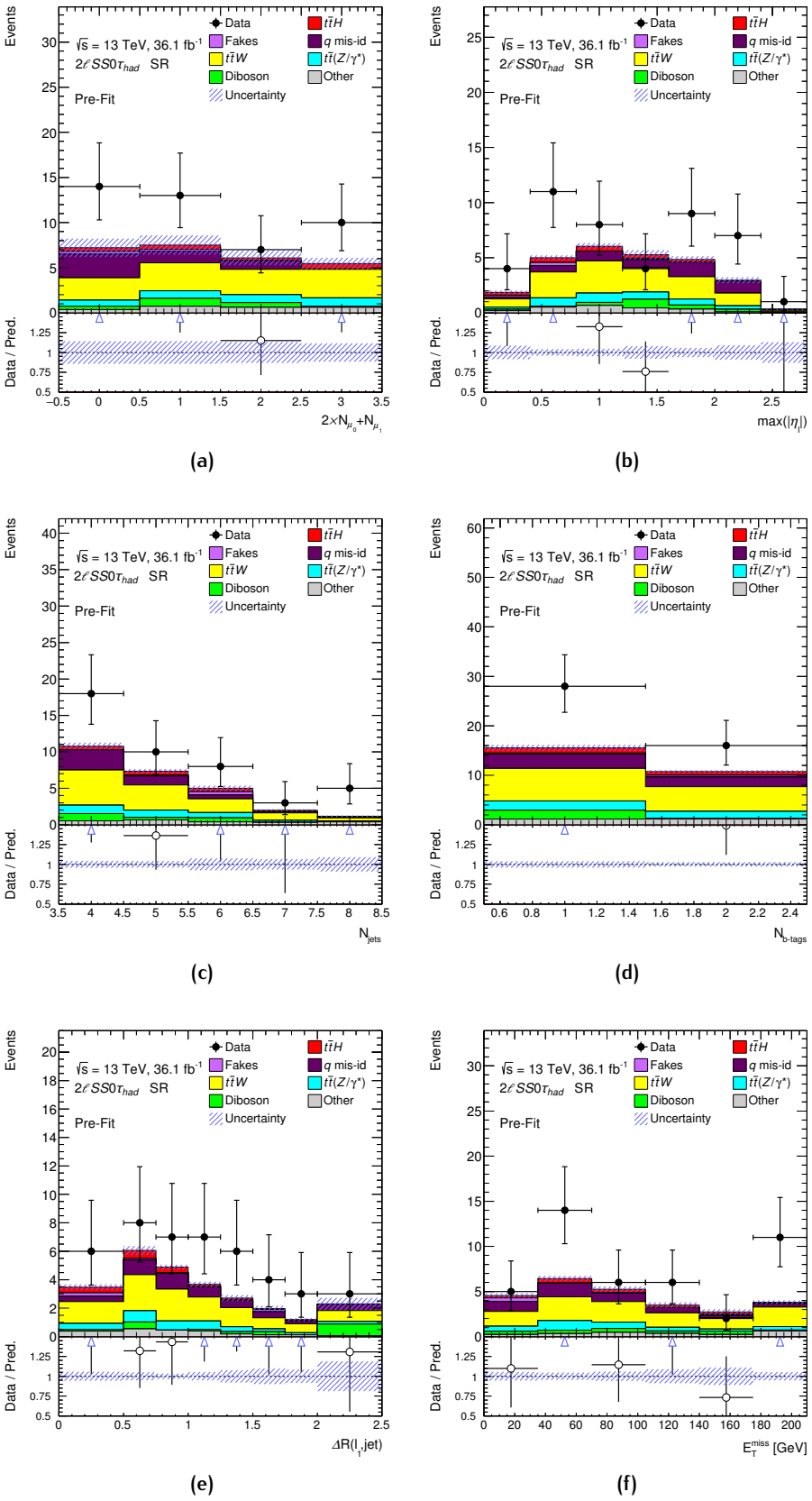
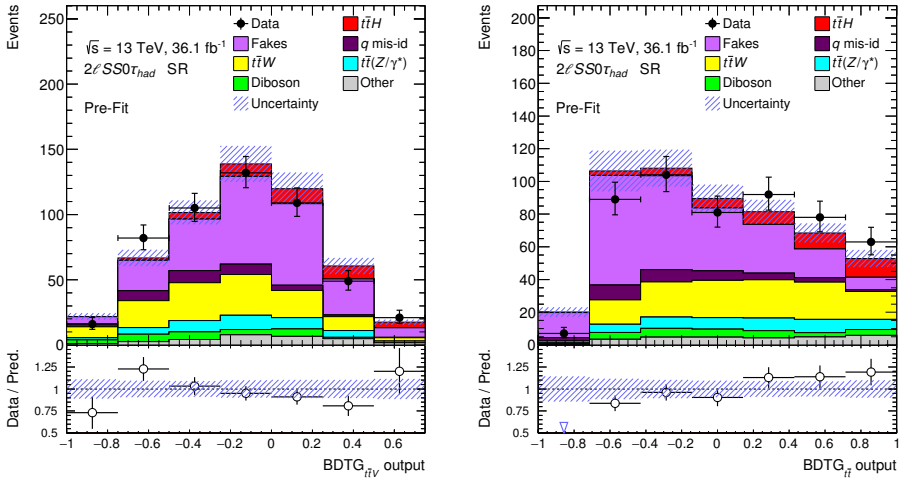


Figure 74: Pre-fit distributions of several input variables to the BDTG algorithm in the 2ℓ SS $0\tau_{had}$ signal region for events with $p_T(\ell_1) > 90$ GeV.



(a) $BDTG_{t\bar{t}V}$ **(b)** $BDTG_{t\bar{t}}$

Figure 75: Pre-fit distributions of the $BDTG_{t\bar{t}V}$ and $BDTG_{t\bar{t}}$ output in the 2ℓ SS $0\tau_{had}$ signal region. Events with fake leptons and charge flip electrons are estimated in a data-driven way, whereas irreducible background are simulated. Errors include the statistical uncertainty and all systematic uncertainties described in Table 27.

verse momentum of (mostly) electron fakes. The impact on the expected sensitivity of the search of removing such events from the fitted distribution was checked. This eventually leads to almost no change in sensitivity, which gets reduced by only about 2%. Hence, the effect of this mis-modelling is considered marginal.

5.3 RESULTS FOR THE 2ℓ SS $0\tau_{had}$ CHANNEL

This results presented in this section are relative to the fit of Eq. (44) performed in 2ℓ SS $0\tau_{had}$ category only. Nuisance parameters for systematic uncertainties leading to a variation in the yield of less than 1% in all bins are discarded to improve the speed of the fit. Likewise, the statistical uncertainty NPs are considered for a given bin only if the induced variation in the yield is greater than 1%. Out of the initial 238, only 181 nuisance parameters are eventually considered. Their naming scheme in the fit model is outlined in Table 29.

When appropriate, predicted correlations between uncertainties are accounted for in the model. In particular, as the residual prompt (mainly $t\bar{t}W$ and diboson) background contribution is subtracted from the control regions to extract the fake lepton rates, the associated nuisance parameters are

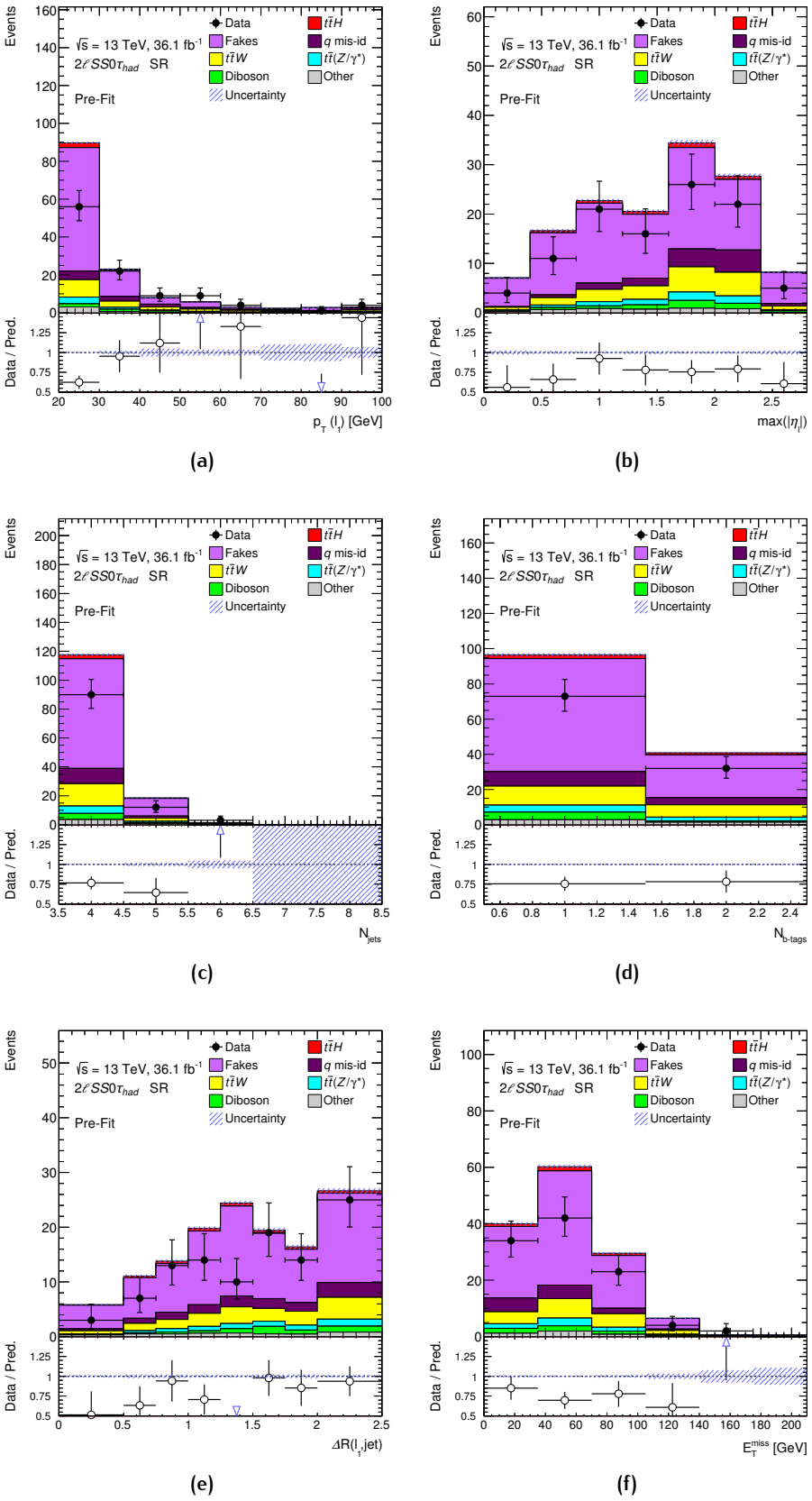
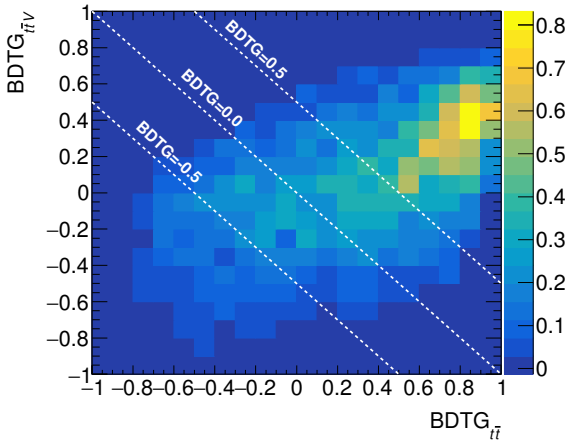
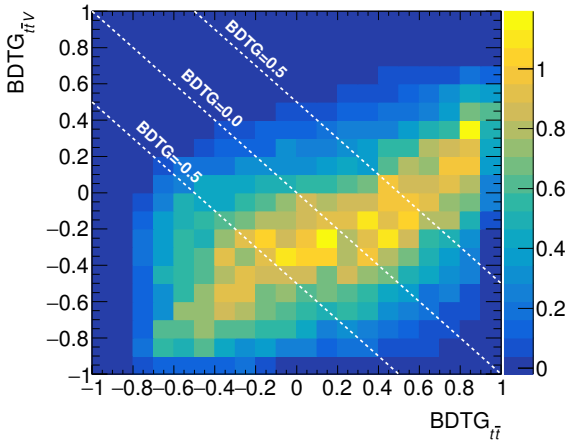


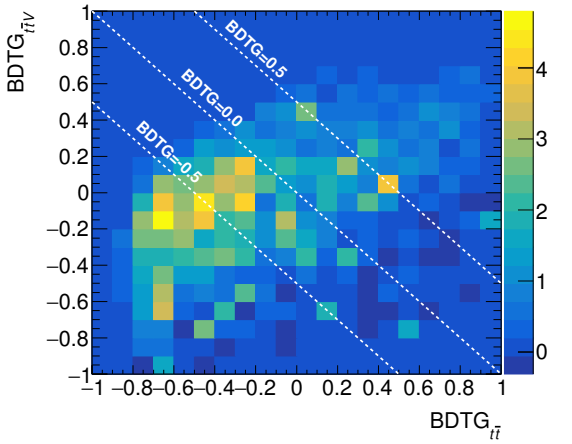
Figure 76: Pre-fit distributions of several input variables to the BDTG algorithm in the 2ℓ SS $0\tau_{had}$ signal region, for $-1 < BDTG_{t\bar{t}} < -0.4$.



