A Search for the Higgs boson produced in association with top quarks in multilepton final states at ATLAS

Chris Lester

A DISSERTATION

in

Physics and Astronomy

Presented to the Faculties of The Univeristy of Pennsylvania

in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy

2014

I. Joseph Kroll, Professor, Physics

Supervisor of Dissertation

Elliot Lipeles, Associate Professor, Physics

Graduate Group Chairperson

Dissertation Committee

Justin Khoury, Associate Professor, Physics

I. Joseph Kroll, Professor, Physics

Burt Ovrut, Professor, Physics

Evelyn Thompson, Associate Professor, Physics

I. Joseph Kroll, Professor, Physics

A Search for the Higgs boson produced in association with top quarks in multilepton final states at ATLAS

Acknowledgements

Acknowledgements acknowledgements.

Acknowledgements acknowledgements.

Acknowledgements acknowledgements.

ABSTRACT

A Search for the Higgs boson produced in association with top quarks in multilepton final states at ATLAS

Chris Lester

Joseph Kroll

Abstract abstract.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

Chapter 2

Theoretical Background

The Standard Model of particle physics (SM) is an extraordinarily successful description of the fundamental constituents of matter and their interactions. Many experiments have verified the extremely precise prediction of the SM. This success has culminated most recently in the discovery of the Higgs Boson. This chapter provides a brief introduction to the structure of the SM and how scientists are able to test it using hadron colliders. It focuses primarily on the physics of the Higgs boson and its decay to top quarks. I stress the importance of a measurement of the rate at which Higgs Bosons are produced in association of top quarks, as a new, rigorous test of the SM. The experimental search for this production mode in multi-lepton final states is the general subject of this thesis.

2.1 The Standard Model

2.1.1 The Standard Model Structure

The Standard Model (SM) [1, 2, 3, 4] is an example of a quantum field theory that describes the interactions of all of the known fundamental particles. Particles are understood to be excitations of the more fundamental object of the theory, the field. The dynamics and interactions of the fields are derived from the Standard Model Lagrangian, which is constructed to be symmetric under transformations of the group $SU(3) \times SU(2)_L \times U(1)$. $SU(3)$ is the group for the color, $SU(2)_L$ is the group for weak iso spin, and $U(1)$ is the group for weak hyper-charge.

Demanding these symmetries be local, gauge symmetries allows the theory to be re-normalizable [5], meaning that unwanted infinities can be absorbed into observables from theory in a way that allows the theory to be able to predict physics at multiple energy scales. Gauging the symmetries results in the introduction of 8 massless gluons, or the boson[1] carriers of the strong force [6] from the generators $SU(3)$ symmetry, and the 4 massless bosons, carriers for the weak and electromagnetic forces from the 3 generators of the $SU(2)$ and 1 generator of the $U(1)$ group. The weak and the electromagnetic forces are considered part of a larger single unified electroweak group $SU(2) \times U(1)$ and the associated generators mix.

Matter particles are half-integer spin fermions and are represenations of the symmetry groups. Singlets of the $SU(3)$ are called leptons, do not have a color charge, and, therefore, do not interact with the strong force. Quarks, as triplets of the $SU(3)$ group, do interact with the strong force. The SM is a chiral theory: the weak force violates parity, as it only couples to left-chiral particles or right-chiral antiparticles. This means that right-chiral and left-chiral fermions arise from different fields, which are different representations of the $SU(2)_L$ group.

The discovery of particles and new interactions in various experiments is intertwined with the development of the theory that spans many decades and is not discussed in detail here. But these experiments have proven the above model and symmetries to be an overwhemlimg success. So far, 3 separate generations of both quarks and leptons have been discovered, differing only by mass. The gluons and the 4 electroweak bosons have also been discovered(W^+, W^-, Z^0, and γ). The reason for this 3-fold replication is not known. Figure 2.1 shows a table of the known SM particle content.

2.1.2 Electroweak Symmetry Breaking and the Higgs

Despite the simple structure of theory, the discovery of massive fundamental particles creates two sets of problems both related to $SU(2)_L \times U(1)$ symmetry. First, the force-carrying bosons must enter the theory without mass or the symmetries will be explicitly broken in the Lagrangian. Second, adding fermion masses to theory in an ad-hoc way allows the right-chiral and left-chiral fermions to mix. Since they possesses different quantum numbers, as different representations of the weak-isospin group, this too breaks gauge invariance.

To solve these problems, spontaneous electro-weak symmetry breaking (EWSB) is introduced via the Brout-Englert-Higgs mechanism [7, 8, 9]. A massive scalar field in an electro-weak doublet is added to the theory with 4 new degrees of freedom and a potential which includes a quartic self-interaction term. Each fermion field interacts with the scalar field via a different Yukawa coupling, which unites the left and right chiral fields of a single particle type. This field explicitly preserves all of the symmetries, but the minimum of the potential does not occur when the expectation of the

[1]bosons are full integer spin particles that obey Bose-Einstein statistics, while fermions are half-integer spin particles that obey Fermi-Dirac statistics

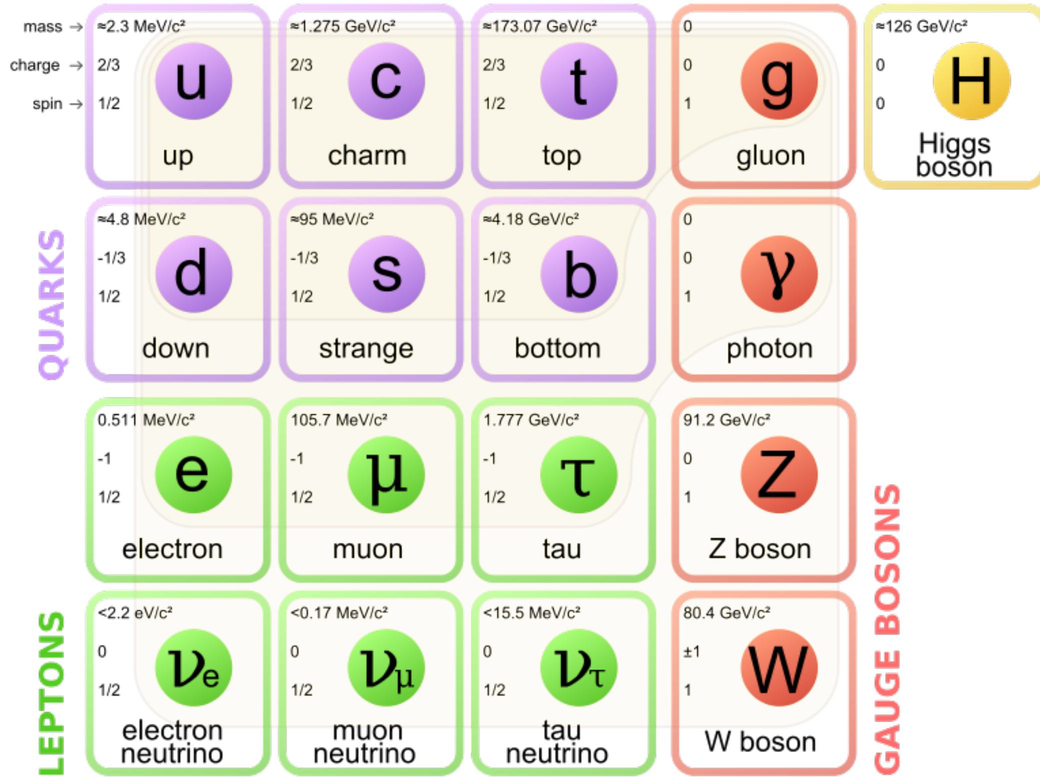


Figure 2.1: The Standard Model Particle Content

field is zero. The field eventually falls to a state, where it acquires a non-zero vacuum-expectation value. A non-vanishing field must point in a particular direction of weak-isospin space, breaking the symmetry.

The consequences of this spontaneous symmetry breaking are tremendous. The universe is filled with a field that has a non-zero expectation value. The theory can be expanded around this new value and 3 of the degrees of freedom can be interpreted as the longitudinal polarizations of the W^+, W^-, and Z^0, while the 4th remains a scalar field, called the Higgs field with an associated particle called the Higgs particle or "Higgs". The weak bosons acquire a mass via their longitudinal polarizations and the Yukawa couplings of the scalar field to the fermions now behave like a mass term at this new minimum.

2.1.3 The Standard Model Parameters

Confronting the SM with experiment requires the measurement of 17[2] free parameters, which are unconstrained from the theory. These free parameters include the fermion masses from the Yukawa couplings, the force coupling constants, the angles and phase of the mixing between quarks, and constants from the Higgs and electroweak sector[3].

Experiments have provided a number of measurements of the parameters of the SM[10]. With the discovery of the Higgs boson and the measurement of the Higgs mass, all of the parameters of the SM can be estimated and statistical procedures can assess the relative agreement of overlapping measurements to test the self-consistency of the SM. The GFitter collaboration assembles all relevant electroweak observable measurements into a statistical model and then allows certain measurements to float within their uncertainty to allow for a fit among multiple correlated measurements[11]. These correlations arise for two reasons. First, measurments are made that often depend on multiple SM parameters. Second, radiative corrections often cause parameters to depend on each other. For instance, the Higgs mass is sensitive to both the W mass and top mass, through loop level corrections.

Figure 2.2 shows the fitted constraints on 4 key SM parameters (M_H, M_W, M_t, $sin^2\theta_w$) with actual measurements overlaid. The plots show both the removal and inclusion in the fit of key measurements to assess their overall impact. The addition to the fit of the measured Higgs mass from the ATLAS and CMS collaborations creates a small tension, as the other observables prefer the mass to be much lower ($\sim$ 80 GeV). This tension in the combined electroweak fit as a result is not statistically significant with a p-value of 0.07. The SM seems to be self-consistent.

2.2 Collider Physics and the Higgs

To test the theory, physicists accelerate particles to extremely high energies and force them to interact through collisions. Typically, the particles accelerated are electrons or protons, since they are stable. Electron-positron collider machines have a rich history of discovery and measurement in particle physics. The advantage of electron accelerators is that the colliding element is itself a fundamental particle. However, due to synchrotron radiation, the curvature of the beam line becomes problematic for high energy beams. On the other hand, proton-proton and proton-anti-proton colliders can be accelerated in rings without large losses due to synchrotron radiation, but the actual colliding objects

[2]There are additional parameters from neutrino mass terms and mixing but it is unclear how to include these into the Standard Model, since it does not predict right-chiral neutrinos

[3] The electroweak sector includes parameters like mass of the $W^{\pm}$ and Z^0 bosons, the weak mixing angle,$sin^2\theta_w$, the fermi constant G_F, and Higgs Mass and vacuum expectation value. These parameters however are not wholly independent. As discussed above, it is only necessary theoretically to specify the two parameters relevant to the Higgs potential and the two coupling associated with the gauge groups

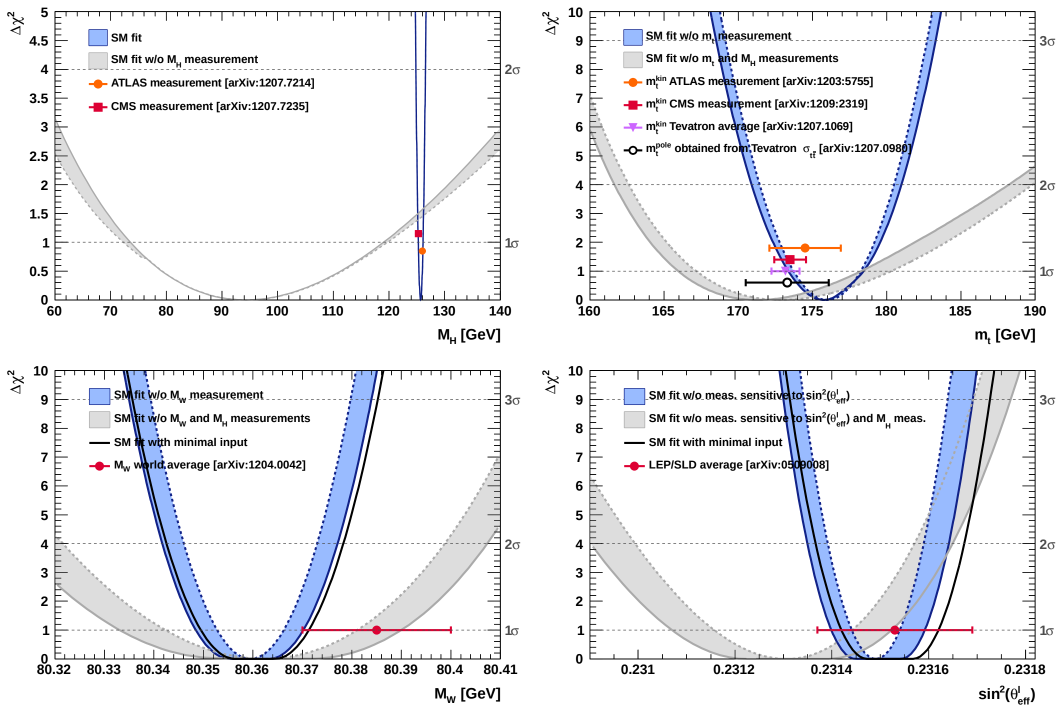


Figure 2.2: χ^2 as a function of the Higgs mass (top left), the top quark mass (top right), the W boson mass (bottom left) and the effective weak mixing angle (bottom right) for the combined SM fit from the GFitter group. The data points placed along $\chi^2 = 1$ represent direct measurements of the respective observable and their $\pm 1\sigma$ uncertainties. The grey (blue) bands show the results when excluding (including) the new M_H measurements from (in) the fits.

at high energies are the constituent quarks and gluons. This complicates analysis because the initial state of the hard-scattering system is not known on a per-collision basis and the momentum of hard-scattering system is unknown along the beam direction.

For hadron colliders, physicists must rely on form-factor descriptions of the colliding hadrons that describe the fraction of momentum carried by the hadrons constituent 'partons'. These are called parton-distribution functions, seen in Figure 2.3, and are factorized and integrated through the theoretical calculations of various collision processes [12].

At the Large Hadron Collider (LHC) protons are collided.The types of initial hard-scattering states at the LHC are quark-quark, quark-gluon, and gluon-gluon. Gluon collisions dominate overall, due to the large number of gluons inside the proton, though the relative importance of different initial states changes with the energy scale of the collision and the type of final state selected.

A prime motivation for the construction of the Large Hadron Collider was the discovery or exclusion of the Higgs boson[13]. LEP and the Tevatron excluded large swaths of possible Higgs boson masses,

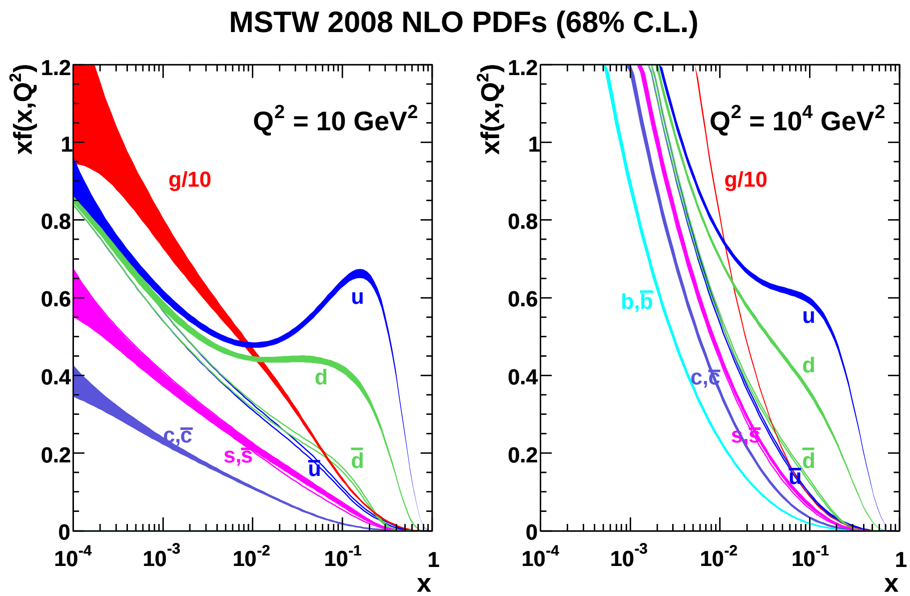


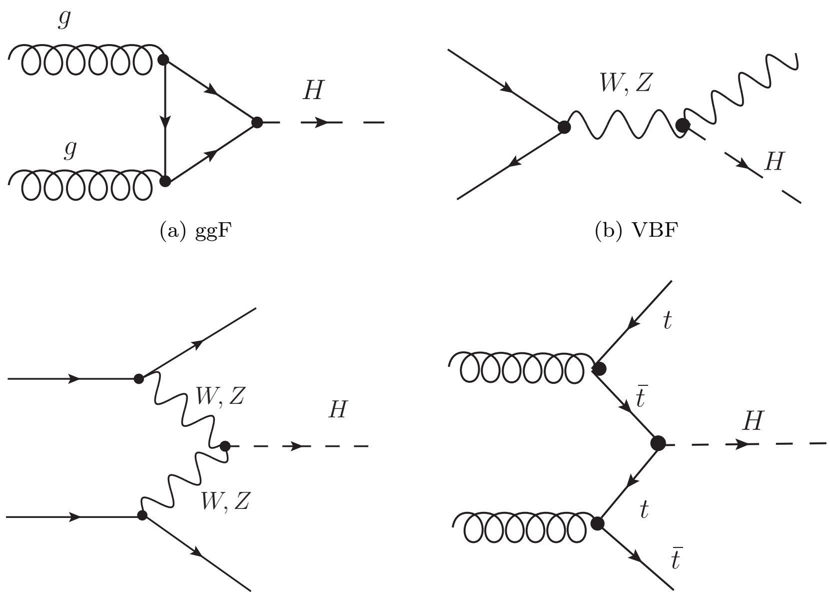
Figure 2.3: Proton Parton Distribution Functions (PDFs) from the MSTW Collaboration at $Q^2 = 10$ GeV^2 and $Q^2 = 10^4$ GeV^2

.

especially below 114 GeV. The Higgs mass was also known to have a theoretically motivated upper bound. The unitarity of diagrams including the $WWWW$ vertex required the Higgs mass to be below about 1 TeV. The LHC was designed to be able to eventually find or exclude a Higgs particle in this range [10].

Reaching this discovery or exclusion required an enormous dataset with collisions at high energies. Despite the fact that the Higgs couples to nearly every particle, Higgs boson production at the LHC is a low rate process. Because it couples to fermions proportional to mass and because the colliding particles must be stable and therefore light, production of the Higgs must occur through virtual states.

The Higgs boson can be produced through collision at the LHC via 4 mechanisms: gluon-fusion (ggF), vector-boson fusion (VBF), Higgsstrahlung (VH), and production in association with top quarks ($t\bar{t}H$). The diagrams are shown in Figure 2.4 and the production cross-sections as a function of Higgs mass for the 8 TeV LHC proton-proton running are shown in Figure 2.5 [14]. The largest production cross-section is via the gluon fusion channel at 20 pb, which proceeds through a fermion loop that is dominated by the top quark, because of its large Yukawa coupling to the Higgs. Because the Higgs couples to every massive particle, it has a rich set of decays also seen in Figure 2.5, especially for $m_H = 125$ GeV. Studies of Higgs properties at hadron colliders offers many tests of the Standard Model and ample room for new physics searches. These tests specifically can verify



(a) ggF

(b) VBF

(c) VH

(d) $t\bar{t}H$

Figure 2.4: Dominant Higgs production modes at the LHC

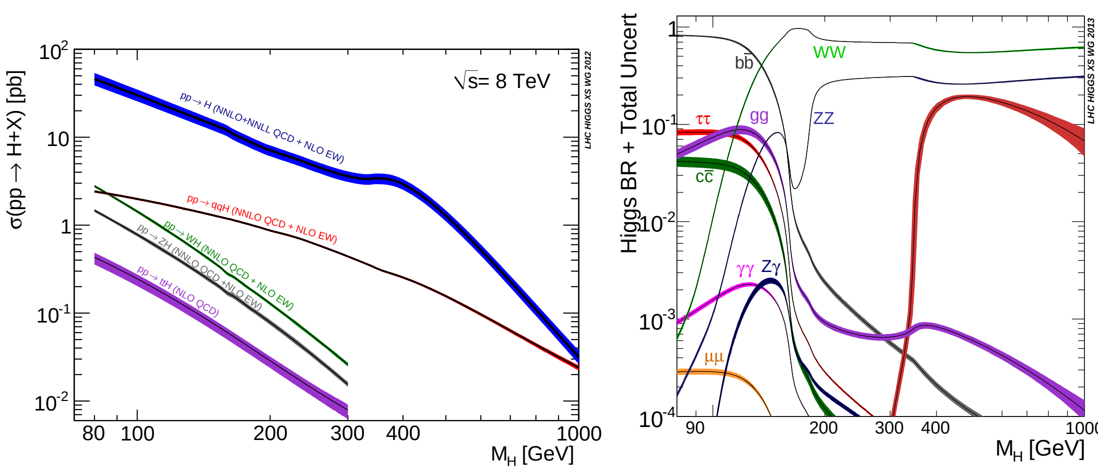


Figure 2.5: 8 TeV LHC Higgs production cross-sections (left) and decay branching fractions

.

441 the link between Yukawa coupling and the particles mass and further constrain details of EWSB by examining Higgs coupling to the weak bosons.

2.2.1 Higgs Discovery at the LHC

In 2012 both ATLAS and CMS announced the discovery of a new boson consistent with the Higgs by examining the results of Higgs searches in a number of decay channels ($H \to W^+W^-$,$H \to Z^0Z^0$, and $H \to \gamma\gamma$) in the 2011 dataset at $\sqrt{s}$ =7 TeV and part of the 2012 dataset at $\sqrt{s}$ =8 TeV. By 2013 and 2014, both experiments have updated and/or finalized their results for the full 2011 and 2012 datasets [15, 16]. I will focus on the ATLAS results, which measured both the Higgs mass[17] and spin[18], as well as provided initial constraints of the Higgs couplings to different particles.

Figure 2.6 show the results of the searches in all of the measurement channels as well as constraints on the SM Higgs coupling parameters in an example fit, where the couplings to the top-quark, bottom-quark, W,Z, and τ are allowed to fluctuate independently. These rely on measurements binned in different production and decay channels. They are dominated by higher statistics results in the gluon-fusion production modes, but measurements in the VH and VBF modes are close to SM sensitivity.

The combined results show basic agreement with the SM with much room for improvement with the addition of new production and decay modes and higher statistics. The coupling constraints are particularly strong for the W and Z, which are the most sensitive decay channels, and top quark due to the dominance of the top Yukawa in the ggF loop.

2.2.2 The Importance $t\bar{t}H$ Production

Notably absent thus far in the SM are searches for the Higgs in the $t\bar{t}H$ production channel, due to the low production rate and lack of statistics. Searches are underway and initial results are close to SM sensitivity for ATLAS and CMS.

Measuring the $t\bar{t}H$ production rate is important, because $t\bar{t}H$ production depends on the top Yukawa coupling at tree level. Comparing the predicted Yukawa coupling from top mass measurements to the coupling from the wholly independent Higgs production measurements is a very direct test of the Higgs' involvement in providing mass for the fermions in the SM.

The top Yukawa coupling is already constrained from current measurements of the ggF production process, since the ggF loop is dominated by top quarks. However, new, colored particles could be present in the loop. Comparison of the gluon-fusion and the $t\bar{t}H$ modes would allow for disentangling the effects of these possible new particles[19]. The simplest of new phyiscs models, allowing for the modification of the ggF loop, introduce a new generation of quarks. However, fourth generation quarks, which obtain mass from a Higgs Yukawa coupling, are already largely excluded due to their enormous effects on the Higgs production cross-section[20]. Other exotic scenarios allow for new

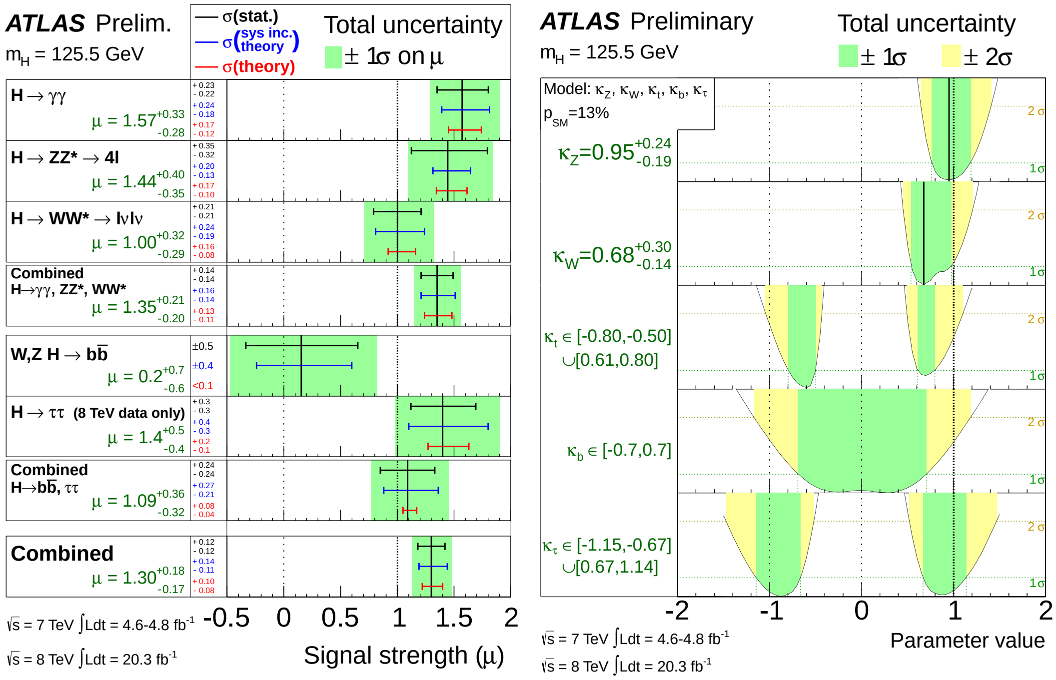


Figure 2.6: ATLAS Higgs combination results for all SM measurement channels as ratios of the measured to SM production cross-sections (left) and extracted Higgs coupling constraint scale-factors for a combined fit to the measurement channels, where the W,Z, top-quark, b-quark, and τ couplings are allowed to float. The p-value of this particular model is 0.13 and in agreement with SM expectations

colored particles, which are not entirely constrained by present measurements[21, 22, 23]. These include, for instance, supersymmetric models involving the stop quark.

With the level of statistics available in Run I dataset, very strict constraints on the top Yukawa coupling are simply not possible and the measurment presented in this thesis is a first step. Future, high-statistics datasets will have the ability to provide better measurments and $t\bar{t}H$ production will become very important. Despite similar uncertainties on the overall production cross-sections for $t\bar{t}H$ and the ggF, $t\bar{t}H$ has the advantage that most of these uncertainties would cancel for $t\bar{t}H$ if normalized to the topologically similar $t\bar{t}Z$. Finally, the uniqueness of the experimental signature means that searches for $t\bar{t}$ signatures can be performed for a variety of Higgs decays ($\gamma\gamma$, $b\bar{b}$, WW,ZZ, and $\tau\bar{\tau}$ with roughly similar degrees of sensitivity (within a factor of 10)[19].

It is important to note the importance of the top Yukawa coupling due to its enormous size compared to other couplings. For instance, the top Yukawa is 350000x as large as the electron Yukawa coupling. The top Yukawa coupling, along with the Higgs mass, is one of the most important pieces of the renormalization group equations (RGE) responsible for the running of the parameter that determines the Higgs self-coupling λ. If this parameter runs negative, then the potential responsible

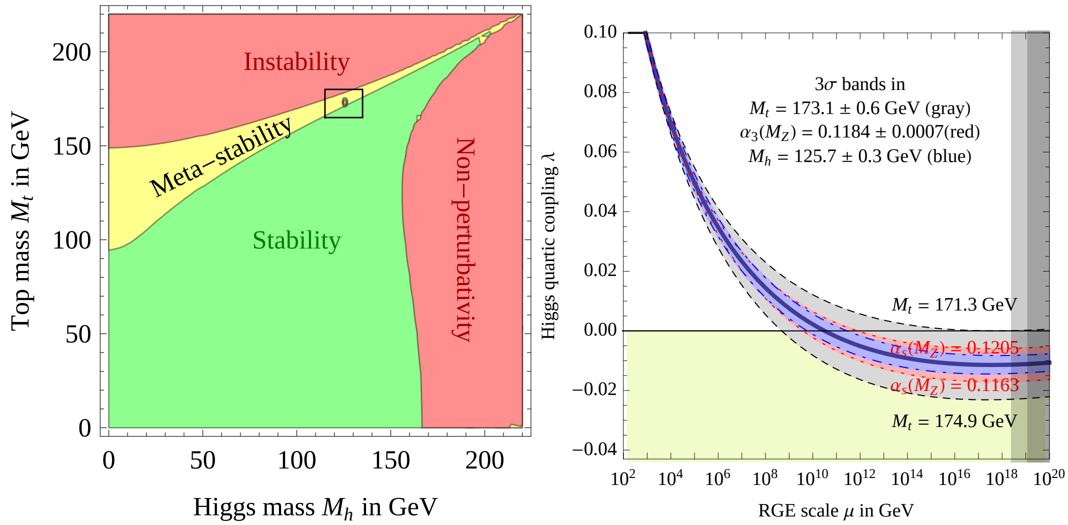


Figure 2.7: RGE for the running of the SM parameter, λ for the Higgs self-coupling term with present values and uncertainty bands for M_H and M_t (left). The two-dimensional plot colored (right) shows regions for which the SM is stable, unstable and metastable based on this RGE.

for the entire mechanism of EWSB no longer has a minimum and becomes unbounded, resulting in instability in the universe [24]. Metastability occurs when the shape of the potential allows for a false local minimum. Figure 2.7 shows the running of this parameter, the regions for which the universe is stable, unstable and metastable. Current measurements suggest that universe lies in a metastable island[4]. This is a sort of fanciful aside, intended only to highlight the importance of the top Yukawa coupling and to suggest that new discoveries in the top-Higgs sector have far reaching consequences.

2.3 Conclusion

The Standard Model, despite its success in providing a unified description of fundamental particles and interactions into single theory, has its flaws. These have been discussed in depth elsewhere, but include issues like the description of massive neutrinos, the failure to include gravity, and the unnaturalness of large quantum corrections to Higgs parameters. For these reasons, it seems the SM might be a lower energy approximation to a more fundamental theory. The discovery of the Higgs boson at the LHC provided a stunning verification of one of the fundamental aspects of the theory but at the same time offers new area to search for glimpses of something more fundamental. The production of samples of Higgs bosons allows for a rich array of new tests of the Standard Model, which is now finally over-constrained by experiment. Searches for the $t\bar{t}H$ production, one category of

[4]The RGE assumed that there is no new physics at all energy scales

which is the topic of this thesis, provide tree-level access to a central parameter of the theory, the top Yukawa coupling, as well as access a variety of Higgs decays, which will eventually provide a rigorous new test of the SM.

Chapter 3

The Large Hadron Collider and the ATLAS Experiment

3.1 The Large Hadron Collider

Production of a sufficient number of high energy collisions to adequately explore particle physics at the electro-weak scale required the development of one of the most complex machines ever built, the Large Hadron Collider or LHC.

The LHC is the world's highest energy particle accelerator and is located 100m underneath the Franco-Swiss border at the European Organization for Nuclear Research (CERN) in a 26.7 km tunnel. The technology involved in the development of the LHC is an enormous achievement it its own right and is documented in detail here [25, 26, 27]. The LHC is a circular machine capable of accelerating beams of protons and colliding them at center of mass energies up to $\sqrt{s} = 14$ TeV at 4 collision sites around the ring, where 4 experiments are housed (ATLAS[28], CMS[29], LHCb[30], and ALICE[31]). Figure 3.1 is a diagram of the layout of the LHC and its experiments[32]. The LHC also operates in a mode with beams of heavy ions. The LHC is composed of thousands of super-conducting Niobium-Titantium magnets, cooled to 1.9° K with liquid Helium, which steer and focus the particle beams, and a superconducting resonant-frequency (RF) cavity, which boosts the beam to higher energies.

3.1.1 The Accelerator Complex

The accelerator complex is a progressive series of machines with the LHC as the final stage. Protons are obtained from hydrogen atoms and are accelerated to 50 MeV using the Linac2, a linear accelerator, before being injected into the Proton-Synchrotron Booster (PSB). In the PSB the protons are accelerated to energies of 1.4 GeV for injection in to the Proton-Synchrotron (PS). The PS accelerates

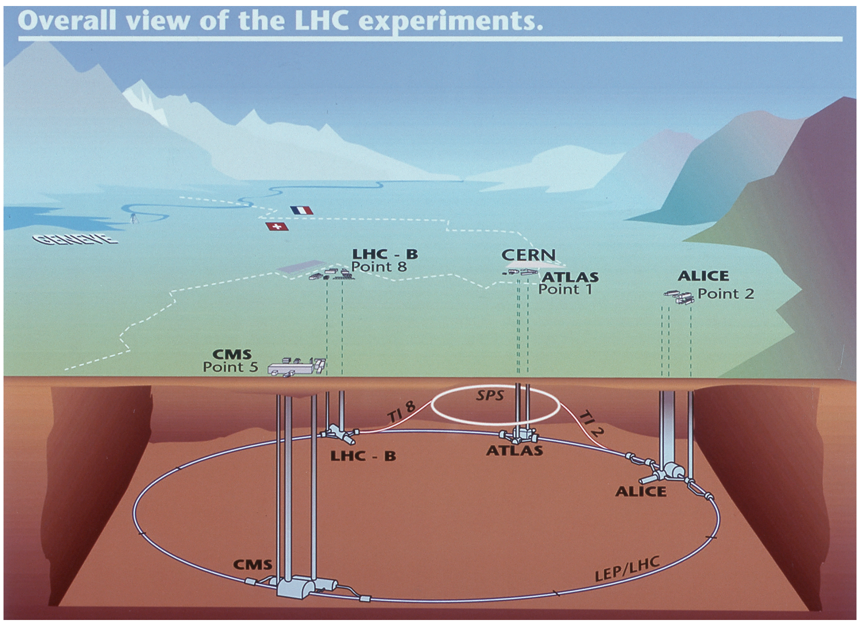


Figure 3.1: Diagram of the Large Hadron collider and location of the 4 main experiments (ATLAS, CMS, LHCb, and ALICE) around the ring. The diagram also shows the location of the SPS, the final booster ring in the accelerator complex that accelerates the protons to 450 GeV before injection into the LHC.

the protons to 25 GeV and dumps bunches into the Super Proton Synchrotron (SPS), where they are accelerated to 450 GeV and finally dumped into the LHC for full acceleration. The PS and SPS are circular accelerators that were important in past physics discoveries and have been re-purposed for use in the LHC complex.

3.1.2 Beam Parameters and Collisions

For the physics studied at the ATLAS experiment, the two most important parameters of the collisions are the center of mass energy (CME) and instantaneous luminosity ($\mathcal{L}$). High center of mass energies are necessary for the production of new high mass particles, and, because the constituents of the actual collisions are the partons of the proton, the CME of the collisions must in general be much higher than the mass of the particles produced.

The instantaneous luminosity of the collisions is a measure of the collision rate. The integrated luminosity over time is a measure of the size of the dataset and when multiplied by the cross-section of a particular process gives the total number of expected events produced for that process. Instantaneous luminosity depends on the number of colliding bunches of protons, the intensity of those bunches, the revolution frequency, and the normalized transverse spread of the beam in momentum and position phase space, called the emittance, and the transverse beam size. The LHC has the option for colliding beams with 2808 bunches of protons, each with around 10^{11} protons, at a rate of one bunch collision every 25 ns, or 40 MHz. These parameters correspond to a design luminosity of around 10^{34} cm^2 s^{-1} or 10 nb^{-1} s^{-1}, equivalent to 1 Higgs every 5 seconds.

3.2 The ATLAS Experiment

This section provides a brief overview of the ATLAS experiment. The ATLAS detector is centered on one of the LHC collisions points, located 100m underground. Through the combination of a number of subsystems, it designed to identify the particles arising from these collisions, measure the energy and momentum of these particles, and make fast decisions about the content of each collision, in order to save a small fraction of measured collision events for offline study.

ATLAS, shown in Figure 3.2, possesses cylindrical symmetry around the beam pipe. It weights 7000 tons, has a diameter of roughly 25m and length of 46m. ATLAS was designed to be a multi-purpose hermetic, particle detector, able to identify many types of particles, and designed to provide a snapshot of the entire collision event. The detector sub-systems form concentric rings around the beam-line at increasing distance. From closest to the beam outward, they are:

- **Inner Detector:** The inner detector (ID)[33, 34] is immersed in a solenoidal magnetic field[35] and provides measurements of charge particle tracks, through three subsystems: the Pixel Detector[36, 37], the Semi-Conductor Tracker (SCT)[38, 39], and Transition Radiation Tracker(TRT) [40, 41, 42].

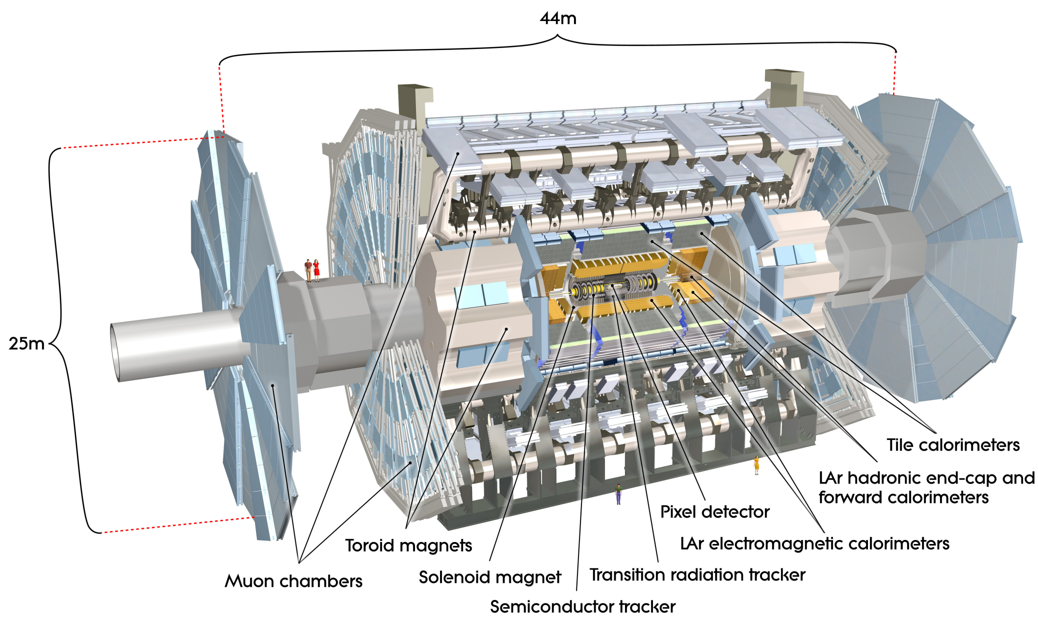


Figure 3.2: Diagram of the ATLAS detector and subsystems

- **Calorimeter:** The calorimeters measure the energy of particles that participate in the electromagnetic (photons, electrons) and hadronic interactions (pions, protons, neutrons, etc.), by forcing them to shower in dense material. The hermeticity of the calorimeters allows for missing transverse energy measurements. The calorimeter is composed of the liquid argon electromagnetic calorimeter (LAr)[43], the hadronic tile calorimeter[44], the liquid argon hadronic endcap calorimeter, and the forward calorimeters.

- **Muon Spectrometer:** The muon spectrometer (MS) sub-systems[45] form the outermost detector systems and measure the momentum of minimum ionizing muon tracks, as all other particles are stopped by the calorimeters. The muon systems are immersed in a toroidal magnetic field [35] and are composed of 4 different sub-systems for triggering and tracking measurements [46, 47, 48].

- **Triggering Systems:** The trigger and data acquisition systems[49, 50] read out data from the detector through a three-tiered hardware and software decision making framework to record the most interesting physical processes for a broad physics analysis program.

These systems are discussed in depth in the following sections.

3.2.1 Detector Coordinate System

ATLAS uses a right-handed coordinate system centered at the nominal proton interaction point. The beam line defines the z-axis. The $x-y$ plane is perpendicular to the beam line and is referred to as the transverse plane. The transverse plane holds special significance in reporting measurements, because the initial momentum of the hard collision system is 0 along the transverse plane in the laboratory rest frame. Particle momenta measured along the transverse plane is called transverse momenta, and labeled p_T. The momentum of the colliding proton-proton system is also 0 along the z-axis but the colliding partons may have vastly different momenta. Thus, momentum of the hard colliding system along the z-axis differs collision to collision.

Because ATLAS possesses a rough cylindrical symmetry, cylindrical and polar coordinates are used to describe particle trajectories and detector positions. The radial coordinate, R, describes transverse distances from the beam line. An azimuthal angle, ϕ, describes angles around the z-axis, and a polar coordinate θ describes angles away from the z-axis. The polar angle is often expressed in terms of pseudo-rapidity, defined as $\eta = -\ln(\tan(\theta/2))$. Distances in $\eta - \phi$ space are often used to describe the proximity of objects in the detector, $\Delta R = \sqrt{\eta^2 + \phi^2}$.

The 'barrel' and 'endcap' are classifications that are used to label the position of sub-detectors. Barrel sub-detectors occupy positions more central to the detector at $|\eta|$ values roughly less than 1-2, while the endcap calorimeters extend farther in $|\eta|$. The barrel-endcap transition region contains detector services. Also, the orientation of the detector elements are often different in the barrel and endcap to have optimal particle flux.