(a) $t\bar{t}H$ events



(b) $t\bar{t}V$ events



(c) Fake events

Figure 77: Distribution of 77a: $t\bar{t}H$ signal events, 77b: $t\bar{t}V$ events, 77c: fake background events in the ($BDTG_{t\bar{t}}$,$BDTG_{t\bar{t}V}$) plane. The white lines correspond to three example values (-0.5,0,0.5) of the combined BDTG output defined in Eq. (52), in the case $a = 1$.

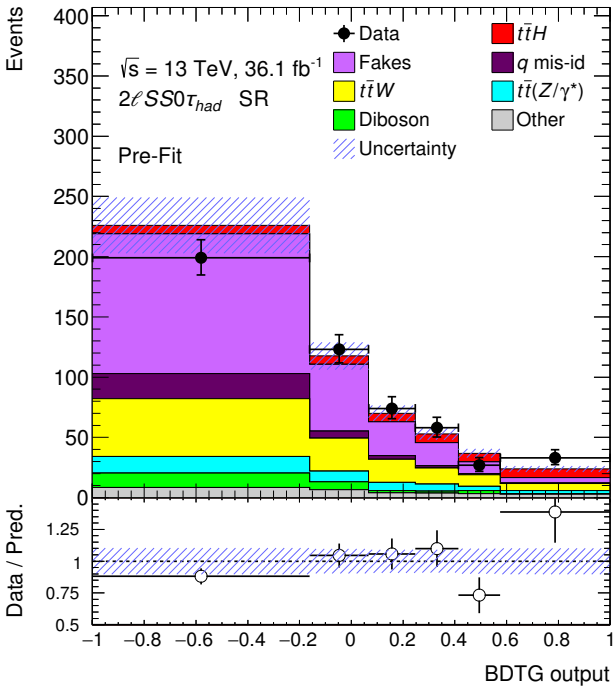


Figure 78: Pre-fit distribution of the combination of the $BDTG_{t\bar{t}V}$ and $BDTG_{t\bar{t}}$ outputs in the 2ℓ SS $0\tau_{had}$ signal region, defined as in Eq. (53). Events with fake leptons and charge flip electrons are estimated in a data-driven way, whereas irreducible background are simulated. Errors include the statistical uncertainty and all systematic uncertainties described in Table 27.

Statistical/systematic uncertainty	Components	NP names
Experimental		
Luminosity	1	ATLAS_lumi
Pile-up modelling	1	ATLAS_PRW_DATASF
e (ε's, scale, resolution)	5	ATLAS_EL_SF_ID, ATLAS_EL_SF_ISO, ATLAS_EL_SF_RECO, ATLAS_EM_RES, ATLAS_EM_SCALE
μ (ε's, scale, resolution)	7	ATLAS_MU_SF_ID_Stat, ATLAS_MU_SF_ID_Syst, ATLAS_MU_SF_ISO_Syst, ATLAS_MU_SF_TTVA_Stat, ATLAS_MU_SF_TTVA_Syst ATLAS_MU_ID, ATLAS_MU_MS
JVT	1	ATLAS_JVT
JES, JER	21	ATLAS_JER, ATLAS_JES_BJES, ATLAS_JES_EffectiveNP_*, ATLAS_JES_EtaInter_Model, ATLAS_JES_EtaInter_NonClosure, ATLAS_JES_EtaInter_Stat, ATLAS_JES_Flavor_Comp_l2otau_VV, ATLAS_JES_Flavor_Comp_l2otau_ttZ, ATLAS_JES_Flavor_Comp_l2otau_Other, ATLAS_JES_Flavor_Resp, ATLAS_JES_PU_OffsetMu, ATLAS_JES_PU_OffsetNPV, ATLAS_JES_PU_PtTerm, ATLAS_JES_PU_Rho
b-tagging	87	ATLAS_FTAG_B*, ATLAS_FTAG_C*, ATLAS_FTAG_L*
Reducible backgrounds		
SR and sidebands statistics	6	Fakes_CR_Stat_l2otau_bin*
Background sub.	4	Fakes_MM_SubtrOther, ttW_XS_QCDscale (*), Diboson_XS (*), QMisId_l2otau (*)
Fake light lepton efficiencies	6	Fakes_Mu_Stat_13, Fakes_Mu_Stat_14, Fakes_Mu_Stat_17, Fakes_Mu_Stat_18, Fakes_El_Stat_13, Fakes_El_Stat_14
Matrix method non-closure	3	Fakes_l2otau_MM_Closure_ee, Fakes_l2otau_MM_Closure_em, Fakes_l2otau_MM_Closure_mm
α conversion fraction for conversions	1	Fakes_l2otau_MM_FracConv_ee
Electron charge mis-assignment	1	QMisId_l2otau (**)
Signal and irreducible backgrounds		
SR background statistics	6	gamma_stat_l2otau_bin_*
Cross section (QCD, PDF+α_s)	13	ttH_XS_PDFunc, ttH_XS_QCDscale, ttW_XS_PDFunc, ttW_XS_QCDscale (**), ttZ_XS_PDFunc, ttZ_XS_QCDscale, Diboson_XS (**), VVV_XS, WtZ_XS, rareTop_XS, tZ_XS, threeTop_XS, tttt_XS
Higgs branching ratios	4	ATLAS_BR_WW, ATLAS_BR_ZZ, ATLAS_BR_tautau, ATLAS_BR_Others
QCD scales	8	ttH_varR, ttH_varRF, ttW_varF, ttW_varRF, ttZ_varF, ttZ_varR, ttZ_varRF, Diboson_varSherpa
Generator	2	ttW_Gen, ttZ_Gen
Parton shower model	1	ttH_PS
Shower tune	3	ttH_varA14, ttW_varA14, ttZ_varA14

Table 29: Names of the NPs associated to statistical/systematic uncertainties effectively considered in the fit for the 2ℓ SS $0\tau_{had}$ channel. The NPs indicated with the (*) in the "Reducible backgrounds" set are associated to the backgrounds subtracted in the Fake control regions of the matrix method (see subsection 4.8.2). Their systematic uncertainties are implemented in the model as anti-correlated to the ones of the NPs affecting the same backgrounds in the SR (indicated with (**) in the table).

taken as fully correlated with the theoretical cross-section systematic uncertainties in the SR. The same treatment is used for the uncertainty associated to the measurement of the charge flip background, which is also subtracted from the electron control region.

Figure 79 shows the post-fit BDTG distribution. The event yield with the uncertainties is outlined in Table 30, in comparison with the pre-fit values. The result of the fit for $\mu_{t\bar{t}H}$ as well as the expected and observed significance of the search are highlighted in Table 31.

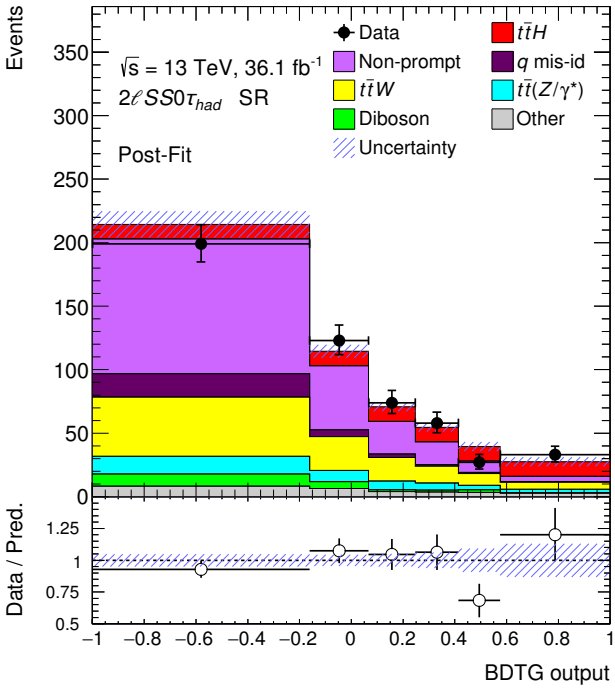


Figure 79: Post-fit distribution of the combination of the $BDTG_{t\bar{t}V}$ and $BDTG_{t\bar{t}}$ outputs in the 2ℓ SS $0\tau_{had}$ signal region, defined as their arithmetic mean. Events with fake leptons and charge flip electrons are estimated in a data-driven way, whereas irreducible background are simulated. Errors include the statistical uncertainty and the systematic uncertainties described in Table 29.

A best-fit of $\mu_{t\bar{t}H} = 1.5^{+0.7}_{-0.6} = 1.5^{+0.4}_{-0.4}$ (stat.) $^{+0.5}_{-0.4}$ (syst.) is found. Statistical and systematic components contribute with approximately equal size to the total uncertainty. The expected significance of the search (see **??**) against the background-only hypothesis - under the hypothesis of the presence of the SM $t\bar{t}H$ process - is found to be 1.9σ; the significance extracted from the observed data is quantified as 2.7σ instead. The significance of the observed data under the assumption of a SM Higgs boson (ie., under the $\mu = 1$ hypothesis) is quantified as 0.9σ, indicating a very low degree of incompatibility with respect to the SM prediction for $t\bar{t}H$ and backgrounds.

	Yield (pre-fit)	Yield (post-fit)
Fakes	233 ± 39	211 ± 26
Charge flips	33 ± 11	28.3 ± 9.4
$t\bar{t}W$	123 ± 18	127 ± 18
$t\bar{t}Z$	41.4 ± 5.6	42.9 ± 5.4
Diboson	25 ± 15	20.0 ± 6.3
Others	28.4 ± 5.9	28.5 ± 5.7
Tot. Background	484 ± 38	459 ± 24
$t\bar{t}H$	42.6 ± 4.2	67 ± 18
Observed	514	

Table 30: Observed and expected event yield for data and background processes in the 2ℓ SS $0\tau_{had}$ signal region before and after the fit. Uncertainties are statistical plus systematic.

Best-fit $\mu_{t\bar{t}H}$	
$\mu = 1.5\ ^{+0.4}_{-0.4}$ (stat.) $^{+0.5}_{-0.4}$ (syst.) $= 1.5\ ^{+0.7}_{-0.6}$	
Significance	
Observed	2.7σ
Expected	1.9σ

Table 31: Best-fit value of $\mu_{t\bar{t}H}$, and expected/observed sensitivity of the $t\bar{t}H \rightarrow$ multi $-$ leptons search in the 2ℓ SS $0\tau_{had}$ channel.

The total statistical-plus-systematic $\pm 1\sigma$ variation of μ is obtained by profiling the likelihood, i.e., minimising with respect to all the parameters but the PoI, and describing the dependence of the minimum as a function of μ: $-2\ln\mathcal{L}_\mu$. By re-defining the profile likelihood to be zero at the minimum:

$$\Delta\left(-\ln\mathcal{L}_\mu\right) = -2\left(\ln\mathcal{L}_\mu - \ln\mathcal{L}_{\mu_{min}}\right), \tag{55}$$

the uncertainty on μ is found by determining the values at which $\Delta\left(-\ln\mathcal{L}_\mu\right) = 1/2$. The profile likelihood as a function of $\mu_{t\bar{t}H}$ is displayed in Figure 80. Its smooth shape and the absence of double minima indicates no issues happen in the minimisation. The systematic error on $\mu_{t\bar{t}H}$ is found by subtracting in quadrature from the total uncertainty the statistical uncertainty, determined by fixing all the NPs in the likelihood to their best-fit values.