3.2.2 The Inner Detector

The ID makes measurements of the position of charged particles as they move through the detectors 3 sub-systems (Pixel Detector, SCT, TRT). The individual position measurements can be strung together to form a particle track. The entire ID is immersed in a 2T solenoidal magnetic field allowing for measurements of particle momenta through the curvature of the tracks. The ID is contained with a radius of 1.15 m and has a total length of 7m, allowing for particle tracking out to $|\eta| < 2.5$. Figures 3.3 and 3.4 show the placement of the ID sub-systems in the $R-\phi$ and $R-z$ planes.

The Pixel Detector has 80 million silicon read out channels (pixels) and is closest to the interaction point with the finest granularity. As charged particles traverse the silicon, they create electron-hole pairs, which subsequently drift in an electric field and can be captured and registered as a current pulse. The detector has three concentric layers of pixels in the barrel (to $|\eta| < 1.9$) and three endcap disks on each side of the barrel (to $|\eta| < 2.5$). The closest barrel layer to the beam pipe is called the

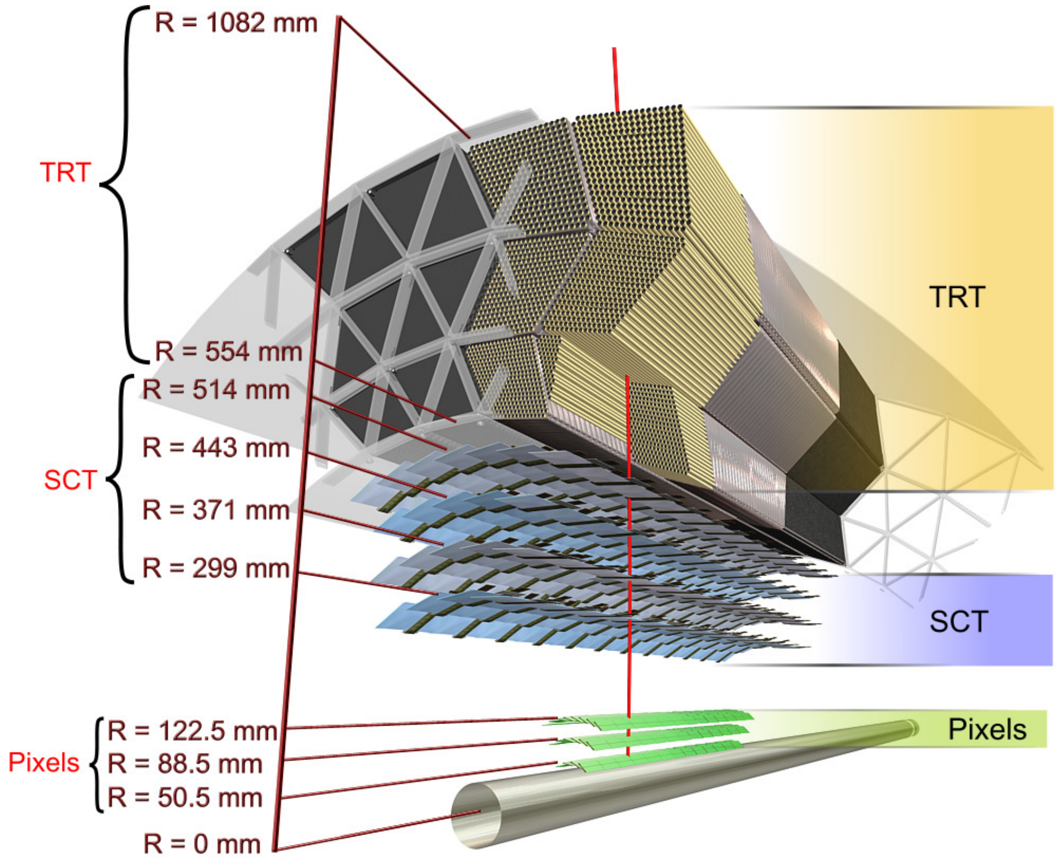


Figure 3.3: Diagram of the ATLAS ID in the $R-\phi$ plane showing the barrel view of the Pixel Detector, SCT, and TRT.

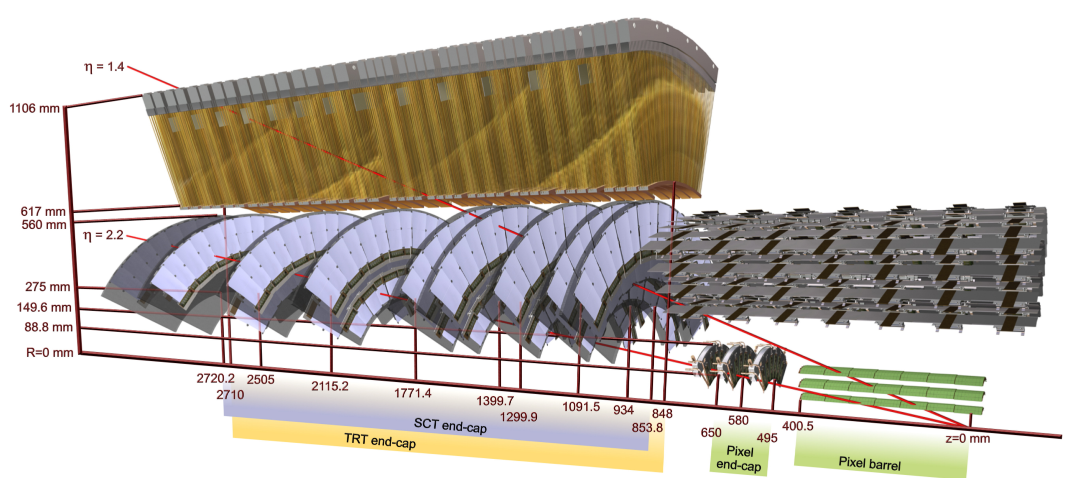


Figure 3.4: Diagram of the ATLAS ID in the $R-z$ plane showing the endcap view of the Pixel Detector, SCT, and TRT. Only one side of the endcap is shown.

b-layer. The pixels provide excellent hit resolution ($R-\phi$ accuracy of 10 μm and $z(R)$ accuracy of 115 μm in the barrel (endcap)).

The SCT uses similar silicon technology to the pixels, but the each SCT layer contains a double layer of silicon strips, which are much longer in length than width. The SCT has 4 million read out channels and is comprises 4 barrel layers and 9 endcap layers with coverage to $|\eta| < 2.5$. The double layers are inclined slightly with respect to each other so that these 1D sensors have 2D resolution for coincident hits. The resolutions are 580 μm in $z(R)$ for the barrel(endcap) and 17 μm in $R-\phi$.

The TRT is comprised of straw drift tubes filled with a gas mixture, containing mostly Xenon gas. Particles traversing the straws ionize the gas, and the liberated electrons drift to a wire at the center of the straw, which has an applied voltage, and induce an signal on the wire. The TRT has ∼300,000 straws. The barrel straws are arranged cylindrically along the z direction out to $\sim \eta < 1$ and the endcap straws point radially outward in the R direction. For this reason, the barrel(endcap) straws provide no measurement in the $R(z)$ directions. The drift tubes provide individual position measurements with resolutions of $\sim$ 130 μm. Each particle track has on average 35 hits, which is large compared to the Pixel and SCT tracks, which have on average 7 hits.

The TRT is unique in that it also provides particle identification measurements via transition radiation. Charged particles emit transition radiation when traversing a boundary between materials of different dielectric constants. The volume between the straws is filled with a radiator material, a polymer foil or foam, to provide this boundary condition. Transition radiation photons are emitted in the direction of the particle trajectory in the keV range and cause a much larger signal amplitude within the straw. Hits that cause a signal at a higher threshold are thus indicative of transition radiation. The probability for emission transition radiation depends on the relativistic γ of the traversing particle. Because electrons are much lighter than any other charged particle, their γ-factors tend to be high enough to induce transition radiations, as opposed to pions, muons and other particles.

Combined tracking of particles through the 3 sub-detectors results in track momentum measurements from 500 MeV, the minimum energy need to leave the ID due to the magnetic field, to a few TeV. The track p_{T} resolution is roughly $0.05\% \cdot p_{\mathrm{T}} \oplus 1\%$.

3.2.3 The Calorimeter

The ATLAS calorimeters measure the energy of electrons, photons and hadrons with $|\eta| < 4.5$. They induce a particle shower via electro-magnetic and nuclear interactions with the detector material and are deep enough to ensure that all or most of the shower energy remains contained. Exceptioanlly,

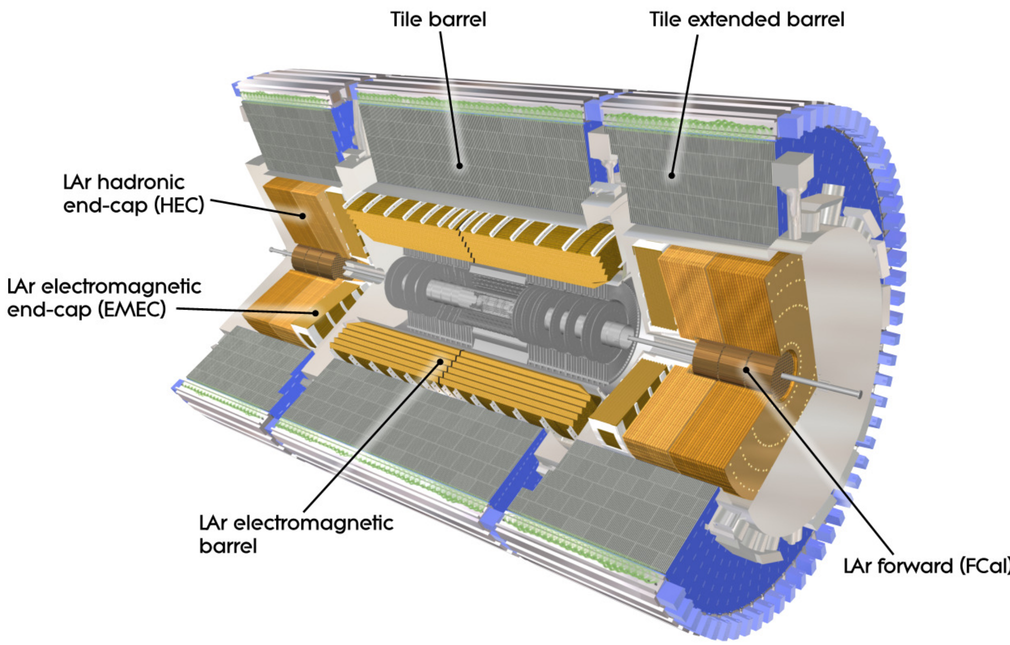


Figure 3.5: Diagram of the ATLAS calorimeters

muons pass through the ATLAS calorimeters leaving relatively little energy behind. ATLAS calorimeters are sampling calorimeters meaning that the active material of the detector only measures a small fraction of the energy produced by the shower. The overall shower energy is inferred from this fractional measurement. The rest of the material is inactive, dense material, designed to induce showers. The calorimetry system is grossly divided longitudinally (radially) into electro-magnetic (EM) and then hadronic segments, operated with different technologies. Figure 3.5 diagrams the layout of the calorimeter system.

The EM calorimeter (LAr), which is located directly outside of the solenoid magnet but within the same cryostat, has a accordion design with lead absorber and liquid argon active material. The accordion design ensures uniform coverage in ϕ. The barrel and endcap LAr extend to $|\eta| < 2.47$. The LAr provides highly granular measurements in $\eta - \phi$ with 4 longitudinal segments, totaling $\sim$25-35 radiation lengths with the exception of the barrel/endcap transition region ($1.37 < |\eta| < 1.52$). The geometry of the barrel LAr calorimeter can be seen in Figure 3.6. The first longitudinal segment is called the pre-sampler, composed of a thin layer of active liquid argon, designed to detect early particle showers. The second segment is the most highly granular segment called the 'strips', as it is composed of thin liquid argon cells. The strips have a size of 0.025/8 x 0.1 in $\eta - \phi$ in the barrel

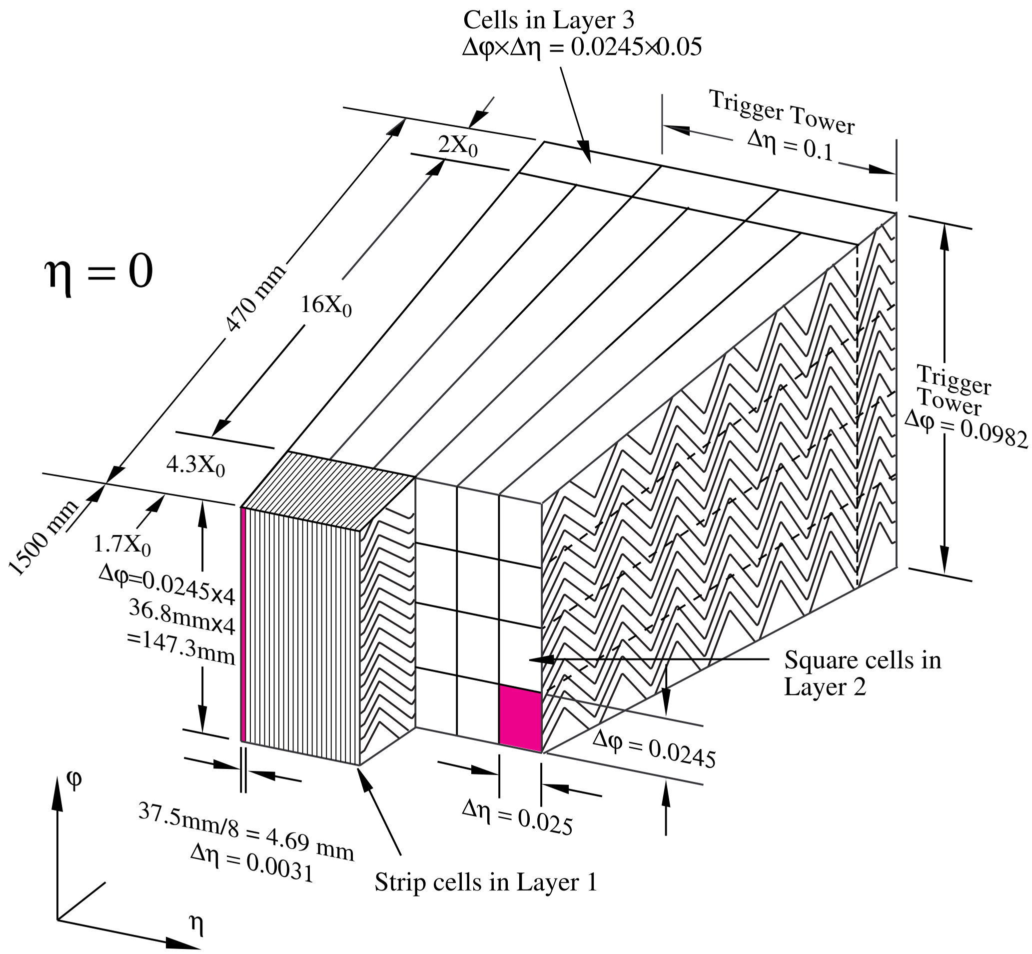


Figure 3.6: Diagram of the ATLAS LAr EM calorimeter showing the longitudinal segmentation and the $\eta - \phi$ cells for the central barrel region

with similar sizes in the endcap and are designed to be able to resolve single and double particle showers. This resolution is particularly useful in distinguishing $\pi^0 \rightarrow \gamma\gamma$ signatures from electron and photon signatures. The bulk of the radiation lengths and therefore the primary energy measurement come from the the third layer[5]. Each cell in this layer is 0.025 x 0.025 in $\eta - \phi$. The final layer is coarser, thinner and designed to estimate energy leaking out of the EM calorimeter. The forward EM calorimeters extend the η range and use the same technology, but are not used in this analysis. The energy resolution of the EM calorimeters is $\sigma_E/E = 10\%/\sqrt{E} \oplus 0.7\%$, measured in test beam data and confirmed in collision data.

The hadronic calorimeter is located directly behind the EM calorimeter. It is composed of tiles

[5]this layer is actually called 'layer 2', since the pre-sampler is referred to as 'layer 0'

of iron absorber and plastic scintillator in the barrel ($|\eta| < 1.6$), called the TileCal, and copper-liquid argon in the endcap ($1.5 < |\eta| < 3.2$), called the HEC. The calorimeters contain ~10-19 hadronic interactions lengths with multiple longitudinal segments to contain showers induced by the nuclear interaction of hadronic particles. The energy resolution of the hadronic calorimeters is $\sigma_E/E = 50\%/\sqrt{E} \oplus 3\%$. The intrinsic resolution of hadron calorimeters is much worse than electromagnetic calorimeters, because much of the energy is lost to the inelasticity of nuclear break-up.

3.2.4 The Muon Spectrometer

The MS measures the trajectory of particles outside of the calorimeters, using multiple different technologies. Generally, all charged particles except for muons are stopped by the calorimeter, and therefore the majority of particles in the MS are muons, with the exception of rare cases of hadronic punch-through. Particle momentum spectroscopy is made possible by an air-core toroidal magnet system, embedded in the barrel MS ($|\eta| < 1.4$), and two smaller end cap toroids that provide fields out to $|\eta| < 2.7$.

In the barrel region, the muon chambers are arranged in three cylindrical layers around the beam, while the in the endcap-regions the the layers are arranged perpendicular to the beam in wheels. The arrangement is depicted in Figure 3.7.

The chambers in the barrel and most of the endcap are constructed from Monitored Drift Tubes (MDTs) with an Argon gas mixture. Each chamber contains 3-7 drift tubes and provide hit resolutions of 80 μm per tube and 35 μm per chamber in the bending plane. For $|\eta| > 2.0$, Cathode Strip Chambers (CSCs) are used, primarily to handle the higher incident particle flux. They are composed of cathode strips crossed with anode wires in the gas mixture, but use similar drift technology as the MDTs and have resolutions in the bending plane 40 μm per chamber.

Separate chambers, called Thin Gap Chambers (TGCs), used in the endcaps, and Resistive Plate Chambers (RPCs), used in the barrel, provide less precise hit information but within a much quicker time window, and are therefore used for triggering, as the CSCs and MDTs are too slow.

3.2.5 The Trigger System

The ATLAS trigger system is designed to make quick decisions about individual particle collisions to reduce the enormous collision rate of 20 MHz to a much more manageable 400 Hz to be stored for offline analysis. Saving the full ATLAS data-stream would require space for 40 TB of raw data per second, but, more importantly, most of these collisions result in the uninteresting inelastic break-up of the colliding protons. To select out collisions to allow for a diverse physics program, ATLAS devotes a

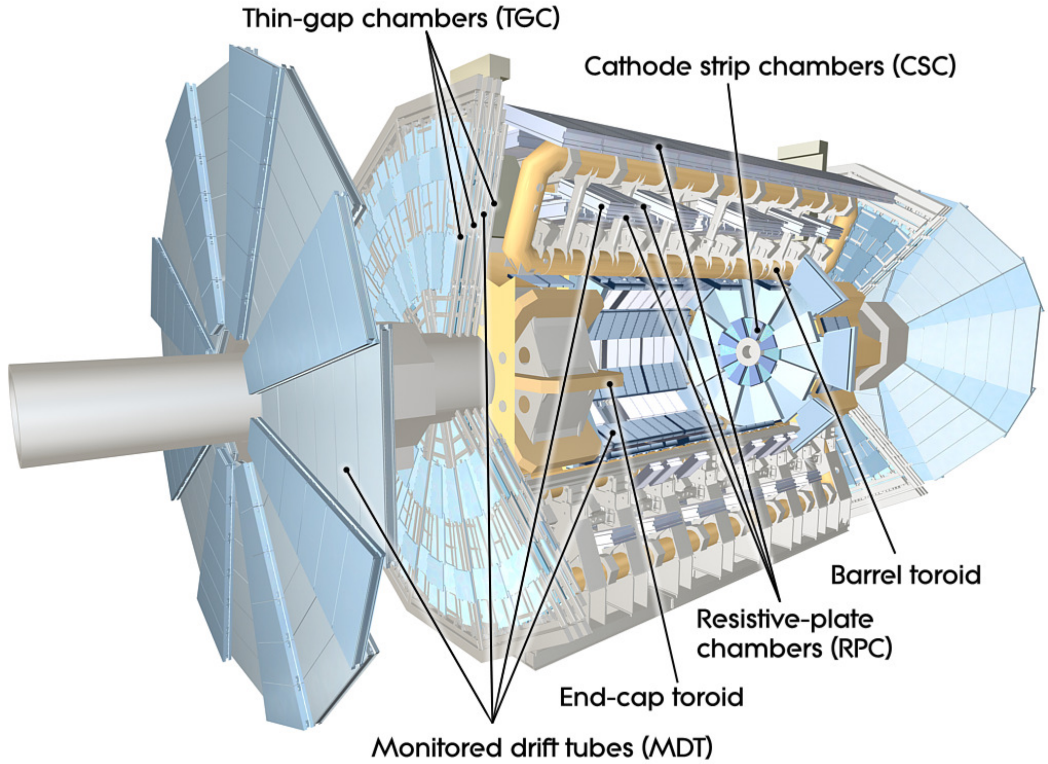


Figure 3.7: Diagram of the ATLAS muon system

large portion of the bandwidth to general purpose single lepton triggers ($\sim$ 250 Hz). The presence of leptons in the event indicates the presence of the weak or electro-magnetic interaction and therefore occurs at many order of magnitude less frequently then interactions involving the strong interaction. Moreover, many interesting physics signatures that are analyzable by ATLAS involve leptonic final states. The remaining bandwidth is allocated to jet, missing energy, tau, and unbiased supporting triggers.

The ATLAS trigger system is composed of 3 levels: level-1 (L1), level (L2), and the event filter (EF). The first level is hardware only trigger that reduces the input 20 MHz rate to $\sim$ 75 kHz, selecting 1 out of every 250 collisions. The available buffering on the FPGA chips means that the decisions need to be made within 2.5 μs. The L1 selection is based on calorimeter clustering and tracking finding in the MS for small areas of the detector called regions-of-interest (ROIs). It selects ROIs with significant energy.

The second and third stages L2 and EF are software based. The L2 algorithms perform more detailed object reconstruction for leptons, jets and photons inside of the ROIs provided by L1, by performing tracking and in depth calorimeter clustering algorithms. The decisions are made within 50 ms per event and pass 1 out of every 15 events to the EF. At the EF, events undergo full reconstruction using similar but faster versions of the algorithms used offline. The EF makes decisions on the presence of fully id-objects in the event and event topological quantities within 4s to reduce the L2 output by a factor of 10. The events that pass this stage are then written to tape for offline study.

3.2.6 Reconstruction: Jets, Muons and Electrons

Physicists analyze the collision event as a collection of identified objects, expressed as momentum 4-vectors. The process of converting the disparate detector signatures and signals into a unified 4-momentum description of individual objects is called reconstruction. These objects arise from the final state particles in the event, which can be combined and counted to infer properties of the hard scatter. The particles that make detectable signatures are those that are stable enough to pass through the detector: muons, electrons, photons, and quarks and gluons. Jets and b-tagged jets, muons and electrons are used in the $t\bar{t}H$ analysis to define our search regions and to separate the Higgs signal from backgrounds. Other analyses may used photons, taus and missing energy[6], but these are not discussed in depth here. Figure 3.8 shows an $R - \phi$ schematic of the interaction of various particle signatures in the ATLAS detector.

3.2.6.1 Tracks and Clusters

The basic components of reconstruction are sensor measurements, or hits, in trackers (ID, MS) and energy measurements in the calorimeter. Hits in the ID and MS undergo pattern recognition, which identifies hits that belong to a single track, and fitting, which fits a curve to the track to assess the particle trajectory. Charged particle trajectories are generally helical in a magnetic field, but the fitting algorithm takes into more detailed information about energy loss to material along the tracks length. The result of the fitting is an estimation of particle momentum 3-vector. Electrons, photons and hadronic particles leave clustered deposits of energy in the EM and hadronic calorimeters from their showers. Electron and photon showers are primarily contained with the EM calorimeter, while hadronic showers are deposit most of their energy in the hadronic calorimeters. The process of associating individual read-out cells of energy in the calorimeter to clusters of energy from the showers of individual particles is called clustering. From the basic pieces of tracks and clusters, more complex objects can be created.

[6] missing energy is the presence of momentum imbalance in the transverse plane of the calorimeter due to escaping neutrinos

3.2.6.2 Electrons

Electrons leave both a track in the ID and a narrow and isolated cluster of energy in the EM calorimeter, $\Delta R < 0.1$. Electron reconstruction proceeds using a sliding window algorithm, which scans a fixed size rectangle in $\eta - \phi$ space over the EM calorimeter cells to find relative maxima of energy in the window[51]. These maxima seed the clustering algorithms. Because electrons are light, they both lose energy to the material gradually through scattering and more catastrophically through the emission of a high energy photon, through interaction with nuclei. This process is called bremsstrahlung. Tracks for electrons are reconstructed differently because they must include the hypothesis that the electron loses significant energy through bremsstrahlung. Generally, the emitted photon is contained within the same energy cluster and therefore the sliding window algorithm is always wider in the direction of bending, ϕ. A single track is then matched to the cluster within certain minimum matching requirements in η,ϕ,and p_{T}. Electrons are distinguished from photon conversions, which also have a track, by lack of association with conversion vertices, found with a dedicated algorithm.

Electron have many lever arms for further identification to suppress backgrounds from fake sources. The narrowness of the shower shape, quality of track, and presence of transition radiation are used by cut-based and multivariate identification algorithms. This is discussed in depth in Chapter 4. Electrons are reliably reconstructed and identified with energies above 7 GeV.

3.2.6.3 Muons

Muons are reconstructed from a combination of ID and MS tracks, when possible. The two tracks must meet matching criteria to ensure they are from the same particle. The muon momentum 3-vector comes from the combined ID/MS fit. Muons leave little energy in the calorimeters and are generally isolated from other particles, when produced from electro-weak bosons. Identification algorithms make requirements on the number of tracking hits in the ID and MS and the quality of the matching of the two tracks. Muons are reliably reconstructed and identified with energies above 5 GeV. More about muon reconstruction and identification can be found here [52].

3.2.6.4 Jets

Quarks and gluons are colored objects that cannot exist alone on the time scales of detector measurements, due to confinement, a property of the strong force . When emitted, they undergo a process called hadronization, in which the convert into 'jets' of colorless hadrons that emerge collimated from the interaction point. The majority of these hadrons are charged and neutral pions, though other hadrons are often present. Jets are reconstructed using conglomerations of calorimeter energy clusters chosen via an anti-k_t algorithm, with a radius of ΔR <0.4 [53]. The algorithm has been shown to be infrared safe, meaning the jet quantities are not sensitive to low energy, small angle radiative divergences. Jets at ATLAS are reconstructed from 10 GeV, calibration of the energy scale and resolution are only available for energies greater than 20-25 GeV.

3.2.6.5 B-Tagged Jets

Generally, the flavor of the initiating quark is not known from the reconstructed jet, although gluon initiated jets and quark initiated jets have slightly different properties. B-quark jets, however, are unique in that the long life-time of the produced b-mesons allow for measurable decays in flight. This property is used to tag b-quark initiated jets. This analysis uses the MV1 tagging algorithm [54], which is a neural network based algorithm that looks for secondary displaced decay vertices inside the event and takes into account jet track parameters and energy flow with respect to these vertices. B-quark jets often involve b-meson decays to leptons, especially muons, which can be used to tag an orthogonal b-jet sample for studying tagging efficiencies.

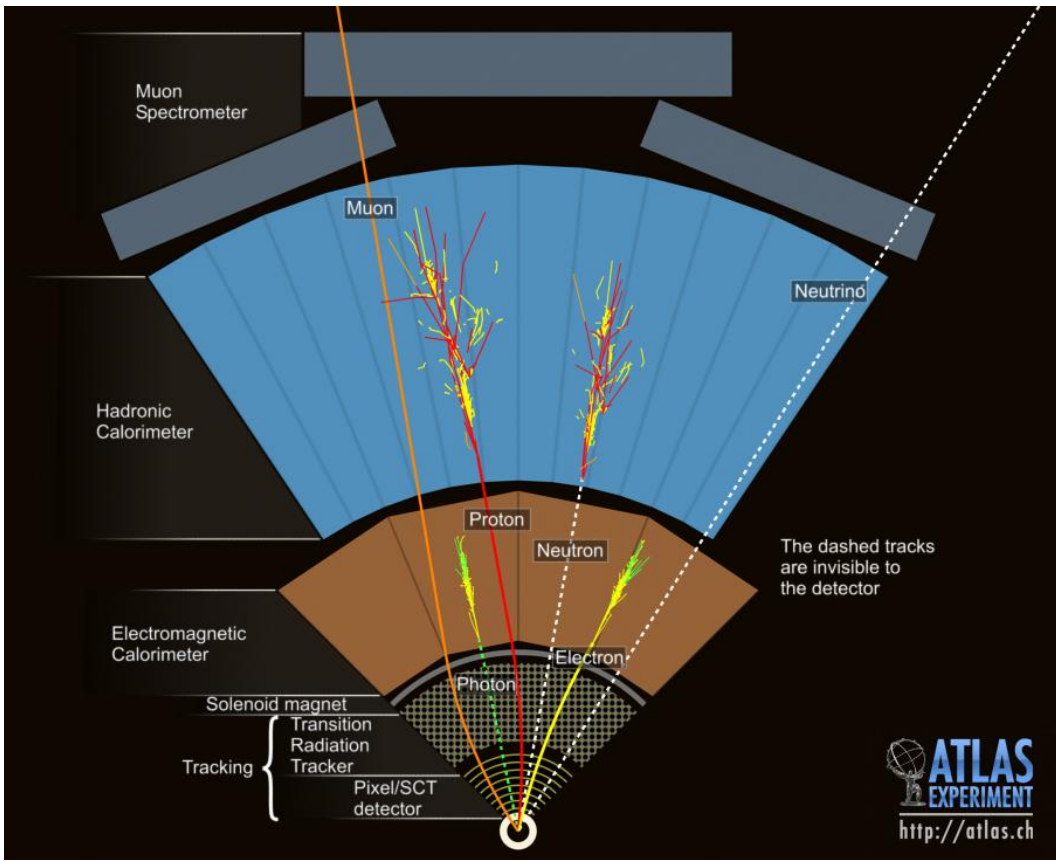


Figure 3.8: $R-\phi$ schematic of the ATLAS detector and various particle signatures

Chapter 4

Electrons

This chapter details the contributions I made to electron identification and efficiency measurments. It is not essential to continuity of the thesis in general but provides in depth documentation of the work I completed for the experiment. I focus on the electorn identification

4.1 Electrons at Hadron Colliders

High energy electron signatures are important elements of seaches and measurements at hadron colliders. The overwhemling majority of collisions that deposit energy in the detectors are the result of strong-force mediate interactions of the constituent partons. These collisions result in the production of high energy jets in the detector. Figure XX shows the cross-sections of various processes as a function of the center of mass energy of the collision. Physics invovling the electroweak interaction or even strong production of massive states occur many orders of magnitdue less frequently than the total inelastic cross-section. Interesting physics signatures, both stardard model and beyond, often involve the production of light leptons as a result of the decay of massive particles. Choosing events that have high energy electrons or muons targets events that contain electroweak vertices and dramatically reduce the background from the more copiously produced strong physics. Electron and muon energy and momenta are also relatively well-measured compared to jets. This allows for the use of well-resovled kinematic shapes used discriminate the sigantures of different processes in analyses.

At ATLAS, the primary datasets for most analyses are collected with electron and muon triggers. Electron triggers are particularly important, because the muon triggers system has a 20% small acceptance than the electrons. The challenge in identifying electrons is distguising the production of electrons from direct production of W and Z decays from electrons produced in the more copiuously produced b-meson decays, fake-electron signatures from rare jet fragmentations into charged and neutral pions, and photon conversions in the inner detector. The identification of electrons, the precise measurement of the identification efficien, and the measurement of the rate of fake electron signatures lead are often the most important and challengin peices of an analyis. The following sections discuss the identification of electrons for the primary electron trigger and offline physics analyses as well as the measurement of the electron identification efficiency in 2012. Because I had a major role in these projects, I will at times discuss their historical evolution and not simply focus of the particular measurement relevant to the $t\bar{t}H$analysis.

4.2 Identification of Electrons at ATLAS

Electron reconstruction

4.2.1 Pile-up and Electron identifiation

Plots of pile up differences in distributions

4.2.2 2011 Menu and Trigger

4.2.3 2012 Menu and Trigger

4.2.4 Electron Likelihood

4.3 Measurement of Electron Efficiency at ATLAS

4.3.1 Techniques

4.3.2 Issues

Chapter 5

Analysis Summary

This chapter provides an overview the of analysis searching for SM production of the Higgs boson in association with top quarks in multi-lepton final states. The analysis searches in signal regions (SRs) with 2 same-sign, 3 and 4 light leptons (e, μ), which are sensitive to Higgs decays to vector bosons, $H \to W^{\pm}W^{\pm}$ and $H \to Z^{\pm}Z^{\pm}$. We refer to these channels as 2ℓ SS, 3ℓ, and 4ℓ through the rest of this document.

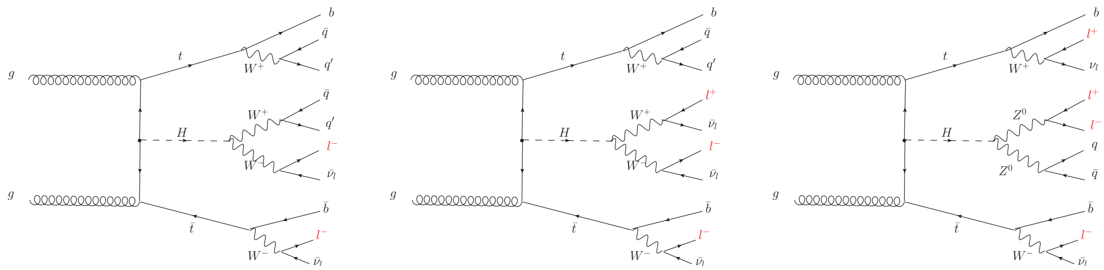
The multi-lepton channels form a complement to already completed $t\bar{t}H$ searches in final states targeting the $H \to b\bar{b}$ [55], $H \to \gamma\gamma$[56]. The $t\bar{t}H$ searches in the $H \to \tau\tau$ decay modes were developed concurrently with the multi-lepton searches, but we do not discuss these here. Of this set of complementary searches, the multi-lepton and $b\bar{b}$ are the most sensitve.

Based on SM production cross-sections, observation lies just outside the sensitivity of the Run I dataset, even when combining all searches. Instead, the analyses provide an opportunity to constrain for the first time the $t\bar{t}H$ production mode with limits reasonably close to the actual production rate. The multi-lepton analysis is therefore optimized to overall sensitivity to the $t\bar{t}H$ production rather than individual decay modes, which would be more useful for constraining Higgs couplings.

Detailed description of the event and objection section are provided in Chapter 7, background modeling in Chapter 9, the effect of systematic errors and the statistical analysis in Chapter 9 and final results in Chapter 10.

5.1 Signal Characteristics

The signal is expected to be characterized by the presence of 2 b-quark jets from the top quark decays, isolated leptons from vector boson and tau decays, a high jet multiplicity and missing energy from neutrinos. Three Higgs boson decays are relevant for this analysis: W^+W^- , $\tau^+\tau^-$ and ZZ .



(a) 2ℓ SS (b) 3ℓ (c) 4ℓ

Figure 5.1: Example Feynman diagrams for the 3 $t\bar{t}H$ multi-lepton categories.

All modes are generally dominated by the WW signature, though the 3ℓ and 4ℓ channels possess some contribution from the$\tau\tau$ and ZZ decays. Table 5.1 provides the fractional contribution of the main Higgs decay modes at the generator level to $t\bar{t}H$ search channels and Figure 5.1 shows example diagrams for each channel. In general, the number of leptons is anti-correlated with the number of jets, since a vector boson can either decay leptonically or hadronically, such that:

- in the 2ℓ SS channel, the $t\bar{t}H$ final state contains 6 quarks[7]. These events are then characterized by the largest jet multiplicity.
- In the 3ℓ, the $t\bar{t}H$ final state contains 4 quarks
- In the 4ℓ channel, the $t\bar{t}H$ final state contains a small number of light quarks, 0 ($H \to W^+W^-$ case), 2 or 4 ($H \to ZZ$ case).

Table 5.1: Contributions of the main Higgs decay modes to the 3 multi-lepton $t\bar{t}H$ signatures at generation level.

Signature	$H \to WW$	$H \to \tau\tau$	$H \to ZZ$
Same-sign	100%	–	–
3 leptons	71%	20%	9%
4 leptons	53%	30%	17%

5.2 Background Overview

For all channels after selection, the size of the signal is of similar order to the expected size of background. Background processes can be sorted into two categories:

[7] this does not include additional quarks from radiation

- **Reducible:** These processes cannot lead to a final state compatible with the signal signature without a mis-reconstructed object. This category includes events with a prompt lepton but with mis-reconstructed charge and events with jets that "fake" leptons. The main backgrounds of this sort are $t\bar{t}$ and Z+jets . Data-driven techniques are used to measure the rate of these processes and strict object selection and isolation requirements are used to reduce their rate.

- **Irreducible:** Events which can lead to the same final state as the signal. The main background of this category are: vector boson production (V) associated with top quarks ($t\bar{t}V$), a Z boson produced in associated with a top quark (tZ), $W^{\pm}Z$, and ZZ . They are modeled using the Monte Carlo simulations. In general, these backgrounds are combatted with jet and b-tagged jet requirements. Although the jet multiplicity of $t\bar{t}V$ is high, the multiplicity of $t\bar{t}H$ events is still higher.

5.3 Analysis Strategy

The analysis search is conducted in 3 channels, based on counting of fully identified leptons: 2ℓ SS, 3ℓ, and 4ℓ, with cuts optimized separately for each. We further divide the 2ℓ SS into sub channels based on the number of jets and flavor of the leptons and the 4ℓ channel into sub-channels enriched and depleted in opposite-sign (OS) leptons arising from Z decays.

This analysis is a counting experiment, meaning that the only quantities signficant to the result are the event counts in the signal regions and not the event shapes. The measured background rates, expected signal rates and systematic uncertainties are fed into a Poisson model and fit to the observed data. The parameter of interest in the fit and the result of this measurment is, μ, the ratio of the fitted number of $t\bar{t}H$ events in the signal regions to expected number of $t\bar{t}H$ events in the signal regions. Since we assume SM branching ratios, μ can be considered the ratio of the measured $t\bar{t}H$ cross-section to the observed $t\bar{t}H$ cross-section, and we expect the fitted μ to be close to 1 with large statistical errors.

We express the final result as a measurment of μ with uncertaintites and 95% upper limit on the value of μ: μ-values higher than this value will be considered excluded. We provide these results for each channel individually and combined.

Chapter 6

Dataset and Simulation

6.1 Data

6.1.1 The 2012 Dataset

The $t\bar{t}H$ analysis uses the entire 2012 ATLAS dataset only, collected from April to December. The size of the dataset corresponds to 20.3 fb^{-1}, after passing data quality requirements, ensuring the proper operation of the tracking, calorimeter and muon subsystems. The LHC successfully produced datasets for physics studies in 2010, 2011 and 2012. The 2012 proton-proton dataset was delivered with collisions with a CME of 8 TeV with bunch collisions every 50 ns[57].

Figure 6.1 shows the accumulation of the 2012 dataset over time. Despite doubling the bunch spacing above the design of 25 ns, the luminosity neared the design luminosity due to unexpected improvements in the transverse beam profile[58]. This increased the amount of pile-up, or number of collisions per bunch crossing and in general collision events were busier due to these multiple interactions. Figure 6.2 shows the average number of interaction per bunch crossing for the 2011 and 2012 datasets. The 2012 dataset shows an average of 20-25 interactions.

The dataset must contain either a primary muon or primary electron trigger (`EF_e24vhi_medium1` OR `EF_e60_medium1` OR `EF_24i_tight` OR `EF_36_tight`). The electron triggers require a electron with at least 25 GeV of calorimeter energy, passing the medium identification requirement and loose tracking isolation. Above 60GeV, the isolation requirement is dropped and the identification is loosened slightly. The muon trigger requires a good inner detector track and matching hits in the muon spectrometer, as well as loose tracking isolation, which is also dropped about 36 GeV.

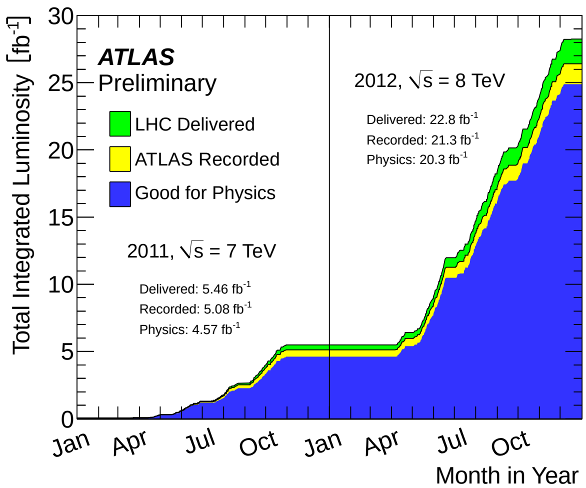


Figure 6.1: Plot showing the accumulation of the integrated luminosity delivered to the ATLAS experiment over 2011 and 2012. The rough size of the usable, physics ready dataset for 2012 is 20 $\text{fb}^{-1}$and is the dataset used.