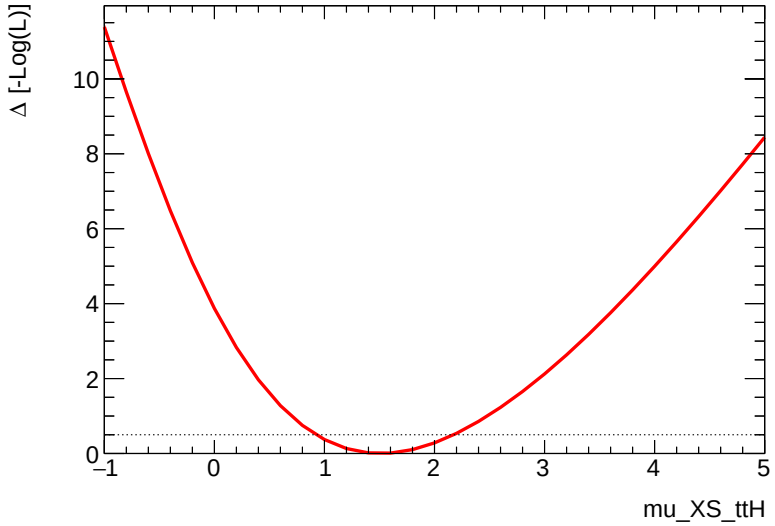


Figure 80: Profile $\Delta\left(-\ln\mathcal{L}_\mu\right)$ as a function of $\mu_{t\bar{t}H}$. The dashed line at $\Delta\left(-\ln\mathcal{L}_\mu\right) = 1/2$ determines the $\pm 1\sigma$ variations of $\mu_{t\bar{t}H}$.

5.3.1 Nuisance parameters pulls and impact on the signal strength uncertainty

As previously stated, the fitted nuisance parameters for systematic uncertainties are expected to be distributed according to a Gaussian model (or log-normal). If that is the case - given the auxiliary measurements for a nuisance parameter and its uncertainty $(\theta_0, \sigma_{\theta_0})$ and their MLE estimators $(\hat{\theta}, \hat{\sigma}_\theta)$ - the *pull*:

$$\Delta\theta = \frac{\hat{\theta} - \theta_0}{\sigma_{\theta_0}} \tag{56}$$

is expected to be Gaussian-distributed with zero mean value and a standard deviation $\sigma_{\Delta\theta} = 1$. If a non-zero pull for a NP is found, it indicates the fit is correcting for a biased initial prediction of that particular parameter. Moreover, the fit can constrain the estimated uncertainty associated to the NP itself if it has more statistical sensitivity to that NP than the auxiliary measurement used to determine its prior uncertainty.

In section 4.9, the impact of the systematic uncertainties on the expected yield for signal and background was discussed and summarised in Table 27. However, that does not represent the most relevant quantity to assess the effect of the systematic uncertainties on the analysis result. Rather, it is the impact of the systematic uncertainties on the fitted signal strength $\mu_{t\bar{t}H}$ that must be evaluated. The procedure to rank the uncertainties upon the size of their effect on the uncertainty on μ is performed according to the following steps, repeated for each NP:

- Perform an unconditional fit to the observed data to measure the nominal value of the signal strength.
- Find the $\pm 1\hat{\sigma}_\theta$ contour for the nuisance parameter θ based on the measured profile likelihood for that parameter.
- Fix the value of θ at the upper and lower $\pm 1\hat{\sigma}_\theta$ limits.
- Repeat the unconditional fit allowing the other nuisance parameters and μ to vary, in order to get an upper and lower estimate of the signal strength $\hat{\mu}^{\pm}$, corresponding to the upper and lower bounds on the systematic contour.
- Estimate the systematic impact on $\hat{\mu}$ of the NP as $\pm\Delta\mu = \pm\hat{\mu} \mp \hat{\mu}^{\pm}$.

The 30 NPs with the largest ranking according to the magnitude of the post-fit impact on $\mu_{t\bar{t}H}$ are shown in Figure 81, from top-ranked to bottom-ranked. The plot also shows the pull for each of those NPs, and its uncertainty.

It can be observed that no nuisance parameters are significantly pulled by the fit, indicating that the pre-fit values and uncertainties are well in agreement with the post-fit results. This is also the case for the other NPs not represented in the plot, including the b-tagging-related ones. Only minor constraints on the uncertainties are found, of $\mathcal{O}(15\%)$ maximum. The two

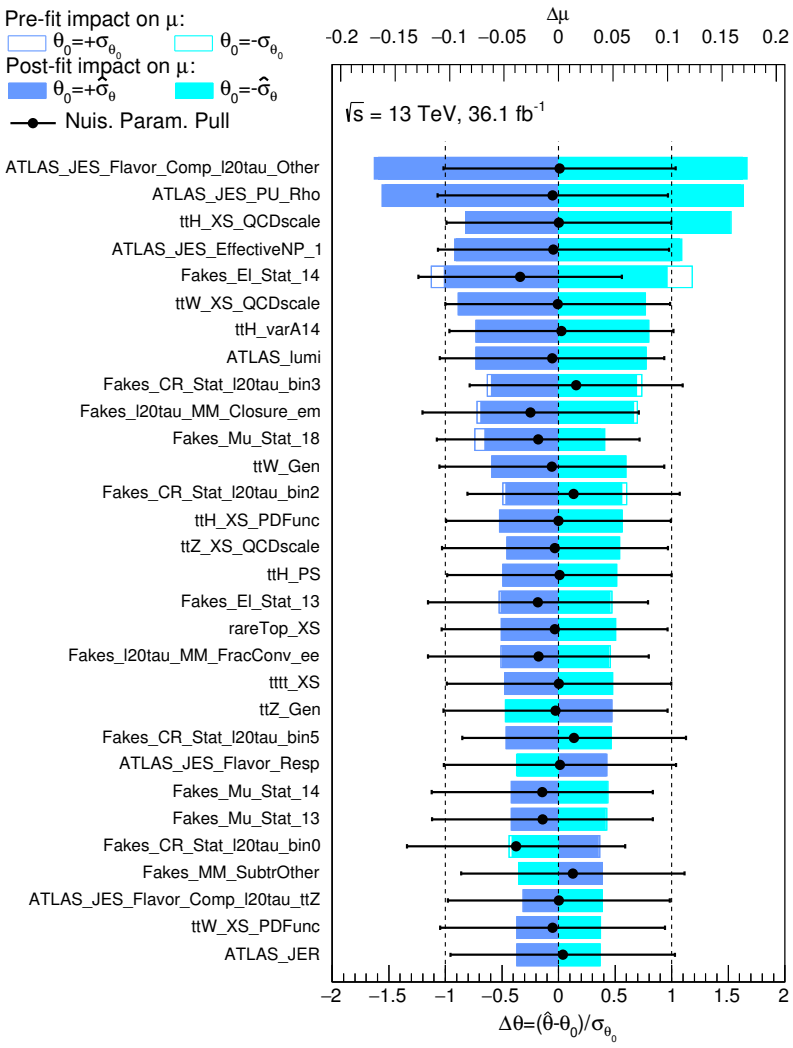


Figure 81: Summary of the impact of systematic uncertainties on $\mu_{t\bar{t}H}$ determined from an unconditional fit to the observed data. Systematic uncertainties are ranked from top to bottom on the left axis according to the post-fit impact shown in solid blue/cyan, and are named according to 29. The pre-fit impact on μ is indicated by the empty blue/cyan boxes. The upper axis represents the scale of $\Delta\mu$. The plot also includes the values of the NP pull (shown as black dots) and of its uncertainty (black lines) determined from the post-fit likelihood contours. The nominal pre-fit bound on each nuisance parameter is represented by the area between the vertical dashed lines. The scale of the pull magnitude is shown in the lower axis.

top-ranked NPs are relative to the jet energy scale uncertainty arising from the jet flavour composition of the combined $t\bar{t}H/t\bar{t}W/$other processes, and the pile-up subtraction. The third highest-ranked NP is the uncertainty on the $t\bar{t}H$ cross section due to QCD scale variations. Fake-related NPs are also found high in the ranking. Overall, the observed ranking reflects our expectation - taking into account the large jet multiplicity characterising signal events and the dominant contribution of the reducible background.

In absence of bias in the fit model, it is expected that the fit result for one NP does not impact in an unexpected way the result for other NPs and - most importantly - for $\mu_{t\bar{t}H}$. This is checked by looking at the correlation coefficients $\rho_{ij} = Cov(\theta_i, \theta_j)/\sigma_{\theta_i}\sigma_{\theta_j}$ between NPs after the fit, and between NPs and $\mu_{t\bar{t}H}$. The correlation matrix is displayed in Figure 82, only for parameters presenting at least one $\geq 10\%$ correlation with another parameter. In general, the fit does not introduce any large or unexpected correlation between NPs. An anti-correlation of around 25% is seen between $\mu_{t\bar{t}H}$ and the JES-related NPs, indicating upward shifts in the jet energy scale lead to a reduction of the signal strength.

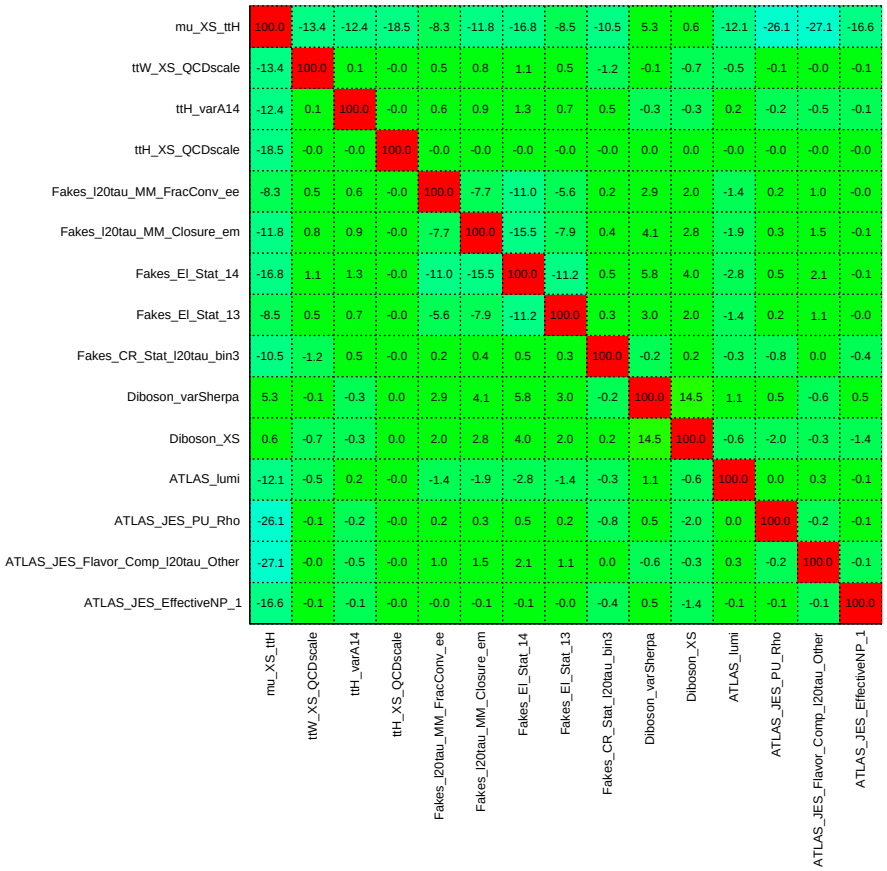


Figure 82: Post-fit correlation matrix for the nuisance parameters and the signal strength. The naming convention of the parameters refers to 29.

5.4 RESULT OF THE COMBINATION OF THE $t\bar{t}H$ MULTI-LEPTONIC CHANNELS

In addition to 2ℓ SS $0\tau_{had}$, seven other independent signal regions of the Run 2 $t\bar{t}H \rightarrow$ multi $-$ leptons analysis are defined according to section 4.2, using lepton and τ_{had} counting to ensure orthogonality. Figure 83 shows the background composition of the eight signal regions and four 3ℓ $0\tau_{had}$ control regions of the $t\bar{t}H$ to multi-leptons analysis, expressed as a fraction of the total background. This indicates that the channels not only differ in terms of signal purity and sensitivity to Higgs decay modes (see Table 6), but also largely differ in terms of background composition. Therefore, a different analysis strategy has been devised for each channel. Their description is beyond the scope of this thesis; details can be found in [125]. The aforementioned document also includes the result of the analysis of the 2ℓ SS $0\tau_{had}$ channel here thoroughly described.

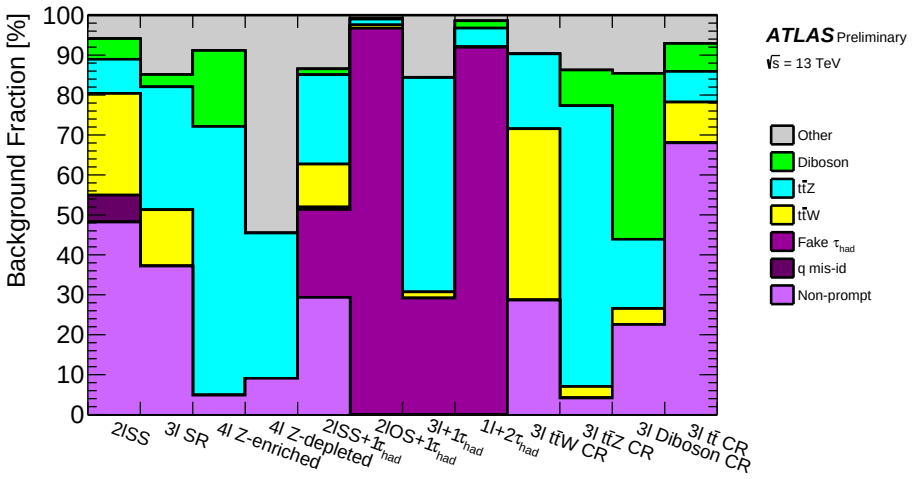


Figure 83: Background composition of the eight signal regions and four 3ℓ $0\tau_{had}$ control regions of the $t\bar{t}H \rightarrow$ multi $-$ leptons analysis, expressed as a fraction of the total background [125].

Similarly to the 2ℓ SS $0\tau_{had}$ channel, most of the other analyses exploit multi-variate analysis techniques to construct discriminants to separate $t\bar{t}H$ from the major backgrounds. In the 3ℓ $0\tau_{had}$ channel, additional four control regions for $t\bar{t}W$, $t\bar{t}Z$, diboson and fakes background are included in the fit model, in order to provide better constraints on the backgrounds normalisation.

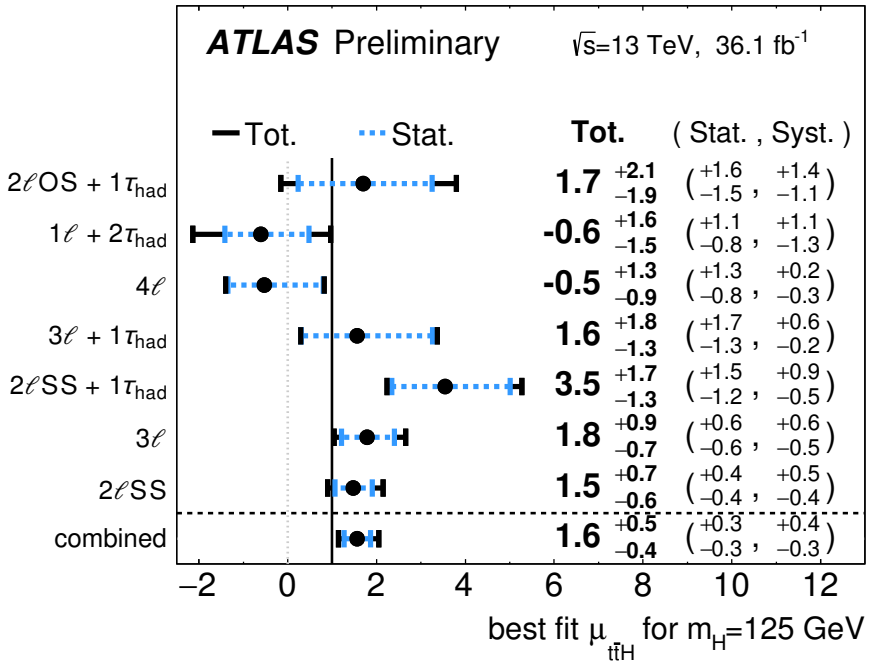
All categories share the same fit model described previously for 2ℓ SS $0\tau_{had}$. The ML fit for $\mu_{t\bar{t}H}$ is performed in each region individually at first. The

combined fit takes into account the product of Poissonian likelihood terms (Eq. (40)) for each region with $\mu_{t\bar{t}H}$ as unique PoI, as well as the product of the probability density functions to include NP constraints from auxiliary measurements. Experimental and theoretical systematic uncertainties are treated as fully correlated across the regions, with the exception of those related to the quark/gluon jet composition. Uncertainties on reducible background modelling are correlated only among regions that share the same methodology for the background estimation. For instance, this the case of the 2ℓ SS $0\tau_{had}$ and 3ℓ $0\tau_{had}$ categories, which are characterised by the same type of fake lepton backgrounds, except for charge flips being negligible in the 3ℓ $0\tau_{had}$ case. Figure 84 shows the best-fit value of $\mu_{t\bar{t}H}$ for each channel and for the combination, and the significance under the SM background-only hypothesis and the SM signal-plus-background hypothesis as well. Figure 85 illustrates the observed and expected yield in each region after the fit.

The observed MLE of μ is $\mu_{t\bar{t}H} = 1.6^{+0.3}_{-0.3}$ (stat.) $^{+0.4}_{-0.3}$ (syst.) $= 1.6\ ^{+0.5}_{-0.4}$. The two most sensitive channels - 2ℓ SS $0\tau_{had}$ and 3ℓ $0\tau_{had}$ - have uncertainties with roughly equal statistical and systematic components, whereas for most the other channels the sensitivity is limited by the large size of the statistical uncertainty. Exceptions are represented by the 1ℓ $2\tau_{had}$ OS and 2ℓ OS $1\tau_{had}$ analyses, which are characterised by large systematic uncertainties originating from the fake τ_{had} estimate. The majority of individual best-fit values of $\mu_{t\bar{t}H}$ are found to be compatible with the combined best-fit result. The 2ℓ SS $1\tau_{had}$ channel features a slightly larger value of $\mu_{t\bar{t}H}$, dominated by the statistical uncertainty. Some tension is also found for the 4ℓ and 1ℓ $2\tau_{had}$ OS channels, likely due to a downward statistical fluctuation of the observed event yield.

After extrapolating to the inclusive phase space under the SM assumption for $\sigma_{t\bar{t}H}$, the $t\bar{t}H$ production cross section is measured to be $\sigma_{t\bar{t}H} = 790^{+150}_{-150}(\text{stat.})^{+170}_{-150}(\text{syst.})$ fb $= 790^{+230}_{-210}$ fb, close to the SM prediction of $\sigma^{SM}_{t\bar{t}H} = 507^{+35}_{-50}$ fb. In the measured value of $\sigma_{t\bar{t}H}$, the uncertainty does not include the systematic uncertainty on the SM $t\bar{t}H$ predicted cross section. Under the hypothesis of absence of signal, an excess of events is observed, which corresponds to an observed (expected) significance of 4.1 (2.8) standard deviations. The significance of the excess is quantified also with respect to the hypothesis of the presence of $t\bar{t}H$ according to the SM, leading to a significance of 1.4σ (see Figure 84). Thus, no significant deviation of the result from the SM prediction is found.

The impact of the most important groups of systematic uncertainties on the combined measurement of μ is shown in Table 32. Again, the uncertain-



Channel	Significance ($\mu = 0$) Observed	Significance ($\mu = 0$) Expected	Significance ($\mu = 1$) Observed
2ℓ SS $0\tau_{had}$	2.7σ	1.9σ	0.9σ
3ℓ $0\tau_{had}$	2.4σ	1.5σ	1.0σ
3ℓ $1\tau_{had}$	1.3σ	0.9σ	0.4σ
2ℓ OS $1\tau_{had}$	0.9σ	0.5σ	0.4σ
1ℓ $2\tau_{had}$ OS	–	0.6σ	–
2ℓ SS $1\tau_{had}$	3.4σ	1.1σ	2.2σ
4ℓ	–	0.8σ	–
Combined	4.1σ	2.8σ	1.4σ

Figure 84: Best-fit values of the signal strength $\mu_{t\bar{t}H}$, and associated expected and observed significance under either the SM background-only hypothesis ($\mu = 0$) or the SM signal-plus-background one ($\mu = 1$), for all the $t\bar{t}H \rightarrow$ multi$-$leptons categories and their combination [125]. The observed significance is indicated with a $-$ for the channels where $\mu_{t\bar{t}H}$ is negative. The plot at the top also includes a graphical representation of the best-fit values of $\mu_{t\bar{t}H}$ with their uncertainties (total and statistical-only component). The vertical black line at $\mu_{t\bar{t}H} = 1$ indicates the SM expectation. In the plot, 2ℓ *SS* is equivalent to the 2ℓ SS $0\tau_{had}$ category described in this thesis.

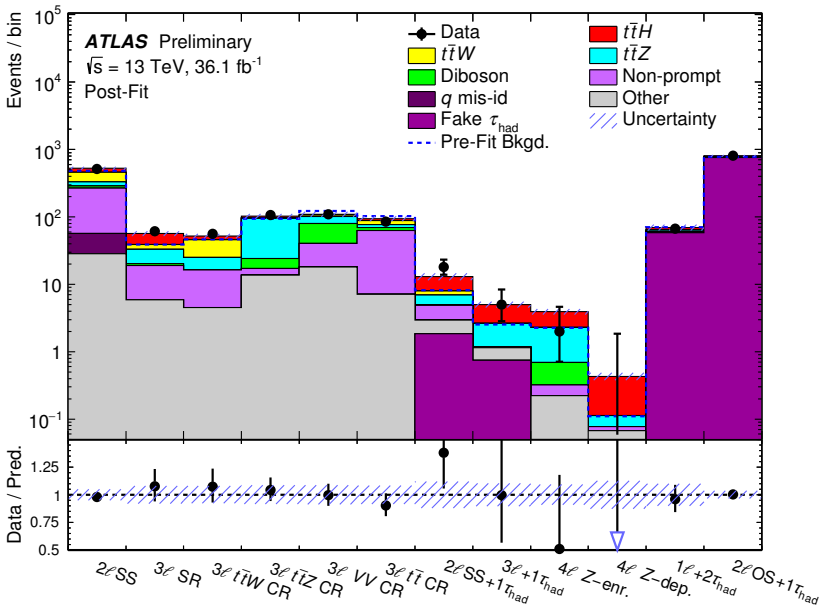


Figure 85: Observed events and post-fit expected yield in the twelve regions of the $t\bar{t}H \rightarrow \text{multi}-\text{leptons}$ analysis [125]. The $t\bar{t}H$ signal yield is scaled by the best-fit value of $\mu_{t\bar{t}H} = 1.6\ ^{+0.5}_{-0.4}$. The hatched band in the lower panel indicates the total uncertainty on the signal-plus-background model as a result of the fit.

ties with the largest impact are those associated with the signal modelling, the jet energy scale and the fake lepton estimate.

Our knowledge of the main irreducible backgrounds $t\bar{t}W$ and $t\bar{t}Z$ is currently limited by the available precision of the measurement of their production cross section, which is of $\mathcal{O}(30\%)$ [148, 149]. Despite the simulation of such processes shows no sign of mis-modelling when checked in dedicated validation regions (see section 4.7), the purity achieved in those regions is not particularly high, especially in the case of the $t\bar{t}W$ background. In order to ensure the result on $\mu_{t\bar{t}H}$ is not dependent on a biased prediction of these processes, an alternative fit setup is tested, in which $\mu_{t\bar{t}H}$ is left free to float alongside the $t\bar{t}W$, $t\bar{t}Z$ cross section modifiers $\mu_{t\bar{t}W}$ and $\mu_{t\bar{t}Z}$. The latter are encoded in the fit model as simple normalisation correction factors, therefore no systematic uncertainties on the cross section of $t\bar{t}W$ and $t\bar{t}Z$ is considered in such setup. The three PoIs are considered to be the same for all the channels entering the fit.