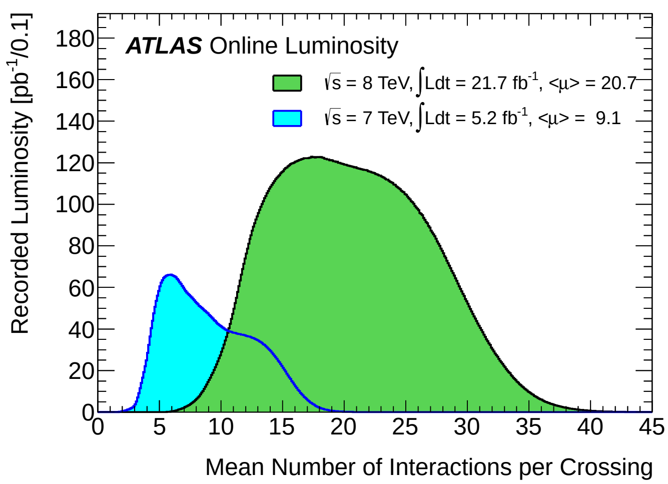


Figure 6.2: The average number of interactions per bunch-crossing for the 2012 and 2011 LHC proton-proton dataset. Most of these interactions are uninteresting but leave energetic signatures in particle detectors called pile-up which interfere with measurements

6.2 Simulation

Simulation samples are used to determine the overall event selection acceptance and efficiency and model the number of events in the signal regions for prompt backgrounds and signal. The simulated samples are created using parton distribution function (PDF) and use Monte Carlo (MC) techniques to model the hard parton scatter, underlying event activity and parton showering and hadronization.

Table 6.1: Monte Carlo samples used for signal description.

Process	Generator	Cross-section [fb]	$\mathcal{L}$ [fb^{-1}]	Detector simulation
ttH→allhad+H	PowHel+Pythia8	59.09	2146.5	Full
ttH→ljets+H	PowHel+Pythia8	56.63	2238.9	Full
ttH→ll+H	PowHel+Pythia8	13.58	9332.0	Full

The samples are then passed through a full ATLAS detector simulation[59] based on GEANT4 [60]. Small corrections are then applied to re-scale object identification efficiencies, energy scales, and the pile-up based on control regions from data. These corrections are discussed in Chapter 9.

6.2.1 Signal Simulation

The $t\bar{t}H$ production is modeled using matrix elements obtained from the HELAC-Oneloop package [61] that corresponds to the next-to-leading order (NLO) QCD accuracy. Powheg BOX [62, 63, 64] serves as an interface to the parton shower Monte Carlo programs. The samples created using this approach are referred to as PowHel samples. CT10NLO PDF sets are used and the factorization (μ_{F}) and renormalization (μ_{R}) scales are set to $\mu_0 = \mu_{\mathrm{F}} = \mu_{\mathrm{R}} = m_t + m_H/2$. Pile-up and the underlying events are simulated by Pythia 8.1 [65] with the CTEQ61L set of parton distribution functions and AU2 underlying event tune. The Higgs boson mass is set to 125 GeV and the top quark mass is set to 172.5 GeV.

The signal Monte Carlo samples are summarized in Table 6.1. These large samples are generated with inclusive Higgs boson decays with branching fractions set to the LHC Higgs Cross Section Working Group (Yellow Report) recommendation for $m_H = 125$ GeV [66]. The inclusive cross section (129.3 fb at $m_H = 125$ GeV) is also obtained from the Yellow Report [66].

6.2.2 Background Simulation

The background simulations used for this analysis are listed in Table 6.2. In general, the Alpgen[67], MadGraph[68], and AcerMC[69] samples use the CTEQ6L1[70] parton distribution function, while the Powheg[71], Sherpa[72], are generated with the CT10 PDF. The exception is the MadGraph $t\bar{t}t\bar{t}$ sample, which is generated with the MSTW2008 PDF[73]. The highest order calculations available are used for the cross sections.

Table 6.2: Monte Carlo samples used for background description. Unless otherwise specified MadGraph samples use Pythia 6 for showering and Alpgen samples use Herwig+Jimmy. $t\bar{t}$, single top, and Z+jets samples are replaced with data-driven estimates for the final result

Process	Generator	Detector simulation
$t\bar{t}W^{\pm}$,$t\bar{t}Z$	MadGraph	Full
tZ	MadGraph	AF2
$t\bar{t}t\bar{t}$	MadGraph	Full
$t\bar{t}W^{\pm}W^{\pm}$	Madgraph+Pythia8	AF2
$t\bar{t}$	Powheg+Pythia6	Full/AF2
single top tchan	AcerMC+Pythia6	Full
single top schan→l	Powheg+Pythia6	Full
single top $W^{\pm}t$	Powheg+Pythia6	Full
Wγ*	MadGraph	Full
Wγ+4p	Alpgen	Full
$W^{+}W^{-}$	Sherpa	Full
$W^{\pm}Z$	Sherpa	Full
Same-sign WW	Madgraph+Pythia8	AF2
ZZ	Powheg+Pythia8,gg2ZZ+Herwig	Full
Zγ*	Sherpa	Full
Z+jets	Sherpa	Full
ggF Higgs	Powheg+Pythia8	Full

Chapter 7

Object and Event Selection

The analysis is divided into 3 signal regions based on lepton counting: 2 same-sign leptons, 3 leptons and 4 leptons. The lepton counting occurs for fully identified leptons with full overlap removal with transverse momenta over 10 GeV to ensure orthogonality. Lepton selections are tightened afterward within each region.

The cuts for each signal region are provided in Table 7.1 and the object selections are detailed in the following selections. The selections are based on optimizations of the region sensitivity performed using MC (event for data driven backgrounds) and ad-hoc values for normalization systematic uncertainties[8] i All signal regions are comprised of three basic requirements: the presence of b-tagged jets, the presence of additional light jets, and a veto of same flavor opposite sign leptons with an invariant mass within the Z window. Additional requirements on the invariant mass of the leptons, the missing transverse energy in the event, and the total object energy (H_T) proved to have negligible additional benefit at our level of statistics. Figure 7.1 shows the background and signal fractions as a function number of jets and number of b-tagged jets for otherwise fully selected events.

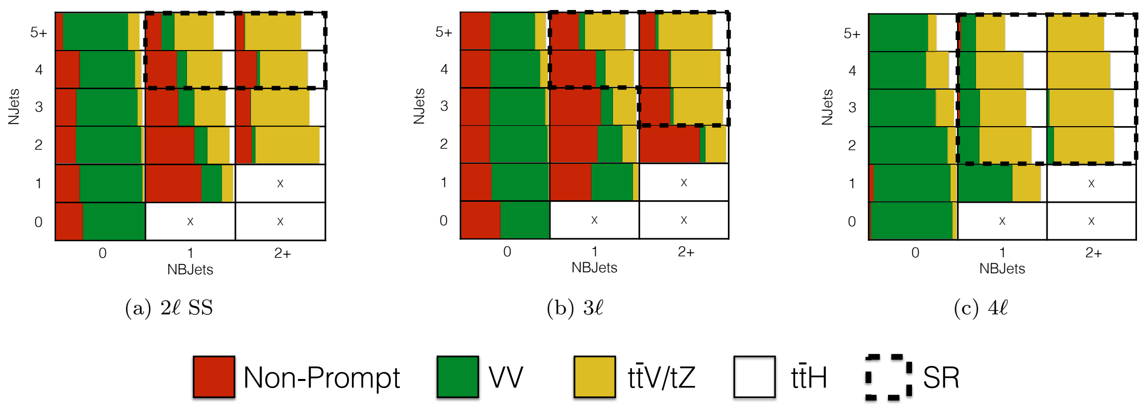
7.1 2ℓ Same-Charge Signal Region

The 2 lepton signal region requires two leptons of similar charge (2ℓ SS). The signal is symmetric in charge but the background from opposite-sign $t\bar{t}$ di-lepton production would be overwhelming. Requiring only two leptons allows the extra 2 W bosons in the event to decay hadronically, resulting in on average 4 additional light jets plus 2 additional b-quark jets from the top decays.

[8] the sensitivity was approximated using the $\frac{s}{\sqrt{b+\Delta b}}$ formula. The systematic errors considered were 20% for $t\bar{t}V$ and VV and 30% for fakes. These ended up being close the final systematic errors assessed in Chapter 9. The objects of optimization were the lepton momenta, identification operating points, isolation and event kinematic variables

Table 7.1: Selections in the 2ℓ SS, 3ℓ and 4ℓ Signal Regions

Signal Region	2ℓ SS	3ℓ	4ℓ
Trigger Matched Lepton	Yes	Yes	Yes
N_l[9]	=2	=3	=4
Lepton Charge Sum	+2 or −2	+1 or −1	0
Lepton Momentum (GeV)[10]	$p_{T0} > 25$ $p_{T1} > 20$	$p_{T0} > 10$ $p_{T1,2} >$25,20	$p_{T0} > 25$ $p_{T1} > 20$ $p_{T2,3} > 10$
Jet Counting	$N_b \geq 1$, $N_{Jet} = 4$	$N_b \geq 1$, $N_{Jet} \geq 4$ or $N_b \geq 2$, $N_{Jet} = 3$	$N_b \geq 1$, $N_{Jet} \geq 2$
Mass Variables (GeV)	$\lvert M_{ee} - M_Z \rvert < 10$	$\lvert M_{SFOS} - M_Z \rvert < 10$	$M_{SFOS} > 10$ $150 < M_{4\ell} < 500$ $\lvert M_{SFOS} - M_Z \rvert < 10$
Sub-channels	2 ($N_{Jet} = 4$, $N_{Jet} \geq 5$) x 3(ee,eμ,$\mu\mu$)	none	2 (No SFOS leps, SFOS leps)



(a) 2ℓ SS (b) 3ℓ (c) 4ℓ

Figure 7.1: Number of jets vs. number of b-tagged jet plot for the fully selected multi-lepton channels. Signal regions are outlined with a dashed line. Sub-channels are defined later in the 2ℓ SS and 4ℓ SRs. The fractional background contribution to each jet and b-tagged jet bin are shown for non-prompt (red), $t\bar{t}V+\ tZ$ (yellow), and VV (green). The expected signal fraction is shown in white. The expected non-prompt fraction contains charge misidentifications and fakes. It is shown for MC only, although data-based methods are used for the final result.

We require a leading lepton with transverse momentum of at least 25 GeV that matches to a trigger and a sub-leading lepton of at least 20 GeV, a b-tagged jet, and at least 4 jets in total.

In order to suppress non-prompt backgrounds, the lepton isolation criteria for tracking and calorimeter are tightened from less than 10% of the lepton momentum to 5%. To suppress charge misidentification, the electron is required to be extremely central ($|\eta| < 1.37$) to avoid the material-rich regions of the detector. Additionally, ee events with a lepton pair invariant mass within 10 GeV of the Z pole are removed. To maintain orthogonality with the τ analyses, events with fully identified taus are vetoed.

For the statistical combination the channel is divided into 6 sub-channels: 2 jets counting bins ($\mathrm{N}_{Jet} = 4$, $\mathrm{N}_{Jet} \geq 5$) x 3 lepton flavor bins (ee,$\mu\mu$,eμ).

7.2 3ℓ Signal Region

The 3 lepton channel requires 3 leptons, whose summed charge is either -1 or $+1$. The leptons are ordered in this way:

- **lep0:** the lepton that is opposite in charge to the other two leptons
- **lep1:** the lepton that is closer in ΔR to lep0
- **lep2:** the lepton that is farther in ΔR from lep1

Since events with a fake lepton arise exclusively from opposite sign di-lepton processes, $t\bar{t}$ and Z+jets, where additional jets are misidentified as the third lepton, lep0 is never the fake lepton. As a result, the transverse momentum requirement of lep0 ($>$ 10 GeV) is lower that the other two, $>$ 20 GeV. One lepton must must match a trigger and have $p_T > 25$ GeV.

The 3ℓ channel further requires at least one b-tagged jet and at least 4 jets in total, or two b-tagged jets and exactly 3 jets in total. Additionally, to suppress $W^{\pm}Z$ and Z+jets events, events with same-flavor opposite-sign (SFOS) pairs within 10 GeV of the Z pole are vetoed.

Additional cuts, including a di-lepton mass cut, and splittings were investigated but low statistics proved to wash out any advantages. The di-lepton mass cut will be a useful discriminator in future analyses since the spin statistics of Higgs decay in W bosons often causes the two emitted opposite-sign leptons to point in the same direction, resulting in a small measured invariant mass.

7.3 4ℓ Signal Region

In the four lepton signal region, selected events must have exactly four leptons with a total charge of zero. At least one lepton must be matched to one of the applied single lepton trigger and have a transverse momentum above 25 GeV. The leading and sub-leading leptons are required to have transverse momentum of 25 and 15 GeV respectively. In order to suppress background contributions from low-mass resonances and Drell-Yan radiation, all SFOS lepton pairs are required to have a dilepton invariant mass of at least 10 GeV.

The four-lepton invariant mass is required to be between 100 and 500 GeV. This choice of mass window suppresses background from the on-shell $Z \to 4\ell$ peak and exploits the high-mass differences between the signal and the dominant $t\bar{t}Z$ background. Events containing an OS-SF lepton pair within 10 GeV of the Z boson mass are discarded. This Z-veto procedure greatly reduces background contributions from ZZ and $t\bar{t}Z$. Finally, selected events are required to have at least two jets, at least one of which must be tagged as a b-quark jet.

The contribution from $t\bar{t}Z$ comprises approximately 75% of the total background in the inclusive signal region. A signal region categorization which factorizes $t\bar{t}Z$ from the remaining backgrounds is thus beneficial. The signal region is accordingly divided into two categories based on the presence of OS-SF lepton pairs in the final state.

7.4 Electron Selection

The electrons are reconstructed by a standard algorithm of the experiment [51] and the electron cluster is required to be fiducial to the barrel or endcap calorimeters: $|\eta_{cluster}| < 2.47$. Electrons in the transition region, $1.37 < |\eta_{cluster}| < 1.52$, are vetoed. Electrons must have p_{T}>10 GeV and pass the VERYTIGHT likelihood identification criteria.

In order to reject jets misidentified as electrons, electron candidates must also be well isolated from additional tracks and calorimeter energy around the electron cluster. Both the tracking and calorimeter energy within $\Delta R = 0.2$ of the electron cluster must be less than 5% of the electron transverse momentum: ptcone20/$\mathrm{P}_T < 0.05$ and Etcone20/$\mathrm{E}_T < 0.05$. All quality tracks with momentum greater than 400 MeV contribute to the isolation energy. Calorimeter isolation energy is calculated using topological clusters with corrections for energy leaked from the electron cluster. Pile-up and underlying event corrections are applied using a median ambient energy density correction.

The electron track must also match the primary vertex. The longitudinal projection of the track along the beam line, $z0 \sin\theta$, must be less than 1 cm) and the transverse projection divided by the parameter error, $d0$ significance, must be less than 4. These cuts are used in particular to suppress backgrounds from conversions, heavy-flavor jets and electron charge-misidentifications.

The electron selection is summarized in Table 7.2.

7.5 Muon Selection

Muons used in the analysis are formed by matching reconstructed inner detector tracks with either a complete track or a track-segment reconstructed in the muon spectrometer (MS), called Chain 2 muons. The muons have p_{T}>10 GeV and satisfy $|\eta| < 2.5$. The muon track are required to be a good quality combined fit of inner detector hits and muon spectrometer segments, unless the muon is not fiducial to the inner detector, $|\eta| > 2.47$. Muons with inner detector tracks are further required to pass standard inner detector track hit requirements [52].

As with electrons, muons are required to be isolated from additional tracking or calorimeter energy: ptcone20/P$_T$ < 0.1, Etcone20/E$_T$ < 0.1. A cell-based Etcone20/P$_T$ relative isolation variable is used. A pile-up energy subtraction based on the number of reconstructed vertices in the event is applied. The subtraction is derived from a Z boson control sample.

The muons must also originate from the primary vertex and have impact parameter requirements, $d0$ significance < 3, and $z0 \sin\theta < 0.1$ cm, similar to the electrons.

The muon selection is summarized in Table 7.2.

7.6 Jet and b-Tagged Jet Selection

Jets are reconstructed in the calorimeter using the anti-k_t [53] algorithm with a distance parameter of 0.4 using locally calibrated topologically clusters as input (LC Jets). Since the jets in the $t\bar{t}H$ signal mostly arise from the decay massive resonances and not radiation, they are expected to be central and high energy. Jets must have p_{T}>25 GeV and $|\eta| < 2.5$.

Jets must also pass loose quality requirement, ensuring the proper functioning of the calorimeter at the time of data taking. Jets near a hot calormeter cell in data periods B1/B2 are rejected. The local hadronic calibration is used for the jet energy scale, and ambient energy corrections are applied to account for energy due to pileup.

Jets within $|\eta| < 2.4$ and $p_T < 50$ GeV are further required to be associated with the primary vertex. The the fraction of track p_T associated with the jet that comes from the primary vertex, must exceed 0.5 (or there must be no track associated to the jet). This requirement rejects jets that arise from pile-up vertices.

B-jets are tagged using a Multi-Variate Analysis (MVA) method called MV1 and relying on information of the impact parameter and the reconstruction of the displaced vertex of the hadron decay inside the jet[54]. The output of the tagger is required to be above 0.8119 which corresponds to a 70% efficient Working Point (WP).

7.7 Object Summary and Overlap

Since many fully identified objects may be reconstructed as two different objects, an overlap removal procedure is applied. Electrons within $\Delta R < 0.1$ of muons are rejected in favor of the muon. Jets within $\Delta R < 0.3$ of electrons are then removed. Finally, muons within $\Delta R < 0.04 + 10GeV/p_T$ of jets are rejected, as these muons are thought to arise from jet fragmentation.

Parameter	Values	Remarks
		Electrons
p_{T}	> 10 GeV	
$\lvert\eta\rvert$	< 2.47 veto crack	< 1.37 for 2ℓ SS channel
ID	Very Tight Likelihood	
Isolation	$E_{\mathrm{T}}Cone20/p_{\mathrm{T}}$,$p_{\mathrm{T}}Cone20/p_{\mathrm{T}} < 0.05$	
Jet overlap removal	$\Delta R > 0.3$	
$\lvert d_0^{sig}\rvert$	$< 4\sigma$	
$z_0 sin\theta$	$< 1\ cm$	
		Muons
p_{T}	> 10 GeV	
$\lvert\eta\rvert$	< 2.5	
ID	Tight	
Jet overlap removal	$\Delta R > 0.04 + 10$ GeV$/p_T$	
Isolation	$E_{\mathrm{T}}Cone20/p_{\mathrm{T}}$,$p_{\mathrm{T}}Cone20/p_{\mathrm{T}} < 0.1$	< 0.05 for 2 leptons
$\lvert d_0^{sig}\rvert$	$< 3\sigma$	
z_0	$< 1\ cm$	
		Jets
p_{T}	> 25 GeV	
$\lvert\eta\rvert$	< 2.5	
JVF	JVF> 0.5 or no associated track or $p_{\mathrm{T}}> 50\ GeV$	
b-Tag	MV1 70% operating point	

Table 7.2: Object identification and selection used to define the 5 channels of the multi-lepton $t\bar{t}H$ analysis. Some channels use a sub-sample of objects as explained in the Remarks column.

Chapter 8

Background Estimation

The $t\bar{t}H$ multi-lepton signal regions discussed in Chapter 5 are contaminated by background contributions at a similar order of magnitude to the signal. The dominant background for each region is $t\bar{t}V$. Sub-dominant but important backgrounds include the production of vector boson pairs in associated with jets and b-quark jets (VV) and $t\bar{t}$ production with a jet misidentified as a lepton (fakes). The 2ℓ SS regions possesses a unique background of charge misidentification from Z and top events. The methods for estimating these backgrounds are discussed in this chapter. Monte Carlo simulation is used for the prompt $t\bar{t}V$and VV contributions. The non-prompt backgrounds from $t\bar{t}$ jet-misidentification and charge-misidentification are estimated using data-driven methods. Table 8.1 provides a summary of the $t\bar{t}H$ signal and background expectation for each of the signal regions, including the data-driven estimates discussed in this section.

8.1 Vector Boson ($W^{\pm}$, Z) production in association with top quarks: $t\bar{t}V$, tZ

Production of top quarks plus vector boson is an important background in all multi-lepton channels. A large part of the $t\bar{t}V$ component, arising from on-shell $Z \to \ell\ell$, can be removed via a Z mass veto on like-flavor, opposite sign leptons. However the $Z \to \tau\tau$ and γ^* components remain. The $t\bar{t}W^{\pm}$ and tZ processes generally require extra jets to reach the multiplicity of our signal regions, as such it is important to ascertain uncertainties associated with QCD radiation. We consider uncertainties on both the $t\bar{t}W^{\pm}$ and $t\bar{t}Z$ production cross-sections of these two processes and event selection efficiencies in the signal regions. The latter is sensitive to the NJet modelling in the MC. We assess the size of these uncertainties by investigating the effects of the choice of the factorization (μ_{F}) and renormalisation μ_{R} scales and PDF sets.