The result of the maximum likelihood fit yields a value of $\mu_{t\bar{t}H} = 1.6^{+0.6}_{-0.5}$, very close to the one obtained from the nominal fit, albeit corresponding to a 15% worse expected sensitivity. The fitted values for the $t\bar{t}V$ cross section modifiers are $\mu_{t\bar{t}W} = 0.92 \pm 0.32$ and $\mu_{t\bar{t}Z} = 1.17^{+0.25}_{-0.22}$, both in agreement with unity as expceted from the SM. The size of the post-fit uncertainties on

Uncertainty Source	$\Delta\mu$	
$t\bar{t}H$ modelling (cross section)	$+0.20$	-0.09
Jet energy scale and resolution	$+0.18$	-0.15
Fake light-lepton estimates	$+0.15$	-0.13
Jet flavour tagging and τ_{had} identification	$+0.11$	-0.09
$t\bar{t}W$ modelling	$+0.10$	-0.09
$t\bar{t}Z$ modelling	$+0.08$	-0.07
Other background modelling	$+0.08$	-0.07
Luminosity	$+0.08$	-0.06
$t\bar{t}H$ modelling (acceptance)	$+0.08$	-0.04
Fake τ_{had} estimates	$+0.07$	-0.07
Other experimental uncertainties	$+0.05$	-0.04
Simulation sample size	$+0.04$	-0.04
Charge misassignment	$+0.01$	-0.01
Total systematic uncertainty	$+0.39$	-0.30

Table 32: Summary of the effects of the most important sources of systematic uncertainties on μ [125]. Due to rounding effects and small correlations between the different sources of uncertainty, the total systematic uncertainty is different from the sum in quadrature of the individual sources.

the cross section modifiers shows no improved precision with respect to the theoretical prediction used in the simulation.

CONCLUSIONS

This thesis presented a search of the $t\bar{t}H$ production mode in the multi-leptonic final state based on 36.1 fb^{-1} of data collected in 2015 and 2016 by the ATLAS detector at the LHC at $\sqrt{s} = 13$ TeV. The search is sensitive to the $H \to WW^*$, $H \to ZZ^*$, and $H \to \tau\tau$ decays.

It also described the work carried out by the author to enhance the performance of the ATLAS tracking algorithm to reconstruct charged particle tracks in high density environments such as high transverse momentum jets and τ_{had} , where distances between particles become comparable to the Pixel detector granularity. Despite not directly utilised by the $t\bar{t}H$ to multi-leptons search, it is foreseen to be beneficial for searches of heavy resonances decaying to boosted particles. In the core of high-p_T jets in simulated $Z' \to t\bar{t}$ events, up to 10% more clusters on track at are found with the new setup. The track reconstruction efficiency is also up to 10% (14%) higher than what achieved in Run 1 for light-flavoured (b) jets. The performance of b-tagging algorithms at high jet momentum is enhanced as well, with an increase of b-tagging efficiency of $7 - 13\%$ for given light-flavoured jet rejection factors for jets with p_T greater than 100 GeV. Furthermore, the choice of using the neural network cluster splitting algorithm directly on track candidates leads to 10% less tracks to be considered in the ambiguity solving stage and the track fitting, thus saving CPU time for the entire tracking software.

The $t\bar{t}H$ search in the 2ℓ SS $0\tau_{had}$ final state with two light leptons of same sign electric charge, no hadronic taus and large jet multiplicity has been described. This channel has a sensitivity of $S/\sqrt{B} \sim 2$, the largest among all channels taken into acount in the analysis. The irreducible $t\bar{t}W$, $t\bar{t}Z$ and diboson backgrounds are estimated from simulation, and their modelling in the signal region is validated in dedicated regions similar to the signal region. For $t\bar{t}W$, two validation regions with limited purity of about 40% could be defined. In both cases, a fit to the observed data yields a value of the $t\bar{t}W$ cross section modifier compatible with the SM value within the uncertainties. Good agreement with data is observed in a $t\bar{t}Z$ validation region with 85% purity. For WZ, a large uncertainty on the production cross section is assigned to take into account the low purity of the control region.

The impact of reducible backgrounds is limited by means of dedicated lepton selection algorithms based on multi-variate analysis techniques. The

algorithm for the suppression of non-prompt leptons from heavy-flavoured hadron decays allows reaching a signal acceptance twice as large as the one achieved in the early Run 2 analysis, given the same signal purity. The algorithm for the identification of leptons whose charge is mis-reconstructed leads to a 60% reduction of the fraction of background from charge flips in $e^{\pm}e^{\pm}$ events with respect to the previous analysis.

The residual fake background contamination in the signal region is estimated in a data-driven way via the matrix method. Real and fake lepton efficiencies to pass the tight lepton selection are estimated in control regions using several parametrisations, and are used to determine the final fake event yield by weighting events in sidebands containing anti-tight leptons. Factors are derived from simulation to correct for the uneven fraction of fake electrons from photon conversions between control and signal region. The matrix method is validated through a closure test based on simulation, and by checking the background modelling in data in low jet multiplicity validation regions. The robustness of the method against potential biases from the trigger selection is checked as well. A good understanding of the physics modelling of fake events is achieved within the estimated uncertainties. The total uncertainty on the predicted fake event yield ranges between 22% to 30%; largest uncertainties are found in events with electrons. The major source of systematic uncertainty is represented by the statistical uncertainty on the fake rates, due to the limited event sample size in the control regions used to extract them. Overall, the size of the uncertainty on the fake lepton background is reduced by $30-48\%$ with respect to the early Run 2 analysis.

Two boosted decision trees relying on kinematic features of the event are trained to discriminate $t\bar{t}H$ against $t\bar{t}V$ and fake backgrounds, respectively. A combination of the output of such algorithms is used to define a binned distribution to be considered in the statistical analysis. A maximum likelihood fit to the observed collision data gives a $t\bar{t}H$ signal strength of $\mu_{t\bar{t}H} = \sigma/\sigma^{SM}_{t\bar{t}H} = 1.5^{+0.4}_{-0.4}$ (stat.) $^{+0.5}_{-0.4}$ (syst.) = $1.5^{+0.7}_{-0.6}$. The observed (expected) significance against the background-only hypothesis is quantified as 2.7σ (1.9σ). Further analysis of the fit results indicates the background model is robust and well understood.

The 2ℓ SS $0\tau_{had}$ signal region is eventually combined with seven other independent signal regions - defined by different lepton and τ_{had} multiplicity - and four control regions. The result of a simultaneous fit in all the signal and control regions leads to a best-fit $\mu_{t\bar{t}H} = 1.6^{+0.3}_{-0.3}$ (stat.) $^{+0.4}_{-0.3}$ (syst.) = $1.6^{+0.5}_{-0.4}$. When extrapolated to an inclusive final state, this results into a measured production cross section of $\sigma_{t\bar{t}H} = 790^{+230}_{-210}$ fb, close to the SM prediction of

$\sigma^{SM}_{t\bar{t}H} = 507^{+35}_{-50}$ fb. The observed (expected) significance of 4.1σ (2.8σ) against the background-only hypothesis provides the strongest evidence to date for the $t\bar{t}H$ production. No significant deviation of the observed data from the SM prediction for $t\bar{t}H$ and backgrounds is found. A recent preliminary result of a $t\bar{t}H$ to multi-leptons search by the CMS collaboration [166] features a best-fit value of $\mu_{t\bar{t}H} = 1.5 \pm 0.5$ - consistent with the ATLAS result - and an observed significance of 3.3σ (2.5σ expected).

This result represents a first step towards the ambitious goal of constraining the measurement of the top Yukawa coupling at the LHC in a direct and most model-independent way via the sole $t\bar{t}H$ production mechanism.

BIBLIOGRAPHY

[1] S. L. Glashow. "Partial Symmetries of Weak Interactions".
In: *Nucl. Phys.* 22 (1961), pp. 579–588.
DOI: 10.1016/0029-5582(61)90469-2.

[2] S. Weinberg. "A Model of Leptons".
In: *Phys. Rev. Lett.* 19 (1967), pp. 1264–1266.
DOI: 10.1103/PhysRevLett.19.1264.

[3] A. Salam and J. C. Ward.
"Gauge Theory of Elementary Interactions".
In: *Phys. Rev.* 136 (3B Nov. 1964), B763–B768.
DOI: 10.1103/PhysRev.136.B763.
URL: https://link.aps.org/doi/10.1103/PhysRev.136.B763.

[4] D. J. Gross and F. Wilczek.
"Ultraviolet Behavior of Non-Abelian Gauge Theories".
In: *Phys. Rev. Lett.* 30 (26 June 1973), pp. 1343–1346.
DOI: 10.1103/PhysRevLett.30.1343.
URL: https://link.aps.org/doi/10.1103/PhysRevLett.30.1343.

[5] G. 't Hooft and M. J. G. Veltman.
"Regularization and Renormalization of Gauge Fields".
In: *Nucl. Phys.* B44 (1972), pp. 189–213.
DOI: 10.1016/0550-3213(72)90279-9.

[6] R. N. Cahn and G. Goldhaber.
The Experimental Foundations of Particle Physics. Second Edition.
Cambridge, UK: Cambridge University Press, 2009.

[7] The ALEPH Collaboration et al.
"Precision Electroweak Measurements on the Z Resonance".
In: *arXiv:hep-ex/0509008v3* (2006).
DOI: 10.1016/j.physrep.2005.12.006.

[8] B. Odom et al. "New Measurement of the Electron Magnetic
Moment Using a One-Electron Quantum Cyclotron".
In: *Phys. Rev. Lett.* 97 (3 July 2006), p. 030801.
DOI: 10.1103/PhysRevLett.97.030801.
URL: https://link.aps.org/doi/10.1103/PhysRevLett.97.030801.

[9] Q. R. Ahmad et al. "Measurement of the Rate of $\nu_e + d \rightarrow p + p + e^-$ Interactions Produced by ^8B Solar Neutrinos at the Sudbury Neutrino Observatory". In: *Phys. Rev. Lett.* 87 (7 July 2001), p. 071301. DOI: 10.1103/PhysRevLett.87.071301. URL: https://link.aps.org/doi/10.1103/PhysRevLett.87.071301.

[10] Q. R. Ahmad et al. "Direct Evidence for Neutrino Flavor Transformation from Neutral-Current Interactions in the Sudbury Neutrino Observatory". In: *Phys. Rev. Lett.* 89 (1 June 2002), p. 011301. DOI: 10.1103/PhysRevLett.89.011301. URL: https://link.aps.org/doi/10.1103/PhysRevLett.89.011301.

[11] Y. Fukuda et al. "Evidence for Oscillation of Atmospheric Neutrinos". In: *Phys. Rev. Lett.* 81 (8 Aug. 1998), pp. 1562–1567. DOI: 10.1103/PhysRevLett.81.1562. URL: https://link.aps.org/doi/10.1103/PhysRevLett.81.1562.

[12] L. Bergstrom. "Dark Matter Evidence, Particle Physics Candidates and Detection Methods". In: *Annalen Phys.* 524 (2012), pp. 479–496. DOI: 10.1002/andp.201200116. arXiv: 1205.4882 [astro-ph.HE].

[13] L. Canetti, M. Drewes, and M. Shaposhnikov. "Matter and Antimatter in the Universe". In: *New J. Phys.* 14 (2012), p. 095012. DOI: 10.1088/1367-2630/14/9/095012. arXiv: 1204.4186 [hep-ph].

[14] C. G. Tully. *Elementary Particle Physics in a Nutshell*. Princeton, New Jersey 08540: Princeton University Press, 2011.

[15] M. E. Peskin and D. V. Schroeder. *An Introduction to quantum field theory*. Reading, USA: Addison-Wesley, 1995. URL: http://www.slac.stanford.edu/~mpeskin/QFT.html.

[16] F. J. Ynduráin. "Limits on the mass of the gluon". In: *Physics Letters B* 345.4 (1995), pp. 524–526. ISSN: 0370-2693. DOI: https://doi.org/10.1016/0370-2693(94)01677-5. URL: http://www.sciencedirect.com/science/article/pii/0370269394016775.

[17] R. K. Ellis, W. J. Stirling, and B. R. Webber. *QCD and Collider Physics*. Cambridge Monographs on Particle Physics, Nuclear Physics and Cosmology. Cambridge University Press, 1996. DOI: 10.1017/CBO9780511628788.

[18] L. Evans and P. Bryant. "LHC Machine". In: *JINST* 3 (2008), S08001. DOI: 10.1088/1748-0221/3/08/S08001.

[19] M. Kuhlen.
QCD at HERA: the hadronic final state in deep inelastic scattering.
Springer tracts in modern physics. Berlin: Springer, 1999.
URL: https://cds.cern.ch/record/390284.

[20] G. Aad et al. "Measurement of the inclusive jet cross-section in proton-proton collisions at $\sqrt{s} = 7$ TeV using 4.5 fb^1 of data with the ATLAS detector".
In: *Journal of High Energy Physics* 2015.2 (Feb. 2015), p. 153.
ISSN: 1029-8479. DOI: 10.1007/JHEP02(2015)153.
URL: https://doi.org/10.1007/JHEP02(2015)153.

[21] V. Khachatryan et al.
"Measurement of electroweak production of a W boson and two forward jets in proton-proton collisions at $\sqrt{s} = 8$ TeV".
In: *Journal of High Energy Physics* 2016.11 (Nov. 2016), p. 147.
ISSN: 1029-8479. DOI: 10.1007/JHEP11(2016)147.
URL: https://doi.org/10.1007/JHEP11(2016)147.

[22] V. Khachatryan et al.
"Measurement of electroweak production of two jets in association with a Z boson in proton-proton collisions at $\sqrt{s} = 8$ TeV".
In: *The European Physical Journal C* 75.2 (Feb. 2015), p. 66.
ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-014-3232-5.
URL: https://doi.org/10.1140/epjc/s10052-014-3232-5.

[23] R. D. Ball et al. "Parton distributions with QED corrections".
In: *Nucl. Phys.* B877 (2013), pp. 290–320.
DOI: 10.1016/j.nuclphysb.2013.10.010.
arXiv: 1308.0598 [hep-ph].

[24] C. Patrignani et al. "Particle Data Group - 2016 Review of Particle Physics - Structure Functions". In: *Chin. Phys. C, 40, 100001* (2016).
URL: http://pdg.lbl.gov/2017/reviews/rpp2016-rev-structure-functions.pdf.

[25] J. Goldstone. "Field Theories with Superconductor Solutions".
In: *Nuovo Cim.* 19 (1961), pp. 154–164. DOI: 10.1007/BF02812722.

[26] Y. Nambu and G. Jona-Lasinio. "Dynamical Model of Elementary Particles Based on an Analogy with Superconductivity. II".
In: *Phys. Rev.* 124 (1 Oct. 1961), pp. 246–254.

DOI: 10.1103/PhysRev.124.246.
URL: https://link.aps.org/doi/10.1103/PhysRev.124.246.

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

[28] P. W. Higgs.
"Spontaneous Symmetry Breakdown without Massless Bosons".
In: *Phys. Rev.* 145 (1966), pp. 1156–1163.
DOI: 10.1103/PhysRev.145.1156.

[29] F. Englert and R. Brout.
"Broken Symmetry and the Mass of Gauge Vector Mesons".
In: *Phys. Rev. Lett.* 13 (1964), pp. 321–323.
DOI: 10.1103/PhysRevLett.13.321.

[30] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble.
"Global Conservation Laws and Massless Particles".
In: *Phys. Rev. Lett.* 13 (20 Nov. 1964), pp. 585–587.
DOI: 10.1103/PhysRevLett.13.585.
URL: https://link.aps.org/doi/10.1103/PhysRevLett.13.585.

[31] B. W. Lee, C. Quigg, and H. B. Thacker. "Weak Interactions at Very High-Energies: The Role of the Higgs Boson Mass".
In: *Phys. Rev.* D16 (1977), p. 1519. DOI: 10.1103/PhysRevD.16.1519.

[32] S. Weinberg. "Mass of the Higgs Boson".
In: *Phys. Rev. Lett.* 36 (1976), pp. 294–296.
DOI: 10.1103/PhysRevLett.36.294.

[33] R. Barate et al. "Search for the standard model Higgs boson at LEP".
In: *Phys. Lett.* B565 (2003), pp. 61–75.
DOI: 10.1016/S0370-2693(03)00614-2.
arXiv: hep-ex/0306033 [hep-ex].

[34] The ALEPH, DELPHI, L3, OPAL, SLD Collaborations, the LEP Electroweak Working Group, the SLD Electroweak and Heavy Flavour Groups.
"Precision Electroweak Measurements on the Z Resonance".
In: *Phys. Rept.* 427 (2006), p. 257. eprint: hep-ex/0509008.
URL: http://lepewwg.web.cern.ch/LEPEWWG/.

[35] "Updated Combination of CDF and Do Searches for Standard Model Higgs Boson Production with up to 10.0 fb^{-1} of Data". In: 2012. arXiv: 1207.0449 [hep-ex].
URL: http://lss.fnal.gov/cgi-bin/find_paper.pl?conf-12-318.

[36] G. Aad et al.
"The ATLAS Experiment at the CERN Large Hadron Collider".
In: *JINST* 3 (2008), S08003. DOI: 10.1088/1748-0221/3/08/S08003.

[37] S. Chatrchyan et al. "The CMS Experiment at the CERN LHC".
In: *JINST* 3 (2008), S08004. DOI: 10.1088/1748-0221/3/08/S08004.

[38] G. Aad et al. "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC".
In: *Phys. Lett.* B716 (2012), pp. 1–29.
DOI: 10.1016/j.physletb.2012.08.020. arXiv: 1207.7214 [hep-ex].

[39] S. Chatrchyan et al. "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC".
In: *Phys. Lett.* B716 (2012), pp. 30–61.
DOI: 10.1016/j.physletb.2012.08.021. arXiv: 1207.7235 [hep-ex].

[40] M. Aaboud et al.
"Evidence for the $H \to b\bar{b}$ decay with the ATLAS detector".
In: (2017). arXiv: 1708.03299 [hep-ex].

[41] G. Aad et al. "Evidence for the Higgs-boson Yukawa coupling to tau leptons with the ATLAS detector". In: *JHEP* 04 (2015), p. 117.
DOI: 10.1007/JHEP04(2015)117. arXiv: 1501.04943 [hep-ex].

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

[43] G. Aad et al. "Study of the spin and parity of the Higgs boson in diboson decays with the ATLAS detector". In: *Eur. Phys. J.* C75.10 (2015). [Erratum: Eur. Phys. J.C76,no.3,152(2016)], p. 476.
DOI: 10.1140/epjc/s10052-015-3685-1,10.1140/epjc/s10052-016-3934-y. arXiv: 1506.05669 [hep-ex].

[44] G. Aad et al.
"Measurements of the Higgs boson production and decay rates and constraints on its couplings from a combined ATLAS and CMS

analysis of the LHC pp collision data at $\sqrt{s} = 7$ and 8 TeV".
In: *JHEP* 08 (2016), p. 045. DOI: 10.1007/JHEP08(2016)045.
arXiv: 1606.02266 [hep-ex].

[45] LHC Higgs Cross Section Working Group et al.
"Handbook of LHC Higgs Cross Sections: 3. Higgs Properties".
In: *CERN-2013-004* (2013). DOI: 10.5170/CERN-2013-004.
arXiv: 1307.1347 [hep-ph].

[46] D. de Florian et al. "Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector". In: (2016).
DOI: 10.23731/CYRM-2017-002. arXiv: 1610.07922 [hep-ph].

[47] C. Quigg. "LHC Physics Potential versus Energy". In: (2009).
arXiv: 0908.3660 [hep-ph].

[48] F. Demartin et al. "Higgs production in association with a single top quark at the LHC". In: *Eur. Phys. J.* C75.6 (2015), p. 267.
DOI: 10.1140/epjc/s10052-015-3475-9.
arXiv: 1504.00611 [hep-ph].

[49] J. Pumplin et al. "New generation of parton distributions with uncertainties from global QCD analysis". In: *JHEP* 07 (2002), p. 012.
DOI: 10.1088/1126-6708/2002/07/012.
arXiv: hep-ph/0201195 [hep-ph].

[50] G. Aad et al. "Measurements of the Higgs boson production and decay rates and coupling strengths using pp collision data at $\sqrt{s} = 7$ and 8 TeV in the ATLAS experiment".
In: *Eur. Phys. J.* C76.1 (2016), p. 6.
DOI: 10.1140/epjc/s10052-015-3769-y.
arXiv: 1507.04548 [hep-ex].

[51] V. Khachatryan et al. "Precise determination of the mass of the Higgs boson and tests of compatibility of its couplings with the standard model predictions using proton collisions at 7 and 8 TeV".
In: *Eur. Phys. J.* C75.5 (2015), p. 212.
DOI: 10.1140/epjc/s10052-015-3351-7. arXiv: 1412.8662 [hep-ex].

[52] J. Ellis and T. You. "Updated Global Analysis of Higgs Couplings".
In: *JHEP* 06 (2013), p. 103. DOI: 10.1007/JHEP06(2013)103.
arXiv: 1303.3879 [hep-ph].

[53] "First combination of Tevatron and LHC measurements of the top-quark mass". In: (2014). arXiv: 1403.4427 [hep-ex].

[54] F. Bezrukov and M. Shaposhnikov.
"Why should we care about the top quark Yukawa coupling?"
In: *J. Exp. Theor. Phys.* 120 (2015). [Zh. Eksp. Teor. Fiz.147,389(2015)], pp. 335–343. DOI: 10.1134/S1063776115030152.
arXiv: 1411.1923 [hep-ph].

[55] T. S. Pettersson and P. Lefèvre.
The Large Hadron Collider: conceptual design.
Tech. rep. CERN-AC-95-05-LHC. Oct. 1995.
URL: https://cds.cern.ch/record/291782.

[56] T. Linnecar et al.
Hardware and Initial Beam Commissioning of the LHC RF Systems.
Tech. rep. LHC-PROJECT-Report-1172.
CERN-LHC-PROJECT-Report-1172. Geneva: CERN, Oct. 2008.
URL: https://cds.cern.ch/record/1176380.

[57] A. Augusto Alves Jr. et al. "The LHCb Detector at the LHC".
In: *JINST* 3 (2008), S08005. DOI: 10.1088/1748-0221/3/08/S08005.

[58] K. Aamodt et al. "The ALICE experiment at the CERN LHC".
In: *JINST* 3 (2008), S08002. DOI: 10.1088/1748-0221/3/08/S08002.

[59] ATLAS Data Preparation. *Collision conditions for 2016 data taking*.
2016. URL: https://atlas.web.cern.ch/Atlas/GROUPS/DATAPREPARATION/DataSummary/2016/records.py.

[60] J. M. Campbell, J. W. Huston, and W. J. Stirling. "Hard Interactions of Quarks and Gluons: A Primer for LHC Physics".
In: *Rept. Prog. Phys.* 70 (2007), p. 89.
DOI: 10.1088/0034-4885/70/1/R02.
arXiv: hep-ph/0611148 [hep-ph].

[61] G. Aad et al.
"The ATLAS Inner Detector commissioning and calibration".
In: *Eur. Phys. J.* C70 (2010), pp. 787–821.
DOI: 10.1140/epjc/s10052-010-1366-7.
arXiv: 1004.5293 [physics.ins-det].

[62] E. Abat et al.
"Combined performance tests before installation of the ATLAS Semiconductor and Transition Radiation Tracking Detectors".
In: *JINST* 3 (2008), P08003. DOI: 10.1088/1748-0221/3/08/P08003.

[63] K. Potamianos. "The upgraded Pixel detector and the commissioning of the Inner Detector tracking of the ATLAS experiment for Run-2 at the Large Hadron Collider". In: *PoS* EPS-HEP2015 (2015), p. 261. arXiv: 1608.07850 [physics.ins-det].

[64] M. Capeans et al. *ATLAS Insertable B-Layer Technical Design Report*. Tech. rep. CERN-LHCC-2010-013. ATLAS-TDR-19. Sept. 2010. URL: https://cds.cern.ch/record/1291633.

[65] D. Rousseau. "The ATLAS liquid argon calorimeter : test beam, installation and commissioning". In: *Proceedings, 2007 IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC 2007): Honolulu, Hawaii, October 28-November 3, 2007*. Vol. 1. 2007, pp. 147–152. DOI: 10.1109/NSSMIC.2007.4436305.

[66] T. Davidek and the ATLAS Tilecal system. "ATLAS Tile Calorimeter performance for single particles in beam tests". In: *Journal of Physics: Conference Series* 160.1 (2009), p. 012057. URL: http://stacks.iop.org/1742-6596/160/i=1/a=012057.

[67] G. Aad et al. "Standalone vertex finding in the ATLAS muon spectrometer". In: *JINST* 9 (2014), P02001. DOI: 10.1088/1748-0221/9/02/P02001. arXiv: 1311.7070 [physics.ins-det].

[68] M. Aaboud et al. "Luminosity determination in pp collisions at $\sqrt{s}$ = 8 TeV using the ATLAS detector at the LHC". In: *Eur. Phys. J.* C76.12 (2016), p. 653. DOI: 10.1140/epjc/s10052-016-4466-1. arXiv: 1608.03953 [hep-ex].

[69] S van der Meer. *Calibration of the effective beam height in the ISR*. Tech. rep. CERN-ISR-PO-68-31. ISR-PO-68-31. Geneva: CERN, 1968. URL: http://cds.cern.ch/record/296752.