Table 8.1: Expected number of signal and background events in 2ℓ SS, 3ℓ and 4ℓ signal regions. For data-driven backgrounds, Monte Carlo only numbers are given for reference. The expected sensitivity $\frac{s}{\sqrt{s+b}}$ is provided for data-driven, MC only and for the inclusion of ad-hoc systematic uncertainties (20% for $t\bar{t}V$ and 30% for $t\bar{t}$ fake).

	Same-sign						3 leptons	4 leptons	
	≥ 5 jets			4 jets				Z enriched	Z depleted
	$e^\pm e^\pm$	$e^\pm \mu^\pm$	$\mu^\pm \mu^\pm$	$e^\pm e^\pm$	$e^\pm \mu^\pm$	$\mu^\pm \mu^\pm$			
$t\bar{t}H$	0.73 ± 0.03	2.13 ± 0.05	1.41 ± 0.04	0.44 ± 0.02	1.16 ± 0.03	0.74 ± 0.03	2.34 ± 0.04	0.19 ± 0.01	0.03 ± 0.00
$t\bar{t}V$	2.60 ± 0.13	7.42 ± 0.17	5.01 ± 0.16	3.05 ± 0.13	8.39 ± 0.24	5.79 ± 0.20	7.21 ± 0.24	0.74 ± 0.05	0.00 ± 0.00
tZ							0.71 ± 0.03	$i$ncl. in $t\bar{t}V$	$i$ncl. in $t\bar{t}V$
VV	0.48 ± 0.25	0.37 ± 0.23	0.68 ± 0.30	0.77 ± 0.27	1.93 ± 0.80	0.54 ± 0.30	0.89 ± 0.25	0.08 ± 0.01	0.00 ± 0.00
$t\bar{t}$, tX (MC)	1.31 ± 0.67	2.55 ± 0.84	1.76 ± 0.67	4.99 ± 1.19	8.19 ± 1.41	3.70 ± 1.03	2.46 ± 0.19	0.00 ± 0.00	0.00 ± 0.00
Z+jets (MC)	0.16 ± 0.16	0.28 ± 0.20	0.12 ± 0.12	1.37 ± 0.78	0	0.23 ± 0.23	0	0.00 ± 0.00	0.00 ± 0.00
fake leptons (DD)	2.31 ± 0.97	3.87 ± 1.01	1.24 ± 0.41	3.43 ± 1.38	6.82 ± 1.63	2.38 ± 0.78	2.62 ± 0.51	$(1.1 \pm 0.6) \cdot 10^{-3}$	$(0.09 \pm 0.03) \cdot 10^{-3}$
Q misid (DD)	1.10 ± 0.09	0.85 ± 0.08	–	1.82 ± 0.11	1.39 ± 0.08	–	–	–	–
Tot Background (fake MC)	4.56 ± 1.17	10.62 ± 1.54	7.57 ± 1.31	10.18 ± 2.43	18.51 ± 2.54	10.26 ± 1.82	11.27 ± 0.40	0.83 ± 0.07	0.01 ± 0.00
Tot Background (fake DD)	6.49 ± 1.04	12.51 ± 1.04	6.93 ± 0.52	9.07 ± 1.42	18.53 ± 1.83	8.71 ± 0.88	11.43 ± 0.62	0.831 ± 0.075	0.0110 ± 0.0003
$s/\sqrt{b}$ (fake MC)	0.34	0.65	0.51	0.14	0.27	0.23	0.70	0.21	0.30
$s/\sqrt{b}$ (fake DD)	0.29	0.60	0.54	0.15	0.27	0.25	0.69	0.21	0.29
$s/\sqrt{b} \oplus 0.3fake(MC) \oplus 0.2ttV$	0.33	0.58	0.47	0.12	0.22	0.21	0.63	0.207	0.30
$s/\sqrt{b} \oplus 0.3fake(DD) \oplus 0.2ttV$	0.27	0.53	0.50	0.14	0.23	0.23	0.62	0.207	0.286

Table 8.2: NLO cross section and theoretical uncertainty calculations derived from MadGraph5+aMC@NLO.

Process	σ_{NLO} $[fb]$	Scale Uncertainty [%]		PDF Uncertainty [%]		Total symmetrized uncertainty [%]
$t\bar{t}W^+$	144.9	+10	-11	+7.7	-8.7	13.3
$t\bar{t}W^-$	61.4	+11	-12	+6.3	-8.4	13.6
$t\bar{t}Z$	206.7	+9	-13	+8.0	-9.2	14.0
tZ	160.0	+4	-4	+7	-7	8.0
$\bar{t}Z$	76.0	+5	-4	+7	-7	8.6

Monte Carlo events for these processes are generated with MadGraph 5 and showered with Pythia 6. $t\bar{t}W^{\pm}$ events are generated with up to two extra partons at matrix element level, while for $t\bar{t}Z$ up to one extra parton at matrix-element level is produced. The tZ process is simulated without extra partons. The next-to-leading-order (NLO) cross sections are implemented by applying a uniform k-factor to the leading-order (LO) events for each process. For $t\bar{t}Z$, the k-factor is determined by comparing LO and NLO cross sections for on-shell Z production only and then applied to the off-shell signal regions.

The $t\bar{t}V$ uncertainties are calculated using the internal QCD scale and PDF re-weighting that is available with MadGraph5+aMC@NLO. The prescription for the scale envelope is taken from [74]: the central value $\mu = \mu_R = \mu_F = m_t + m_V/2$ and the uncertainty envelope is $[\mu_0/2, 2\mu_0]$. The PDF uncertainty prescription used is the recipe from [75]: calculate the PDF uncertainty using the `MSTW2008nlo` [73] PDF for the central value and then the final PDF uncertainty envelope is derived from three PDF error sets each with different α_S values (the central value and the upper and lower 90% CL values). The final NLO cross section central values and uncertainties are given in Table 8.2.

The tZ process is normalized to NLO based on the calculation in Ref. [76]. Here the scales are set to $\mu_0 = m_t$ and the scale variations are by a factor of four; the scale dependence is found to be quite small.

8.1.1 $t\bar{t}Z$ Validation Region

Unlike $t\bar{t}W^{\pm}$, a $t\bar{t}Z$ validation region can be obtained by simply inverting the veto on same-flavor opposite-sign (SFOS) lepton pairs near the Z pole in the 3 lepton signal region. This region thus requires 3 leptons (with momentum and identification cuts discussed in Chapter 7, at least one opposite sign, same-flavor pair of leptons within 10 GeV of the Z mass, and either 4 jets and at least 1 b-tagged jet or exactly 3 jet and 2 or more b-tagged jets. The resulting region has low statistics and is

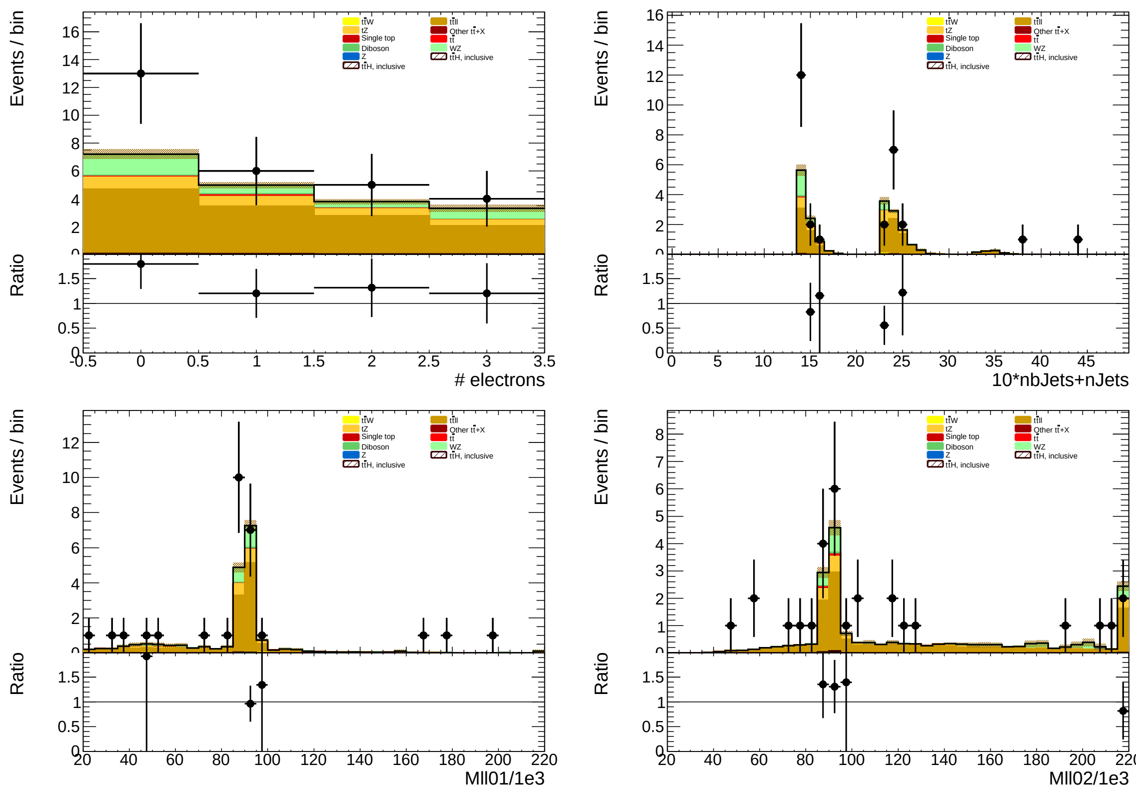


Figure 8.1: Data/MC comparison plots for $t\bar{t}Z$ control region A ($\geq$ 4 jets, $\geq$ 1 b-tag and 3 jets, $\geq$ 2 b-tag). In all plots, the rightmost bin contains any overflows. Top left: number of electrons. Top right: 10*the number of b-tags + the total number of jets. Middle left: the invariant mass of the (0,1) lepton pair (see the text for the definition of the lepton ordering). Middle right: the invariant mass of the (0,2) lepton pair.

not used as a control region but is instead used as a validation to demonstrate that the normalization uncertainty, discussed above, is properly evaluated.

The region defined by this is predicted to be 67% $t\bar{t}Z$, 17% $W^{\pm}Z$, and 13% tZ . We predict 19.3 ± 0.5 events and observe 28, giving a observed-to-predicted ratio of $1.45 \pm 0.27 \pm 0.03$, where the errors are from data and simulation statistics, respectively. Given the large errors, the region is still in agreement with the predictions to within 1-1.5 σ. Distributions of various variables are shown in Fig. 8.1.

8.2 Di-boson Background Estimation: $W^{\pm}Z$,ZZ

$W^{\pm}Z$ and ZZ di-boson production with additional and b-tagged jets constitute small contributions to the 3ℓ and 4ℓ channels. For the 3ℓ case $W^{\pm}Z$ comprises $\sim$ 10% of the total background, while for the 4ℓ case ZZ contribution accounts comprises $\sim$ 10% of the total background. Because of the small size of these contributions, each of the above processes can be assigned a non-aggressive uncertainty based on similar previous analyses with ATLAS and cross-checked with data validation regions and MC truth studies.

Both $W^{\pm}Z$ and ZZ production have been studied by ATLAS [77][78], but neither process has been investigated thoroughly in association with multiple jets and b-quark jets. However, single boson production with b-quark jets has been investiaged. Both $W + b$ [79] and $Z + b$ [80] production in 7 TeV data have been shown to agree with MC models to within 20-30%.

A single W produced in association with b-tagged jets possesses a similar topology to the $W^{\pm}Z+b$ process at a different energy scale and has been shown to be dominated by c mis-tags and b-jets from gluon splitting and multiple parton interaction. The $W + b$ analysis, referenced above, uses Alpgen MC with Herwig PS modeling, only provides results to 1 additional jet, and uses the CombNN tagger (we use MV1). Its results are therefore not directly comparable to this $t\bar{t}H$ analysis (where $W^{\pm}Z$ is modeled using Sherpa with massive c and b quarks). $Z+b$ production originates from slightly different diagrams than $ZZ+b$, but the sources of the b-tags are similar. The 7 TeV analysis, referenced above, provides results with Sherpa MC with an agreement of $\sim$ 30%. However, it also used the CombNN tagger instead of MV1. Beause of the differences of the 2011 single boson analyses (type of tagger used, type of MC and tunes used), we would like to verify the general 20-30% level of agreement in 2012 data with the simulation and tagger used in the $t\bar{t}H$ analysis: Sherpa MC, 2012 tunes, MV1. With the data skims available to use we are able to do this in the $Z + b$ region but not the $W + b$.

Figure 8.2 shows the spectrum of the number of reconstructed and selected jets (NJet) in a $Z + b$ validation region, defined by 2 tight-isolated leptons within 10 GeVof the Z mass and with at least one b-tagged jet, using the $t\bar{t}H$ analysis definitions. The level of agreement in this region confirms at the 30% level seen in the 7 TeV analysis, discussed above.

In the following two sections, we assess the truth origin of jets in the $W^{\pm}Z+b$ and $ZZ+b$ regions and leverage data/MC agreement where we can. We see that the data allows us to constrain the $W^{\pm}Z$ to 50%. We claim this 50% as a systematic. The 20-30% agreement in the single boson regions above bolsters our confidence in this number.

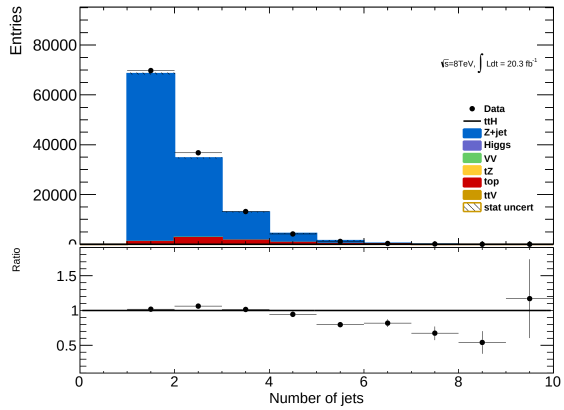


Figure 8.2: NJet spectrum for 2 tight-isolation leptons with 1 b-tagged jet (MV1_70)

8.2.1 $W^{\pm}Z$ Normalization Uncertainty

The $t\bar{t}H$ analyses has two validation regions to test the Sherpa agreement with data for $W^{\pm}Z$: one inclusive 3 lepton region, using the three-lepton channel object and p_{T} cuts and a $W^{\pm}Z + b$ region with 1 b-tagged jet and a requirement that at least one SFOS pair have an invariant mass within 10 GeV of the Z mass. The region with fewer than 4 jets is $W^{\pm}Z$ dominated. Figure 8.3 shows kinematic variables for the inclusive region. The overall data normalization is $\sim$10% higher than MC, but this will be well within our systematic uncertainty. The NJet shape shows good agreement across the full spectrum, giving confidence about the Sherpa high NJet SR extrapolation. Figure 8.4 shows NJet spectrum for the $W^{\pm}Z + b$ validation region with agreement with in statistical uncertainties. The region has low statistics and around $\sim$ 60% purity and statistical analysis of the region suggests that a 50% normalization error on the $W^{\pm}Z$ component is enough to cover any possible mismodelings, especially in higher NJet bins, which are closer to the signal regions.

We also examine the $W^{\pm}Z$ truth origins of the b-jet in the $W^{\pm}Z + b$ validation region (VR) and the signal region using MC to assess the validity of the extrapolation from the VR to the SR and to confirm the similarity in jet origin to the single boson analyses, references above. The flavour of the closest matching truth particle ($p_T > 5$ GeV, after FSR) in ΔR determines the true-jet flavor. If there are no quarks, taus or gluons within ΔR of 0.3, the label defaults to light. Table 8.3 shows the origin fraction of b-tagged jets in the various $W^{\pm}Z + b$ VRs and the SR. If there are two b-tagged jets, the highest p_{T} is used, but this is a small fraction of the number of b-tags. As expected the c and b contributions dominate, as was the case with the 2011 single boson analyses referenced above. It is important also that the VR has similar composition to the SR. There is a small dependence on

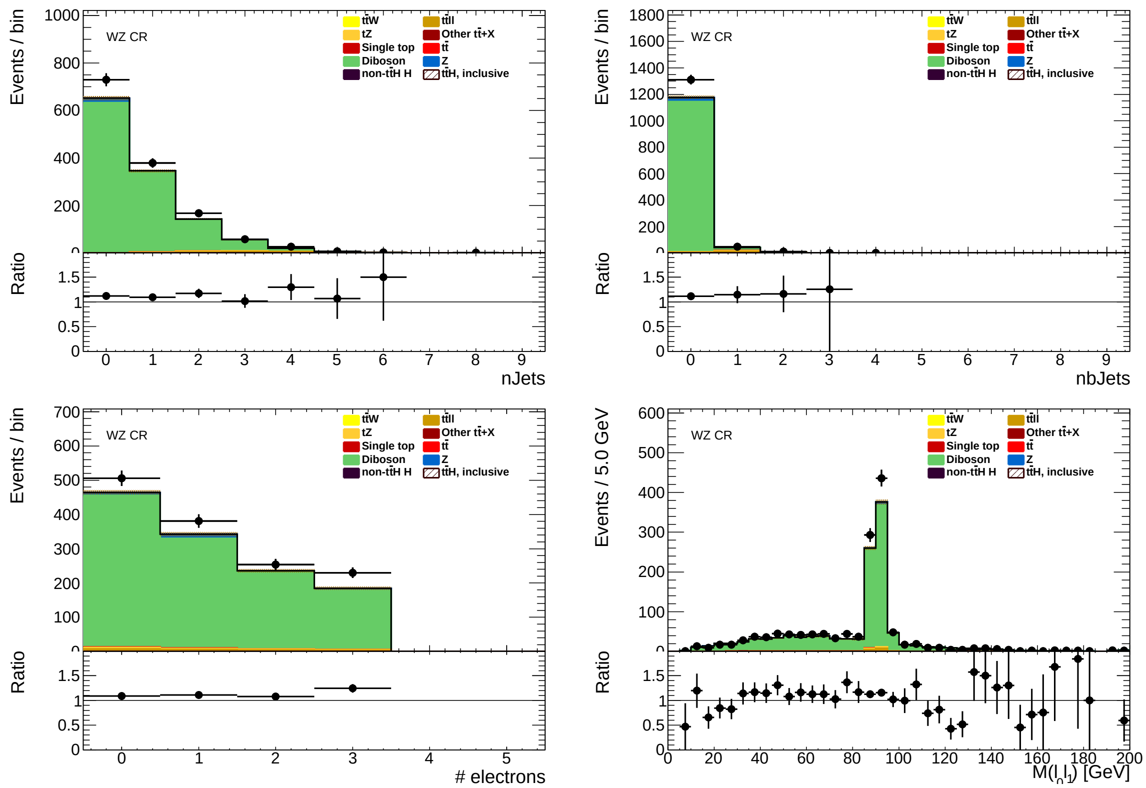


Figure 8.3: 3 lepton $W^{\pm}Z$ validation using the $t\bar{t}H$ lepton identification and momentum cuts: mass, number of jet and flavor variables

the number of jets.