[70] M. Bruschi. *The new ATLAS/LUCID detector*. Tech. rep. ATL-FWD-PROC-2015-002. Geneva: CERN, June 2015. URL: https://cds.cern.ch/record/2025000.

[71] V. Cindro et al. "The ATLAS beam conditions monitor". In: *JINST* 3 (2008), P02004. DOI: 10.1088/1748-0221/3/02/P02004.

[72] ATLAS Luminosity Group. *Luminosity Public Results for Run 2*. 2016. URL: https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2.

[73] ATLAS Collaboration.
"ATLAS event at 13 TeV - 21 May 2015 - Run 265545 Event 5720351".
General Photo. May 2015.
URL: https://cds.cern.ch/record/2017716.

[74] M. Aaboud et al.
"Performance of the ATLAS Trigger System in 2015".
In: *Eur. Phys. J.* C77.5 (2017), p. 317.
DOI: 10.1140/epjc/s10052-017-4852-3.
arXiv: 1611.09661 [hep-ex].

[75] E. Simioni. "The Topological Processor for the future ATLAS Level-1 Trigger: from design to commissioning". In: (2014).
arXiv: 1406.4316 [physics.ins-det].

[76] P. Calafiura et al. "The Athena control framework in production, new developments and lessons learned". In: *Computing in high energy physics and nuclear physics. Proceedings, Conference, CHEP'04, Interlaken, Switzerland, September 27-October 1, 2004*. 2005, pp. 456–458.
URL: http://doc.cern.ch/yellowrep/2005/2005-002/p456.pdf.

[77] ATLAS Collaboration. *The ATLAS Computing Workbook*. 2017.
URL: https://twiki.cern.ch/twiki/bin/view/AtlasComputing/WorkBook.

[78] T. Cornelissen et al. *Concepts, Design and Implementation of the ATLAS New Tracking (NEWT)*.
Tech. rep. ATL-SOFT-PUB-2007-007. ATL-COM-SOFT-2007-002.
Geneva: CERN, Mar. 2007.
URL: https://cds.cern.ch/record/1020106.

[79] T. Cornelissen et al. "The global χ^2 track fitter in ATLAS".
In: *Journal of Physics: Conference Series* 119.3 (2008), p. 032013.
URL: http://stacks.iop.org/1742-6596/119/i=3/a=032013.

[80] G. Piacquadio, K. Prokofiev, and A. Wildauer.
"Primary vertex reconstruction in the ATLAS experiment at LHC".
In: *Journal of Physics: Conference Series* 119.3 (2008), p. 032033.
URL: http://stacks.iop.org/1742-6596/119/i=3/a=032033.

[81] N. C. Benekos et al. "ATLAS muon spectrometer simulation and its validation algorithms".
In: *Journal of Physics: Conference Series* 119.3 (2008), p. 032009.
URL: http://stacks.iop.org/1742-6596/119/i=3/a=032009.

[82] W. Lampl et al.
Calorimeter Clustering Algorithms: Description and Performance.
Tech. rep. ATL-LARG-PUB-2008-002. ATL-COM-LARG-2008-003.
Geneva: CERN, Apr. 2008.
URL: https://cds.cern.ch/record/1099735.

[83] *Measurements of the photon identification efficiency with the ATLAS detector using 4.9 fb^{-1} of pp collision data collected in 2011*.
Tech. rep. ATLAS-CONF-2012-123. Geneva: CERN, 2012.
URL: https://cds.cern.ch/record/1473426.

[84] *Improved electron reconstruction in ATLAS using the Gaussian Sum Filter-based model for bremsstrahlung*.
Tech. rep. ATLAS-CONF-2012-047. Geneva: CERN, May 2012.
URL: http://cds.cern.ch/record/1449796.

[85] G. Aad et al. "Electron and photon energy calibration with the ATLAS detector using LHC Run 1 data".
In: *Eur. Phys. J.* C74.10 (2014), p. 3071.
DOI: 10.1140/epjc/s10052-014-3071-4. arXiv: 1407.5063 [hep-ex].

[86] *Electron efficiency measurements with the ATLAS detector using the 2015 LHC proton-proton collision data*. Tech. rep. ATLAS-CONF-2016-024.
Geneva: CERN, June 2016.
URL: http://cds.cern.ch/record/2157687.

[87] G. Aad et al. "Muon reconstruction performance of the ATLAS detector in proton-proton collision data at $\sqrt{s}$ =13 TeV".
In: *Eur. Phys. J.* C76.5 (2016), p. 292.
DOI: 10.1140/epjc/s10052-016-4120-y.
arXiv: 1603.05598 [hep-ex].

[88] ATLAS Collaboration.
ATLAS Muon Combined Performance with the full 2016 dataset. 2016.
URL:
https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PLOTS/MUON-2017-001/index.html.

[89] ATLAS Collaboration. *Muon trigger public results for Run 2*. 2016.
URL: https://twiki.cern.ch/twiki/bin/view/AtlasPublic/MuonTriggerPublicResults.

[90] G. Aad et al. "Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1". In: (2016).
arXiv: 1603.02934 [hep-ex].

[91] *Jet energy scale and its systematic uncertainty in proton-proton collisions at sqrt(s)=7 TeV with ATLAS 2011 data*.
Tech. rep. ATLAS-CONF-2013-004. Geneva: CERN, Jan. 2013.
URL: https://cds.cern.ch/record/1509552.

[92] M. Cacciari, G. P. Salam, and G. Soyez.
"The Anti-k(t) jet clustering algorithm". In: *JHEP* 04 (2008), p. 063.
DOI: 10.1088/1126-6708/2008/04/063. arXiv: 0802.1189 [hep-ph].

[93] *Monte Carlo Calibration and Combination of In-situ Measurements of Jet Energy Scale, Jet Energy Resolution and Jet Mass in ATLAS*.
Tech. rep. ATLAS-CONF-2015-037. Geneva: CERN, Aug. 2015.
URL: https://cds.cern.ch/record/2044941.

[94] *Pile-up subtraction and suppression for jets in ATLAS*.
Tech. rep. ATLAS-CONF-2013-083. Geneva: CERN, Aug. 2013.
URL: https://cds.cern.ch/record/1570994.

[95] *Tagging and suppression of pileup jets with the ATLAS detector*.
Tech. rep. ATLAS-CONF-2014-018. Geneva: CERN, May 2014.
URL: https://cds.cern.ch/record/1700870.

[96] *Expected performance of the ATLAS b-tagging algorithms in Run-2*.
Tech. rep. ATL-PHYS-PUB-2015-022. Geneva: CERN, July 2015.
URL: https://cds.cern.ch/record/2037697.

[97] *Optimisation of the ATLAS b-tagging performance for the 2016 LHC Run*.
Tech. rep. ATL-PHYS-PUB-2016-012. Geneva: CERN, June 2016.
URL: http://cds.cern.ch/record/2160731.

[98] G. Aad et al.
"Performance of *b*-Jet Identification in the ATLAS Experiment".
In: *JINST* 11.04 (2016), P04008.
DOI: 10.1088/1748-0221/11/04/P04008.
arXiv: 1512.01094 [hep-ex].

[99] *Reconstruction, Energy Calibration, and Identification of Hadronically Decaying Tau Leptons in the ATLAS Experiment for Run-2 of the LHC*.
Tech. rep. ATL-PHYS-PUB-2015-045. Geneva: CERN, Nov. 2015.
URL: https://cds.cern.ch/record/2064383.

[100] *Expected performance of missing transverse momentum reconstruction for the ATLAS detector at* $\sqrt{s} = 13$ *TeV*.
Tech. rep. ATL-PHYS-PUB-2015-023. Geneva: CERN, July 2015.
URL: http://cds.cern.ch/record/2037700.

[101] A. Buckley et al.
"General-purpose event generators for LHC physics".
In: *Phys. Rept.* 504 (2011), pp. 145–233.
DOI: 10.1016/j.physrep.2011.03.005. arXiv: 1101.2599 [hep-ph].

[102] S. Agostinelli et al. "GEANT4: A Simulation toolkit".
In: *Nucl. Instrum. Meth.* A506 (2003), pp. 250–303.
DOI: 10.1016/S0168-9002(03)01368-8.

[103] G. Aad et al. "The ATLAS Simulation Infrastructure".
In: *Eur. Phys. J.* C70 (2010), pp. 823–874.
DOI: 10.1140/epjc/s10052-010-1429-9.
arXiv: 1005.4568 [physics.ins-det].

[104] W. Lukas. *Fast Simulation for ATLAS: Atlfast-II and ISF*.
Tech. rep. ATL-SOFT-PROC-2012-065. Geneva: CERN, June 2012.
URL: https://cds.cern.ch/record/1458503.

[105] G. Aad et al. "A search for high-mass resonances decaying to $\tau^+\tau^-$ in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector".
In: *JHEP* 07 (2015), p. 157. DOI: 10.1007/JHEP07(2015)157.
arXiv: 1502.07177 [hep-ex].

[106] *Search for heavy particles decaying to pairs of highly-boosted top quarks using lepton-plus-jets events in proton–proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*. Tech. rep. ATLAS-CONF-2016-014.
Geneva: CERN, Mar. 2016.
URL: https://cds.cern.ch/record/2141001.

[107] *The Optimization of ATLAS Track Reconstruction in Dense Environments*.
Tech. rep. ATL-PHYS-PUB-2015-006. Geneva: CERN, Mar. 2015.
URL: https://cds.cern.ch/record/2002609.

[108] M. Aaboud et al. "Performance of the ATLAS Track Reconstruction Algorithms in Dense Environments in LHC Run 2".
In: *Eur. Phys. J.* C77.10 (2017), p. 673.
DOI: 10.1140/epjc/s10052-017-5225-7.
arXiv: 1704.07983 [hep-ex].

[109] G. Aad et al. "ATLAS pixel detector electronics and sensors".
In: *JINST* 3 (2008), P07007. DOI: 10.1088/1748-0221/3/07/P07007.

[110] E. Belau et al. "The Charge Collection in Silicon Strip Detectors".
In: *Nucl. Instrum. Meth.* 214 (1983), p. 253.
DOI: 10.1016/0167-5087(83)90591-4.

[111] G. Aad et al. "A neural network clustering algorithm for the ATLAS silicon pixel detector". In: *JINST* 9 (2014), P09009. DOI: 10.1088/1748-0221/9/09/P09009. arXiv: 1406.7690 [hep-ex].

[112] A. Salzburger. "Optimisation of the ATLAS Track Reconstruction Software for Run-2". In: *Journal of Physics: Conference Series* 664.7 (2015), p. 072042. URL: http://stacks.iop.org/1742-6596/664/i=7/a=072042.

[113] G. Aad et al. "Measurement of the muon reconstruction performance of the ATLAS detector using 2011 and 2012 LHC protonproton collision data". In: *Eur. Phys. J.* C74.11 (2014), p. 3130. DOI: 10.1140/epjc/s10052-014-3130-x. arXiv: 1407.3935 [hep-ex].

[114] R. Fruhwirth. "Application of Kalman filtering to track and vertex fitting". In: *Nucl. Instrum. Meth.* A262 (1987), pp. 444–450. DOI: 10.1016/0168-9002(87)90887-4.

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

[116] A. D. Martin et al. "Parton distributions for the LHC". In: *Eur. Phys. J.* C63 (2009), pp. 189–285. DOI: 10.1140/epjc/s10052-009-1072-5. arXiv: 0901.0002 [hep-ph].

[117] *Commissioning of the ATLAS high-performance b-tagging algorithms in the 7 TeV collision data*. Tech. rep. ATLAS-CONF-2011-102. Geneva: CERN, July 2011. URL: https://cds.cern.ch/record/1369219.

[118] G. Aad et al. "Search for the Standard Model Higgs boson decaying into $b\bar{b}$ produced in association with top quarks decaying hadronically in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector". In: *Journal of High Energy Physics* 2016.5 (May 2016), p. 160. ISSN: 1029-8479. DOI: 10.1007/JHEP05(2016)160. URL: https://doi.org/10.1007/JHEP05(2016)160.

[119] *Search for the Standard Model Higgs boson produced in association with top quarks and decaying into $b\bar{b}$ in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*. Tech. rep. ATLAS-CONF-2016-080. Geneva: CERN, Aug. 2016. URL: https://cds.cern.ch/record/2206255.

[120] G. Aad et al. "Search for $H \to \gamma\gamma$ produced in association with top quarks and constraints on the Yukawa coupling between the top quark and the Higgs boson using data taken at 7 TeV and 8 TeV with the ATLAS detector". In: *Phys. Lett.* B740 (2015), pp. 222–242. DOI: 10.1016/j.physletb.2014.11.049. arXiv: 1409.3122 [hep-ex].

[121] G. Aad et al. "Search for the associated production of the Higgs boson with a top quark pair in multilepton final states with the ATLAS detector". In: *Phys. Lett.* B749 (2015), pp. 519–541. DOI: 10.1016/j.physletb.2015.07.079. arXiv: 1506.05988 [hep-ex].

[122] V. Khachatryan et al. "Search for the associated production of the Higgs boson with a top-quark pair". In: *JHEP* 09 (2014). [Erratum: JHEP10,106(2014)], p. 087. DOI: 10.1007/JHEP09(2014)087,10.1007/JHEP10(2014)106. arXiv: 1408.1682 [hep-ex].

[123] W.J. Stirling. *Private communication*. 2016. URL: https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2.

[124] *Search for the Associated Production of a Higgs Boson and a Top Quark Pair in Multilepton Final States with the ATLAS Detector*. Tech. rep. ATLAS-CONF-2016-058. Geneva: CERN, Aug. 2016. URL: https://cds.cern.ch/record/2206153.

[125] *Evidence for the associated production of the Higgs boson and a top quark pair with the ATLAS detector*. Tech. rep. ATLAS-CONF-2017-077. Geneva: CERN, Nov. 2017. URL: https://cds.cern.ch/record/2291405.

[126] *Multi-Boson Simulation for 13 TeV ATLAS Analyses*. Tech. rep. ATL-PHYS-PUB-2016-002. Geneva: CERN, Jan. 2016. URL: https://cds.cern.ch/record/2119986.

[127] *Modelling of the $t\bar{t}H$ and $t\bar{t}V$ $(V = W, Z)$ processes for $\sqrt{s} = 13$ TeV ATLAS analyses*. Tech. rep. ATL-PHYS-PUB-2016-005. Geneva: CERN, Jan. 2016. URL: https://cds.cern.ch/record/2120826.

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

[129] T. Sjostrand, S. Mrenna, and P. Z. Skands.
"A Brief Introduction to PYTHIA 8.1".
In: *Comput. Phys. Commun.* 178 (2008), pp. 852–867.
DOI: 10.1016/j.cpc.2008.01.036. arXiv: 0710.3820 [hep-ph].

[130] *Summary of ATLAS Pythia 8 tunes*.
Tech. rep. ATL-PHYS-PUB-2012-003. Geneva: CERN, Aug. 2012.
URL: https://cds.cern.ch/record/1474107.

[131] P. Golonka and Z. Was. "PHOTOS Monte Carlo: A Precision tool for QED corrections in Z and W decays".
In: *Eur. Phys. J.* C45 (2006), pp. 97–107.
DOI: 10.1140/epjc/s2005-02396-4.
arXiv: hep-ph/0506026 [hep-ph].

[132] *ATLAS Run 1 Pythia8 tunes*. Tech. rep. ATL-PHYS-PUB-2014-021.
Geneva: CERN, Nov. 2014.
URL: https://cds.cern.ch/record/1966419.

[133] R. D. Ball et al. "Parton distributions for the LHC Run II".
In: *JHEP* 04 (2015), p. 040. DOI: 10.1007/JHEP04(2015)040.
arXiv: 1410.8849 [hep-ph].

[134] J. Butterworth et al. "PDF4LHC recommendations for LHC Run II".
In: *J. Phys.* G43 (2016), p. 023001.
DOI: 10.1088/0954-3899/43/2/023001. arXiv: 1510.03865 [hep-ph].

[135] S. Frixione et al. "Electroweak and QCD corrections to top-pair hadroproduction in association with heavy bosons".
In: *JHEP* 06 (2015), p. 184. DOI: 10.1007/JHEP06(2015)184.
arXiv: 1504.03446 [hep-ph].

[136] T. Gleisberg et al. "Event generation with SHERPA 1.1".
In: *JHEP* 02 (2009), p. 007. DOI: 10.1088/1126-6708/2009/02/007.
arXiv: 0811.4622 [hep-ph].

[137] H.-L. Lai et al. "New parton distributions for collider physics".
In: *Phys. Rev.* D82 (2010), p. 074024.
DOI: 10.1103/PhysRevD.82.074024. arXiv: 1007.2241 [hep-ph].

[138] S. Frixione, P. Nason, and G. Ridolfi.
"A Positive-weight next-to-leading-order Monte Carlo for heavy flavour hadroproduction". In: *JHEP* 09 (2007), p. 126.
DOI: 10.1088/1126-6708/2007/09/126. arXiv: 0707.3088 [hep-ph].

[139] D. J. Lange. "The EvtGen particle decay simulation package". In: *Nucl. Instrum. Meth.* A462 (2001), pp. 152–155. DOI: 10.1016/S0168-9002(01)00089-4.

[140] M. Bahr et al. "Herwig++ Physics and Manual". In: *Eur. Phys. J.* C58 (2008), pp. 639–707. DOI: 10.1140/epjc/s10052-008-0798-9. arXiv: 0803.0883 [hep-ph].

[141] R. Ospanov, R. T. Roberts, and T. R. Wyatt. *Tagging non-prompt electrons and muons*. Tech. rep. ATL-COM-PHYS-2016-1444. ATLAS Internal note. Geneva: CERN, Oct. 2016. URL: https://cds.cern.ch/record/2220954.

[142] K. Mochizuki et al. *Electron Charge Flip Tagger*. ATLAS e/gamma workshop 2016. Nov. 2016. URL: https://indico.cern.ch/event/567645/contributions/2349785/.

[143] M. Aaboud et al. "Electron efficiency measurements with the ATLAS detector using 2012 LHC protonproton collision data". In: *Eur. Phys. J.* C77.3 (2017), p. 195. DOI: 10.1140/epjc/s10052-017-4756-2. arXiv: 1612.01456 [hep-ex].

[144] *ATLAS electron, photon and muon isolation in Run 2*. Tech. rep. ATL-COM-PHYS-2017-290. Geneva: CERN, Mar. 2017. URL: https://cds.cern.ch/record/2256658.

[145] J. H. Friedman. "Greedy function approximation: A gradient boosting machine". In: *The Annals of Statistics* 29.5 (Nov. 2001), pp. 1189–1232.

[146] A. Hocker et al. "TMVA - Toolkit for Multivariate Data Analysis". In: *PoS* ACAT (2007), p. 040. arXiv: physics/0703039 [PHYSICS].

[147] Y. Coadou. "Boosted Decision Trees and Applications". In: *EPJ Web Conf.* 55 (2013), p. 02004. DOI: 10.1051/epjconf/20135502004.

[148] G. Aad et al. "Measurement of the $t\bar{t}W$ and $t\bar{t}Z$ production cross sections in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector". In: *JHEP* 11 (2015), p. 172. DOI: 10.1007/JHEP11(2015)172. arXiv: 1509.05276 [hep-ex].

[149] V. Khachatryan. "Observation of top quark pairs produced in association with a vector boson in pp collisions at $\sqrt{s} = 8$ TeV". In: *JHEP* 01 (2016), p. 096. DOI: 10.1007/JHEP01(2016)096. arXiv: 1510.01131 [hep-ex].

[150] G. Aad et al. "Measurement of the top quark-pair production cross section with ATLAS in pp collisions at $\sqrt{s} = 7$ TeV". In: *Eur. Phys. J.* C71 (2011), p. 1577. DOI: 10.1140/epjc/s10052-011-1577-6. arXiv: 1012.1792 [hep-ex].

[151] M. Aaboud et al. "Measurement of the $t\bar{t}Z$ and $t\bar{t}W$ production cross sections in multilepton final states using 3.2 fb^{-1} of pp collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector". In: *Eur. Phys. J.* C77.1 (2017), p. 40. DOI: 10.1140/epjc/s10052-016-4574-y. arXiv: 1609.01599 [hep-ex].

[152] M. Aaboud et al. "Study of the material of the ATLAS inner detector for Run 2 of the LHC". In: (2017). arXiv: 1707.02826 [hep-ex].

[153] M. Aaboud et al. "Measurement of the $t\bar{t}\gamma$ production cross section in proton-proton collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector". In: (2017). arXiv: 1706.03046 [hep-ex].

[154] *Search for the Standard Model Higgs boson produced in association with top quarks and decaying to $b\bar{b}$ in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector at the LHC*. Tech. rep. ATLAS-COM-CONF-2014-004. Internal supporting documentation to ATLAS-CONF-2014-011. Geneva: CERN, Feb. 2014. URL: https://cds.cern.ch/record/1648282.

[155] M. Aaboud et al. "Jet energy scale measurements and their systematic uncertainties in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector". In: *Phys. Rev. D* 96 (7 Oct. 2017), p. 072002. DOI: 10.1103/PhysRevD.96.072002. URL: https://link.aps.org/doi/10.1103/PhysRevD.96.072002.

[156] G. Aad et al. "Performance of pile-up mitigation techniques for jets in pp collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector". In: *Eur. Phys. J.* C76.11 (2016), p. 581. DOI: 10.1140/epjc/s10052-016-4395-z. arXiv: 1510.03823 [hep-ex].

[157] *Jet Calibration and Systematic Uncertainties for Jets Reconstructed in the ATLAS Detector at* $\sqrt{s} = 13$ *TeV*. Tech. rep. ATL-PHYS-PUB-2015-015. Geneva: CERN, July 2015.
URL: https://cds.cern.ch/record/2037613.

[158] M. Pinamonti, L. Valery, et al.
The TRExFitter package for statistical analysis. 2016. URL: https://twiki.cern.ch/twiki/bin/view/AtlasProtected/TtHFitter.

[159] K. Cranmer et al. *HistFactory: A tool for creating statistical models for use with RooFit and RooStats*. Tech. rep. CERN-OPEN-2012-016. New York: New York U., Jan. 2012.
URL: https://cds.cern.ch/record/1456844.

[160] G. Cowan et al.
"Asymptotic formulae for likelihood-based tests of new physics".
In: *Eur. Phys. J.* C71 (2011). [Erratum: Eur. Phys. J.C73,2501(2013)], p. 1554. DOI: 10.1140/epjc/s10052-011-1554-0,10.1140/epjc/s10052-013-2501-z.
arXiv: 1007.1727 [physics.data-an].

[161] R. Barlow and C. Beeston.
"Fitting using finite Monte Carlo samples".
In: *Computer Physics Communications* 77.2 (1993), pp. 219–228.
ISSN: 0010-4655.
DOI: https://doi.org/10.1016/0010-4655(93)90005-W.
URL: http://www.sciencedirect.com/science/article/pii/001046559390005W.

[162] F. James and M. Roos. "Minuit - a system for function minimization and analysis of the parameter errors and correlations".
In: *Computer Physics Communications* 10.6 (1975), pp. 343–367.
ISSN: 0010-4655.
DOI: https://doi.org/10.1016/0010-4655(75)90039-9.
URL: http://www.sciencedirect.com/science/article/pii/0010465575900399.

[163] *Procedure for the LHC Higgs boson search combination in Summer 2011*. Tech. rep. CMS-NOTE-2011-005. ATL-PHYS-PUB-2011-11. Geneva: CERN, Aug. 2011.
URL: https://cds.cern.ch/record/1379837.

[164] S. S. Wilks. "The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses".

In: *Annals Math. Statist.* 9.1 (1938), pp. 60–62.
DOI: 10.1214/aoms/1177732360.

[165] A. Wald. "Tests of Statistical Hypotheses Concerning Several Parameters When the Number of Observations is Large".
In: *Transactions of the American Mathematical Society* 54.3 (1943), pp. 426–482. ISSN: 00029947.
URL: http://www.jstor.org/stable/1990256.

[166] *Search for Higgs boson production in association with top quarks in multilepton final states at* $\sqrt{s} = 13$ TeV.
Tech. rep. CMS-PAS-HIG-17-004. Geneva: CERN, 2017.
URL: https://cds.cern.ch/record/2256103.

[167] G. Aad et al. "Search for the Standard Model Higgs boson produced in association with top quarks and decaying into $b\bar{b}$ in pp collisions at $\sqrt{s}$ = 8 TeV with the ATLAS detector".
In: *Eur. Phys. J.* C75.7 (2015), p. 349.
DOI: 10.1140/epjc/s10052-015-3543-1.
arXiv: 1503.05066 [hep-ex].