	Bottom	Charm	Light
$W^{\pm}Z+b$ VR 1 Jet	0.25 ± 0.03	0.54 ± 0.04	0.20 ± 0.03
$W^{\pm}Z+b$ VR 2 Jet	0.34 ± 0.04	0.52 ± 0.06	0.13 ± 0.03
$W^{\pm}Z+b$ VR 3 Jet	0.40 ± 0.07	0.41 ± 0.07	0.18 ± 0.04
$3l$ SR	0.43 ± 0.14	0.38 ± 0.17	0.18 ± 0.11

Table 8.3: Truth origin of highest energy b-tagged jet in the $W^{\pm}Z+b$ VR and $3l$ SR

8.2.2 ZZ Normalization Uncertainty

In order to investigate the MC agreement with data in the ZZ case, two validation regions similar to the $W^{\pm}Z$ case are defined. First, a 4 lepton ZZ region is constructed using the object selections for the 4-lepton channel and requiring exactly two pairs of opposite sign-same flavor leptons with a di-lepton invariant mass within 10 GeV of the Z mass. Additionally, the $ZZ+b$ process is investigated directly using a similar validation region which again requires exactly two Z-candidate lepton pairs

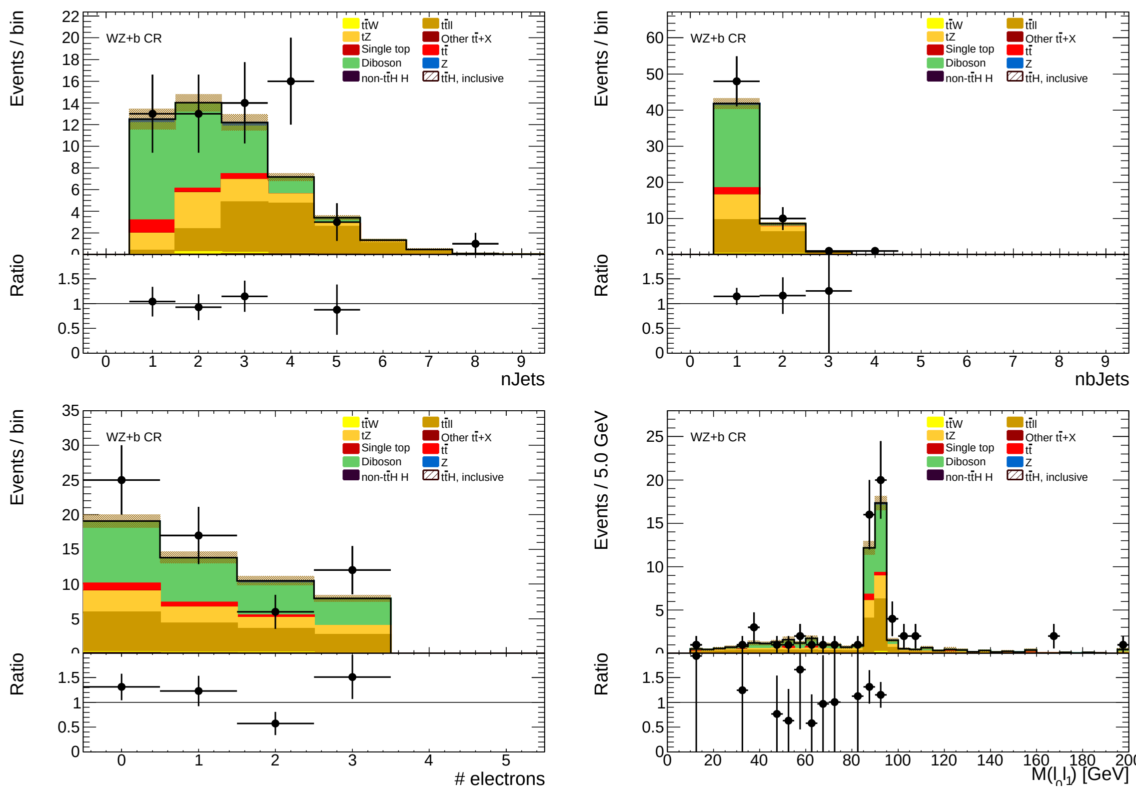


Figure 8.4: $W^{\pm}Z + b$ validation region: NJet, NElec, and Mass Variables

as well as at least 1 b-tagged jet. Some kinematic distributions are shown in Figures 8.5 and 8.6, and particular attention should be paid to the NJet spectrum, which shows good data-MC agreement in the high-jet bins, with a slight discrepancy in the 1-jet bin. The agreement for the region with at least 2 jets yields confidence in the NJet MC modeling in this region which lies close to the 4-lepton signal region.

Based on the study of the ZZ and $ZZ+b$ validation regions and the overall agreement noted with the $Z+b$ analysis, we expect a similar error to $W^{\pm}Z$ to be appropriate in the ZZ case. A truth origin study is undertaken in MC to demonstrate a similar b-jet origin to the $W^{\pm}Z$ case. The true origin of the leading (highest energy) b-tagged jet is shown in Table 8.4 for the 4-lepton signal region as well as the $ZZ + b$ validation region described above divided into jet bins. As it was in the $W^{\pm}Z$ case above, the true origin of the b-jet in $ZZ + b$ is dominated by c and b. Taking this study in tandem with the results from the $W^{\pm}Z$ investigation, it is appropriate to take the central value of the $ZZ+b$ background contribution in the 4-lepton SR from MC and to assign an overall systematic of 50% in order to account for the MC modeling limitations.

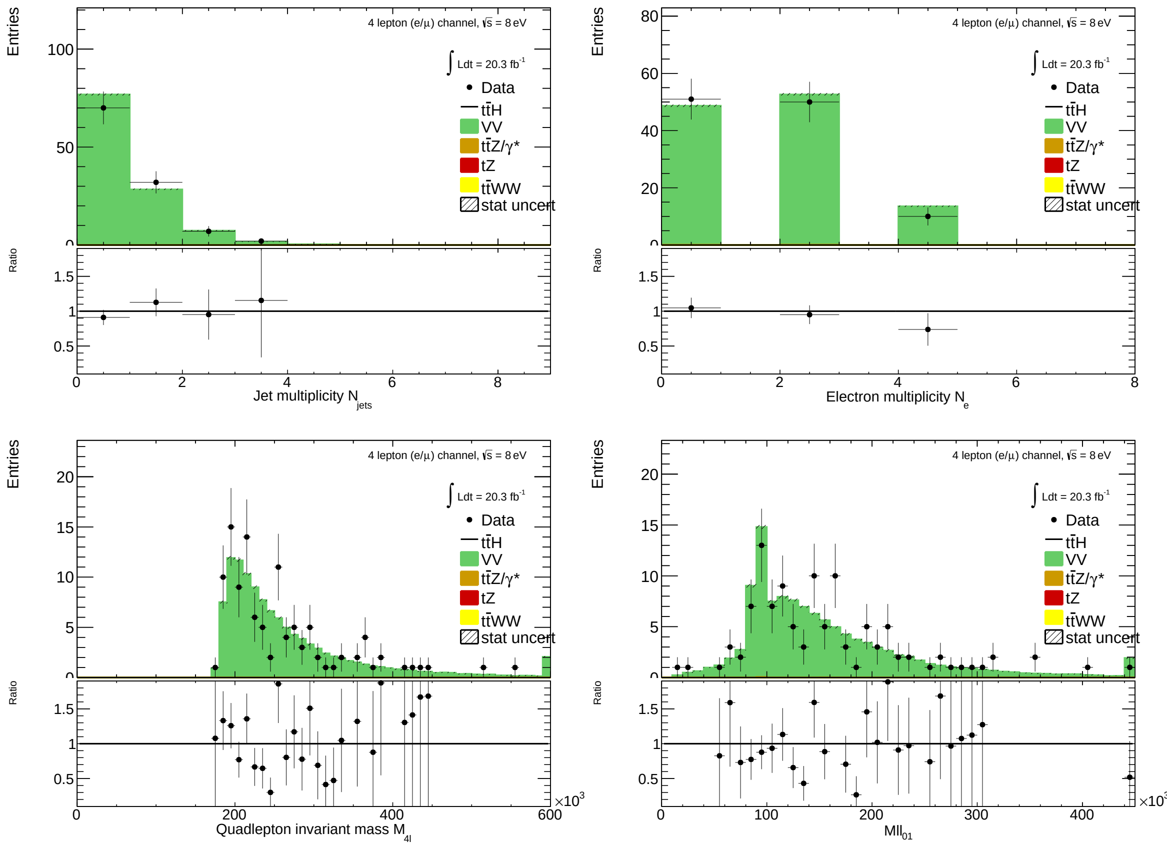


Figure 8.5: Jet-inclusive 4-lepton ZZ validation region using the $t\bar{t}H$ lepton identification and momentum cuts

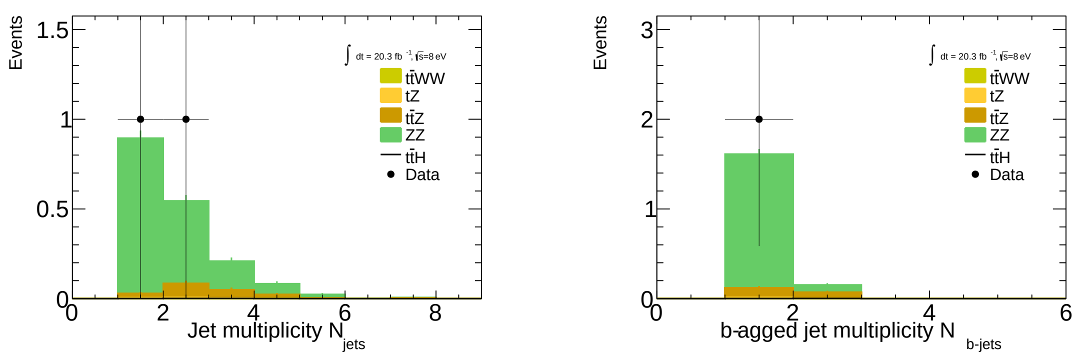


Figure 8.6: $ZZ + b$ validation region using the $t\bar{t}H$ lepton identification and momentum cuts

	Bottom	Charm	Light
$ZZ+b$ VR 1 Jet	0.56 ± 0.03	0.24 ± 0.01	0.20 ± 0.01
$ZZ+b$ VR 2 Jet	0.52 ± 0.05	0.25 ± 0.02	0.23 ± 0.02
$ZZ+b$ VR 3 Jet	0.53 ± 0.11	0.25 ± 0.08	0.22 ± 0.07
$4l$ SR	0.34 ± 0.15	0.42 ± 0.16	0.24 ± 0.10

Table 8.4: Truth origin of highest energy b-tagged jet in the $ZZ+b$ VR and $4l$ SR

8.3 Charge-Misidentification Background

Charge-misidentification contributes to the background for 2ℓ SS case and only for flavor channels, which include electrons. The same-sign requirement is essential in removing large SM opposite sign-backgrounds, but because of their size even small charge misidentification rates result in contamination in same-sign regions. For the 2ℓ SS signal regions, charge-misidentification background arise primarily from $t\bar{t}$ di-lepton events with a smaller contribution from leptonic Z decays.

In general, charge-misidentification can arise in two ways. The first occurs for ultra-high energy electrons and muons, which leave tracks in the detector that are too straight for the fit to determine the direction of curvature with high confidence. This type of charge misidentification is not a concern to the $t\bar{t}H$ multi-lepton analysis, as most of the leptons have transverse momentum < 150 GeV. The second source of charge misidentification is from 'tridents', which only occurs for electrons, because their low mass allows for high rate bremsstrahlung in the detector material. In some cases, after an electron releases a photon through bremsstrahlung, the photon may convert nearby resulting in three electron tracks. The reconstruction algorithms may sometimes match the wrong track to the calorimeter energy deposit, resulting in a possible charge misidentification. As discussed in the selection, tight track-cluster geometric and energy matching requirements are applied on the electron candidates to reduce the overall rate and the electron acceptance is narrowed to ($|\eta| < 1.37$), since most of the material is concentrated more forward in the detector.

We estimate the contribution of charge-misidentification events in our 2ℓ SS signal regions and relevant control regions by applying a weight per electron in the OS region with otherwise identical cuts. The weight is related to the charge-misidentification rates. We measure these rates using a likelihood method in the OS and SS $Z \to ee$ control region in data. The rate measured from these control regions is binned in electron p_{T} and η, to account for dependencies in these variables. The method, validations and associated errors are discussed in detail in the following sub-sections.

8.3.1 Likelihood Method

The number of reconstructed same-sign (N_{ss}) $Z \to ee$ events is related to total number of produced $Z \to ee$ (N) through factors related to the charge misidentification rate, ϵ:

$$N_{ss}^{ij} = N^{ij}(\epsilon_i + \epsilon_j - 2\epsilon_j\epsilon_i) \tag{8.1}$$

where ϵ_i and ϵ_j are the charge misidentification rates for each electron separately. If we drop terms quardratic in ϵ, we have:

$$N_{ss}^{ij} = N^{ij}(\epsilon_i + \epsilon_j). \tag{8.2}$$

Although it is impossible to know event-by-event which electron's charge was misidentified, we can use a likelihood method over the whole Z sample to measure how ϵ depends on the electron p_{T} and $|\eta|$. As illustration, we first consider the case, where ϵ depends on only one variable, $|\eta|$, and then generalize to the two-dimensional case of $|\eta|$ vs p_{T}.

N_{ss}^{ij} is described by a Poisson distribution:

$$f(k,\lambda) = \frac{\lambda^k e^{-\lambda}}{k!}, \tag{8.3}$$

where k is the observed number of occurrences of the event, i.e. $k = N_{ss}^{ij}$, and λ is the expected number, i.e. $\lambda = N^{ij}(\epsilon_i + \epsilon_j)$. Thus, the probability for an observed number of same-sign Z events given the sample size and charge misidentification rates is expressed by:

$$P(N_{ss}^{ij}|N^{ij}, \epsilon_i, \epsilon_j) = \frac{[N^{ij}(\epsilon_i + \epsilon_j)]^{N_{ss}^{ij}} e^{-N^{ij}(\epsilon_i+\epsilon_j)}}{N_{ss}^{ij}!}. \tag{8.4}$$

The likelihood L for all the events is obtained by evaluating all the $|\eta|$ combinations:

$$L(\epsilon|N_{ss}, N) = \prod_{i,j} \frac{[N^{ij}(\epsilon_i + \epsilon_j)]^{N_{ss}^{ij}} e^{-N^{ij}(\epsilon_i+\epsilon_j)}}{N_{ss}^{ij}!}. \tag{8.5}$$

In this process, the $-\ln L$ is used in order to simplify and make easier the minimization. Terms which do not depend on the rates ϵ_i and ϵ_j are removed in this step. This way, the final function to minimize is given by the following expression:

$$-\ln L(\epsilon|N_{ss}, N) \approx \sum_{i,j} \ln[N^{ij}(\epsilon_i + \epsilon_j)] N_{ss}^{ij} - N^{ij}(\epsilon_i + \epsilon_j). \tag{8.6}$$

The likelihood can be easily extended to depend on the charge misidentification rates as a function of two parameters. The probability to find a same-sign event given the rates for each electron is $(\epsilon_{i,k} + \epsilon_{j,l})$, where the two indices represent binned $|\eta|$- and p_{T}-dependence. Thus, the Eq. 8.6

transforms into

$$-\ln L(\epsilon|N_{ss}, N) \approx \sum_{i,j,k,l} \ln[N^{ij,kl}(\epsilon_{i,k} + \epsilon_{j,l})]N_{ss}^{ij,kl} - N^{ij,kl}(\epsilon_{i,k} + \epsilon_{j,l}). \tag{8.7}$$

We use events selected within the Z peak using the $t\bar{t}H$ electron object cuts. The events are stored in two matrices: one for the same-sign events $N_{ss}^{ij,kl}$, and the other one for all events $N^{ij,kl}$. Small backgrounds need to be subtracted. The background subtraction is done using a simple side-band method. This method consists in dividing the Z invariant mass in three regions, i.e. A, B and C, where B is the central region corresponding to the Z peak. The number of events is counted in the regions on the sides of the peak, i.e. n_A and n_C, and removed from the total number of events in the peak region B, n_B. This way, the number of signal events N_Z is given by

$$N_Z = n_B - \frac{n_A + n_C}{2}. \tag{8.8}$$

Once the background has been subtracted, the likelihood is minimized for the 2D matrix of ϵ bins. Knowing ϵ as a function of $|\eta|$ and p_{T} for any single electron, it is now possible to estimate the number of same-sign events from the number of opposite sign events in any sample:

- $N^{ss} = \frac{\epsilon_i + \epsilon_j - 2\epsilon_i\epsilon_j}{1 - \epsilon_i - \epsilon_j + 2\epsilon_i\epsilon_j} N^{os}$ for ee channels
- $N^{ss} = \frac{\epsilon}{1-\epsilon} N^{os}$ for the $e\mu$ channels

8.3.2 Results

The charge misidentification rate is calculated in 7 $|\eta|$ bins [0.0, 0.6, 1.1, 1.37, 1.52, 1.7, 2.3, 2.47] by 4 p_{T}bins [15, 60, 90, 130, 1000] GeV. For p_{T} bins above 130 GeV, the Z dataset becomes too small and the rates are calculated using $t\bar{t}$ MC, scaled by the data-MC ratio of the rates in the lower p_{T} bins, [90-130] GeV. Figure 8.7 shows the extracted rates in all bins.

As a cross-check, we apply the full method to the Z MC samples (extracting rates via a likelihood fit and applying them to opposite sign events) and compare to the MC predicted number of same-sign events. The invariant mass of the Z from our charge misidentification and directly from the MC can be seen on Figure 8.8. In the simulated Z samples, the number of same-sign Z events is 5 049 while the estimation is 5 031^{+375}_{-365}. The uncertainties combine both statistical systematic uncertainties, which are discussed in depth below. The validation gives compatible results within uncertainties.

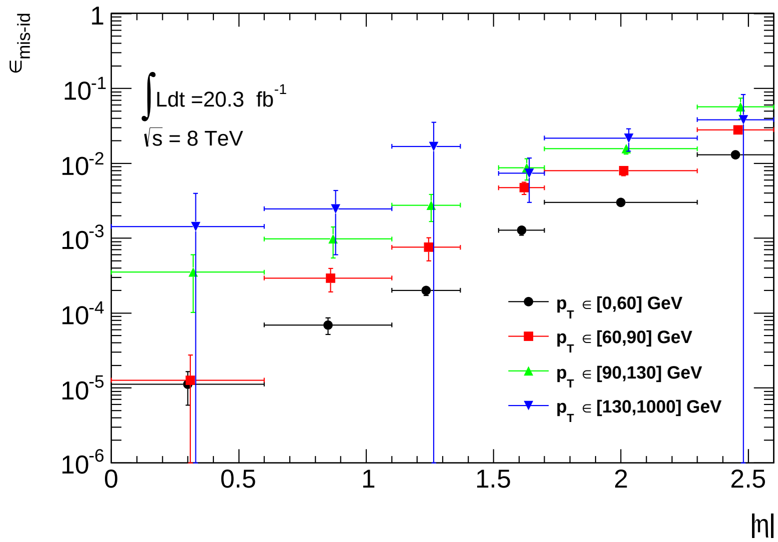


Figure 8.7: Electron charge misidentification rates measured in data with the likelihood method on Z events (black points, red squares and blue triangles) as a function of $|\eta|$ and parametrized in p_{T}. The full 2012 dataset has been used to estimate the rates below 130 GeV. Above this value, the charge misidentification rates have been estimated by extrapolating the rates in the region where the $p_{\mathrm{T}} \in [90, 130]$ GeV with a p_{T} dependent factor extracted from simulated $t\bar{t}$ events (green triangles). Statistical and systematic uncertainties have been included in this plot.

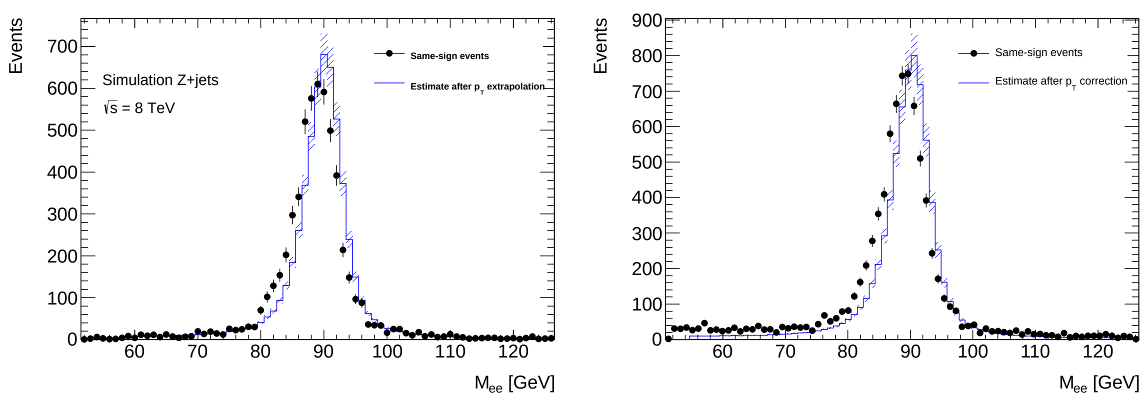


Figure 8.8: Closure test on simulated $Z \to e^+e^-$ events (a) and data (b). The black circles show the distribution of same-sign events while the blue histograms show the distribution of the re-weighted opposite-sign events together with the statistical and systematic uncertainties. The distributions are not expected to overlay exactly, due to the loss of energy of the trident electrons for the same-sign peak.

8.3.3 Systematic and Statistical Uncertainties

Statistical uncertainties dominate the combined uncertainty on the charge misidentification estimate. The statistical uncertainties come primarily from the size of the Z same-sign sample in data and are especially large for central, material-poor regions where the charge misidentification rate is extremely low. Additional systematic uncertainties are included for a comparison between the positron and electron rate, the per-bin MC closure test discussed above, and for the effect of varying the invariant mass window used for the background subtraction for three different cases. Figure 8.9 shows the relative uncertainties for all rate bins.

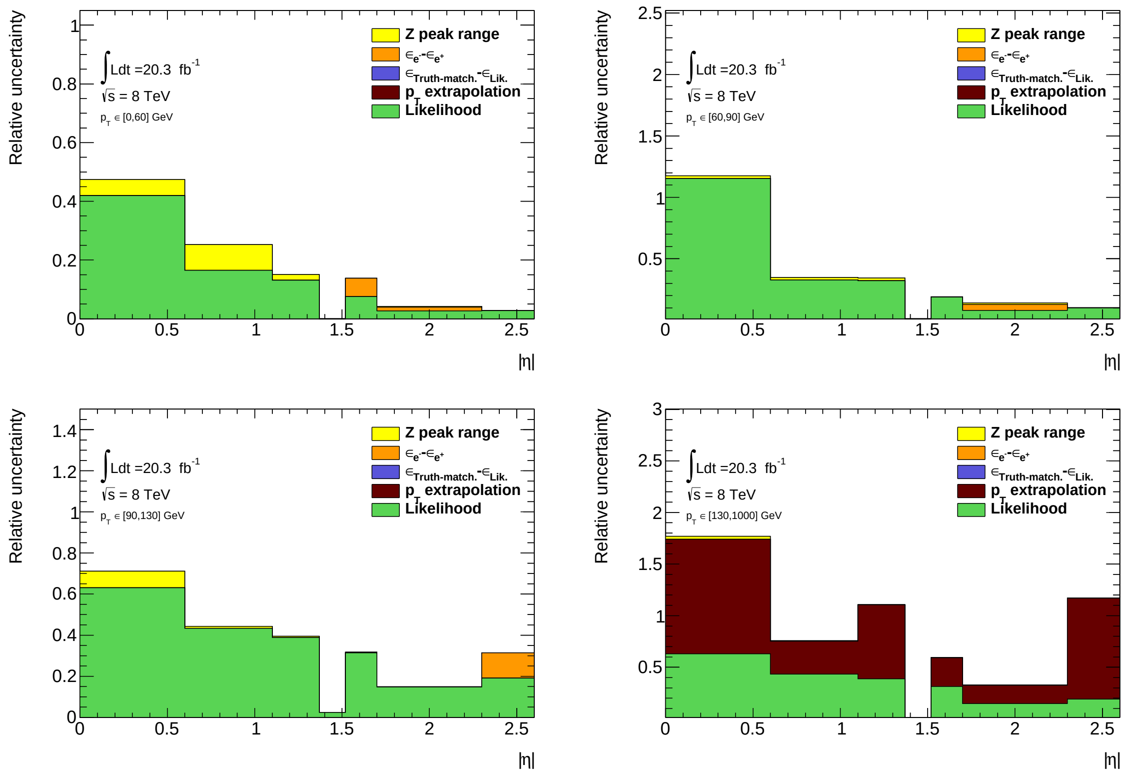


Figure 8.9: Relative systematic uncertainty contributions on the charge misidentification rate, for different bins in p_{T} and $|\eta|$.

We apply the rates to estimate the charge misidentification background in the 2ℓ SS signal regions, and find $\sim 25\%$ contamination in the $e^{\pm}e^{\pm}$ regions and a $\sim 10\%$ contribution to the $e^{\pm}\mu^{\pm}$ regions with a 10% systematic error overall. The low overall error can be attributed to the fact that the statistical error is lowest where the bulk of charge misidentifications occur. The charge flip contribution measured in the signal regions from this method is detailed in Table 8.1.

8.4 Fake Lepton Backgrounds

Fake Leptons, from the misidentification of jets as either electrons or muons, primarily arise from $t\bar{t}$ and single top processes in the 2ℓ SS, 3ℓ and 4ℓ channels. Smaller contributions come from Z+jet events. Fake backgrounds are sub-dominant but important in the 2ℓ SS and 3ℓ channels. They are extremely small in the 4ℓ channels. Truth studies suggest that these misidentified leptons arise overwhelmingly from b-quark initiated jets. The general method for estimating fakes in all channels is to define a reversed object selection control region (usually isolation) for each lepton flavor with otherwise identical signal region selection (N^{e}_{CR}, N^{μ}_{CR}). This region is fake-dominated. The total number of fake events in these regions are then scaled by transfer factors (θ) to estimate the number of fake events of the appropriate flavor in the signal region. The simple formula for determining fakes is defined in Equation 8.9.

$$N_{fake} = \theta_e \cdot N^{e}_{CR} + \theta_\mu \cdot N^{\mu}_{CR} \tag{8.9}$$

This approach factorizes the background model into two separate measurements. N_{CR} is sensitive the overall $t\bar{t}$ production rate, especially in the presence of additional jets from QCD ratio, as well as the object-level misidentification of a jet as a lepton. The transfer factors are sensitive to only the object level properties of the misidentified jet, and in particular only the variables which are reversed in the anti-tight identification.

The transfer factors are obtained obtained in a different way for each channel, due to unique issues with statistics and contamination, but each method relies heavily on the data-based control regions with fewer jets. Figure 8.10 shows a truth study of the stability of the transfer factor for the 2ℓ SS and 3ℓ cases as a function of the number of jets in the event for events with one-b-tagged jet. This suggest that the regions with fewer jets are a good model of the fakes in the signal regions with more jets and is expected because of the homogeneity of origin of the fakes across all jet bins.

The details of the methods for each channel are discussed in depth in the following sections. For all methods, the overall systematic uncertainty on the normalization of the fake estimate is in the range 30%-50% and arise primarily from statistics and the closure on assumptions used to obtain the transfer factor.

Because these methods do not provide a per-object transfer-factor that depends on the properties of the faking object, we must use the MC to model the shapes of the fake kinematic distributions in the signal regions. This is not an essential issue, since the analysis only considers only the total number of events in each signal region in the final measurement of $t\bar{t}H$ production.

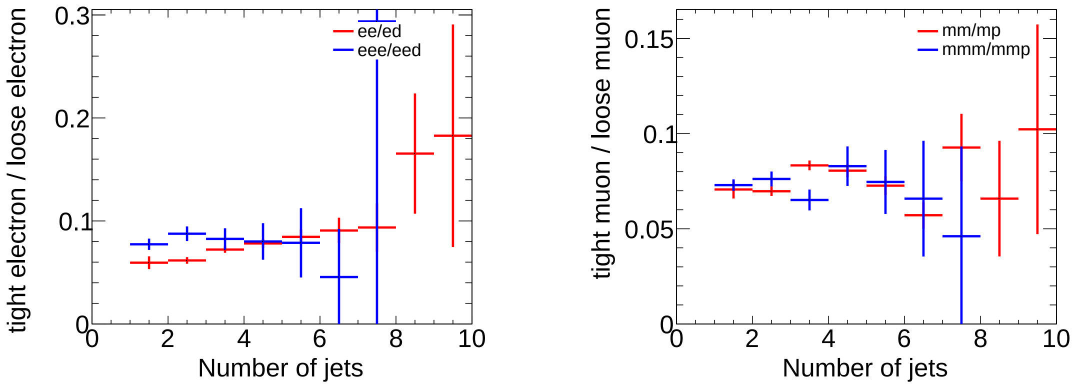


Figure 8.10: Ratios of regions with tight and anti-tight leptons in 2-lepton and 3-lepton channels from $t\bar{t}$ MC. These ratios are the MC calculated transfer factors for each region, i.e. $\theta_e = eee/eed$, ee/ed and θ_μ $=mmp/mmm$, mm/mp, where 'd' refers to anti-tight electrons and 'p' referes to anti-tight muons. The transfer factors are seen to be similar in the 2ℓ and 3ℓ cases and stable as a function of the number of jets

8.4.1 2ℓ SS Fakes

The 2ℓ SS fake method follows the procedure outlined in general above. We define anti-tight electron and muon control regions with reversed particle identification criteria for each signal region, including the 6 flavor and jet-counting sub regions. The anti-tight muon and electron criteria are provided below:

- **anti-tight electron (d):** fails the verytight likelihood operating point, but still passes the veryloose operating point. fails relative tracking and calorimeter isolation, $\mathrm{E}_T^{rel} > 0.05$ and $\mathrm{p}_T^{rel} > 0.05$.
- **anti-tight muon (p):** 6 GeV $< p_\mathrm{T} <$ 10 GeV

The electron and muon transfer factors, θ_e and θ_μ, are calculated in the region with signal region selection but fewer jets, $NJet == 2$ or $NJet == 3$ and are defined as the ratio of the number of events for two fully identified leptons to the number of events with one fully identified lepton and one anti-identified lepton, after the prompt and charge misidentification backgrounds are subtracted. Only same-flavor channels are used to ensure that muon and electron transfer factors maybe estimated separately: on every region, the prompt and charge-misidentification backgrounds are subtracted from the data.

$$\theta_e = \frac{N_{ee}}{N_{ed}} = \frac{N_{ee}^{Data} - N_{ee}^{\mathrm{Prompt}\ SS} - N_{ee}^{\mathrm{QMisId}}}{N_{ed}^{Data} - N_{ed}^{\mathrm{Prompt}\ SS} - N_{ed}^{\mathrm{QMisId}\ MC}} \tag{8.10}$$

Process	N(events)
$ed \leq 3$ jets	
VV	7.13 ± 0.63
$V\gamma$	7.55 ± 1.27
$t\bar{t}V, tV$	6.68 ± 0.18
$V + jets$	59.4 ± 18.51
$t\bar{t}, t + X$	671.26 ± 12.76
$t\bar{t}$ prompts	32.97 ± 2.83
Total MC	752.0 ± 22.5
Data	967
Data fakes (Data - prompts)	912.66
$ee \leq 3$ jets	
VV	3.30 ± 0.42
$V\gamma$	1.31 ± 0.65
$t\bar{t}V, tV$	3.96 ± 0.16
$V + jets$	8.3 ± 8.8
$t\bar{t}, t + X$	11.65 ± 1.67
Charge misID	8.54 ± 0.23
Total MC	28.52 ± 8.96
Data	32
Data fakes (Data - prompts)	14.26

Process	N(events)
$\mu p \leq 3$ jets	
VV	4.34 ± 0.44
$V\gamma$	4.84 ± 2.05
$t\bar{t}V, tV$	0.74 ± 0.06
$V + jets$	21.77 ± 12.24
$t\bar{t}, t + X$	192.71 ± 6.59
Total MC	224.4 ± 14.1
Data	249
Data fakes (Data - prompts)	239.07
$\mu\mu \leq 3$ jets	
VV	6.27 ± 0.56
$V\gamma$	0.06 ± 0.25
$t\bar{t}V, tV$	10.00 ± 0.27
$V + jets$	1.22 ± 11.78
$t\bar{t}, t + X$	15.4 ± 2.1
Total MC	32.95 ± 2.21
Data	44
Data fakes (Data - prompts)	27.65

Table 8.5: Number of events of the main simulated background processes and of the data in the $e^{\pm}e^{\pm}$ and $\mu^{\pm}\mu^{\pm}$ channels used for the measurement of θ_e and θ_μ. VV, $V\gamma$, $t\bar{t}V, tV$ and $t\bar{t}$ prompts (or charge misID) are the backgrounds which lead to prompt same-sign dileptons and are subtracted from the data to get a measured number of fakes. Uncertainties are statistical. The numbers labeled Data fakes are used to measure θ.

$$\theta_\mu = \frac{N_{\mu\mu}}{N_{\mu p}} = \frac{N_{\mu\mu}^{Data} - N_{\mu\mu}^{\text{Prompt}\ SS}}{N_{\mu p}^{Data} - N_{\mu p}^{\text{Prompt}\ SS}} \tag{8.11}$$

The 2,3 jet anti-tight regions used in obtaining the transfer factors are shown in Table 8.5 and the 4,5 jet anti-tight regions, scaled by the transfer factors to get the fake estimates in the SR are shown in Figure 8.11. The $t\bar{t}$ and single top MC are included in the plots and tables for reference, although they are not used in the measurements.

The transfer factors obtained from the 2 and 3 jet regions are shown in 8.6 with statistical errors and propagated systematic errors from the subtraction of relevant backgrounds ($t\bar{t}V$ and charge misidentification). The MC values are just for comparison. An additional systematic error is added by comparing the transfer factors, obtained from the low jet control region, to those obtained from the higher jet signal regions, using $t\bar{t}$ MC. The value of this systematic is about 20 % and can be seen

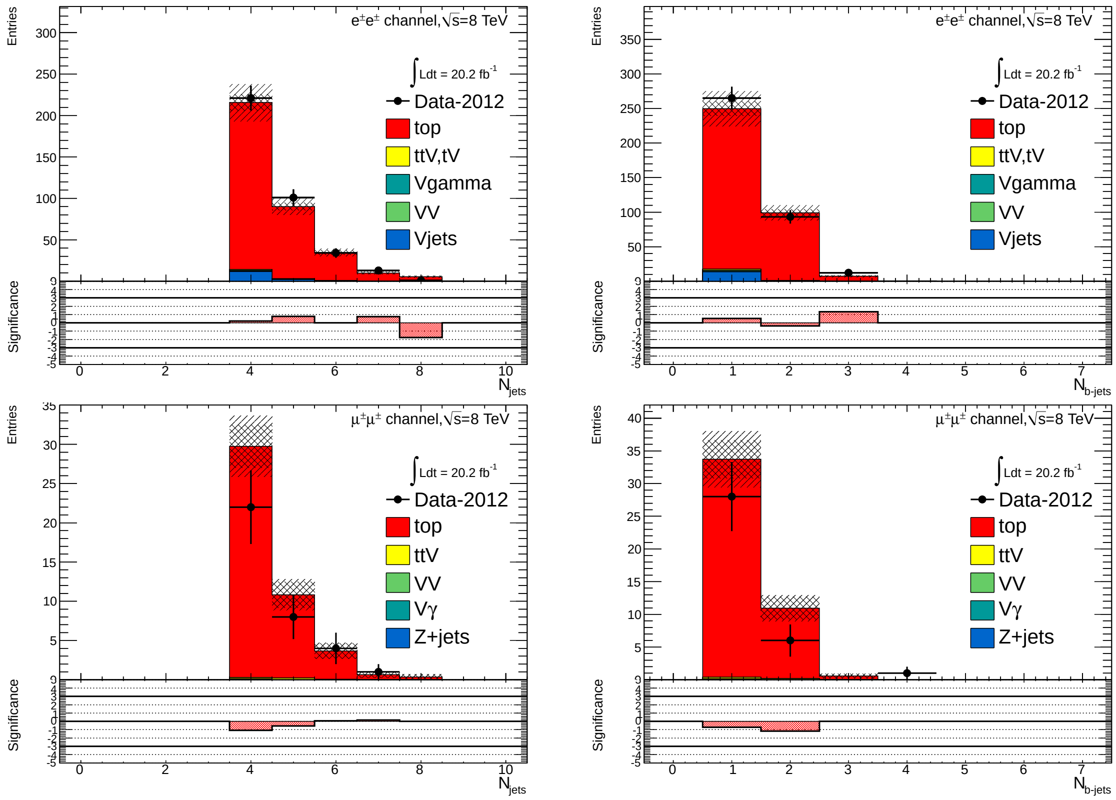


Figure 8.11: 4,5 Jet SS 2ℓ *ed* (above) and μp control regions with at least one b-tagged jet. After subtraction of prompt and charge misidentification backgrounds, these regions are scaled by the transfer factors, θ_e and θ_μ, to obtain the final number of fake events in the CR. The top MC (red) is used for reference but not in the actual calculation.

Factor	Expected (MC)	Measured (data)
θ_e Rev. Id.	0.0136 ± 0.0062	*0.0156±0.0062*
θ_μ Rev. p_{T}	0.078 ± 0.012	*0.1156±0.0288*

Table 8.6: Expected and measured values of the θ factors. NEEDS TO BE UPDATED FOR 2,3

in the above Figure 8.10. The overall systematic uncertainties and contribution from each source in all of the sub-channels of the signal region are shown in Table 8.6 and the final contribution of fake events to the signal region are show in Table 8.1 found at the beginning of the chapter.

8.4.2 3ℓ Fakes

The 3ℓ fake method follows the same general strategy as the 2ℓ SS case. Transfer factors are used to extrapolate from anti-tight, fake-rich control regions in data into the signal region. However, the equivalent low jet control regions are too low in statistics to provide the transfer factors from data

Uncertainties		Channels					
		4 jets			≥ 5 jets		
		$e^{\pm}e^{\pm}$	$\mu^{\pm}\mu^{\pm}$	$e^{\pm}\mu^{\pm}$	$e^{\pm}e^{\pm}$	$\mu^{\pm}\mu^{\pm}$	$e^{\pm}\mu^{\pm}$
Statistical	$\Delta\theta_e^{stat}$	39.6	–	14.2	39.6	–	18.5
	$\Delta\theta_\mu^{stat}$	–	24.7	15.8	–	24.7	13.1
	$\Delta N_{\ell anti-\ell}(n\ jets)$(stat)	6.8	18.1	21.3	8.4	25.9	22.7
Systematics	$\Delta\theta_e^{syst}$ (closure)	21.8	–	7.8	26.7	–	12.5
	$\Delta\theta_\mu^{syst}$ (closure)	–	23.3	18.4	–	31.2	19.7
	Q Mis Id ($\ell anti-\ell$)	2.2	–	1.5	2.4	–	1.5
Total		45.7	38.5 (36.3)	35.7	48.5	47.8 (43.9)	39.6
Systematic	Q Mis Id ($\ell\ell$)	24.0	–	8.6	24.0	–	11.3

Table 8.7: Summary of the uncertainties (in %) in $e^{\pm}e^{\pm}$ (reverse Id + reverse isolation method), $\mu^{\pm}\mu^{\pm}$ and $e^{\pm}\mu^{\pm}$. Statistical uncertainty is split into statistical uncertainties on θ_e and θ_μ and uncertainty due to the Control Region size ($\Delta N_{\ell anti-\ell}(n\ jets)$). For channels with muons, uncertainties between parenthesis are obtained when anti-tight muons are defined such as $p_{\mathrm{T}} < 20$ GeV (includes blinded data)

directly. Instead, the transfer factors are obtained directly from the $t\bar{t}$ simulation. Data control regions, called auxiliary regions, are used to determine the modeling of the identification and isolation variables, used in the transfer factor extrapolation. The low jet regions are still employed in a low statistics validation of the entire fake procedure.

Anti-tight electrons and muons are defined in slightly different ways, compared to the 2ℓ SS case:

- **anti-tight electron (d):** fails to pass the verytight likelihood operating point, but still verifies the veryloose operating point. The isolation selection is released $\mathrm{E}_T^{rel} > 0.05$, $\mathrm{p}_T^{rel} > 0.05$.

- **anti-tight muon (p):** muons must pass identification but the p_{T} cuts is lowered to 6 GeV. The overlap removal with jets and isolation cuts are released.

The transfer factors, θ_e and θ_μ, extracted from MC, is defined as the ratio of the number of top ($t\bar{t}$ + single-top) events in the signal region, to the number of $t\bar{t}$ events in the anti-tight regions. The factors are calculated in separate flavor regions to ensure that the electron jet fakes and muon jet fakes are calculated separately. The calculation follows the same form as for the 2ℓ SS case, but now lep0, which by construction is almost never a fake is allowed to be either electron or muon in both cases, denoted below in Equations 8.12 and 8.13.

$$\theta_e = \frac{N^{top}_{xee}}{N^{top}_{xed}} \tag{8.12}$$

$$\theta_\mu = \frac{N^{top}_{x\mu\mu}}{N^{top}_{x\mu p}} \tag{8.13}$$

The MC modeling of the variables involved in the transfer factor can be verified when another variable fails. For instance, the MC modeling of the electron isolation variable can be compared to data when the particle identification variable fails and vice-versa. The modeling of muon-jet ΔR, involved in the overlap removal, can be compared when either the isolation variable or the p_{T} fails. The comparison of the electron variables in this manner can be seen in Figure 8.13 and the muon variables in Figure 8.12. The regions used have the same selection as the signal region with an added missing transverse energy requirement, > 60 GeV, to ensure only top fakes. 20% and 30% systematic uncertainties are assigned to the muon and electron transfer factors, respectively, to account for data-MC discrepancies. This method for evaluating data-MC agreement for individual electron and muon variables in turn relies on the assumption that these variables are largely uncorrelated and that the transfer factor itself is factorizable into pieces for each variable. Factorized and fully correlated transfer factors have been compared using MC and shown to have differences smaller than the systematic quoted, suggesting that the uncorrelated assumption is reasonable.

The electron and muon anti-tight control regions, which are scaled by the transfer factors are shown in Figure 8.14. The prompt MC subtracted data in these regions are scaled by the transfer factors to obtain the overall contribution of fake electron and muon events in the signal region. The systematic uncertainties are split between the statistical error on the transfer factor and normalization of the anti-tight control regions and the data-MC comparisons outlined above. For muon fakes the total systematic uncertainty is 25% and for electrons it is 34%. The numbers and uncertainties involved in the calculation are shown in Table 8.8.

Stage	Muon	Electron
Anti-Tight CR Normalization	364.62 ± 20.02 (5%)	38.2 ± 6.9 (17%)
Transfer factor θ_μ	0.0047 ± 0.0011 (23%)	0.0240 ± 0.0064 (29%)
SR Contribution	1.71 ± 0.42 (25%)	0.91 ± 0.29 (34%)

Table 8.8: Summary of regions and inputs to the extraction of the number of $t\bar{t}$ events with a fake muon in the SR

The low jet region (1, 2, 3) is used as a validation for the method. The $t\bar{t}$ and single top fakes in this region are estimated using the procedure above. Similar systematics are assessed. This region with the fake estimate is plotted in Figure 8.15. The agreement of data and summed prediction for the fakes and prompt backgrounds is well within the systematic and statistical uncertainties. The figure also shows the same region with relaxed p_{T} cuts on all leptons to 10 GeV, which enriches the fake contributions greatly. The data and summed fake and prompt predictions are also well within the statistical and systematic uncertainties.

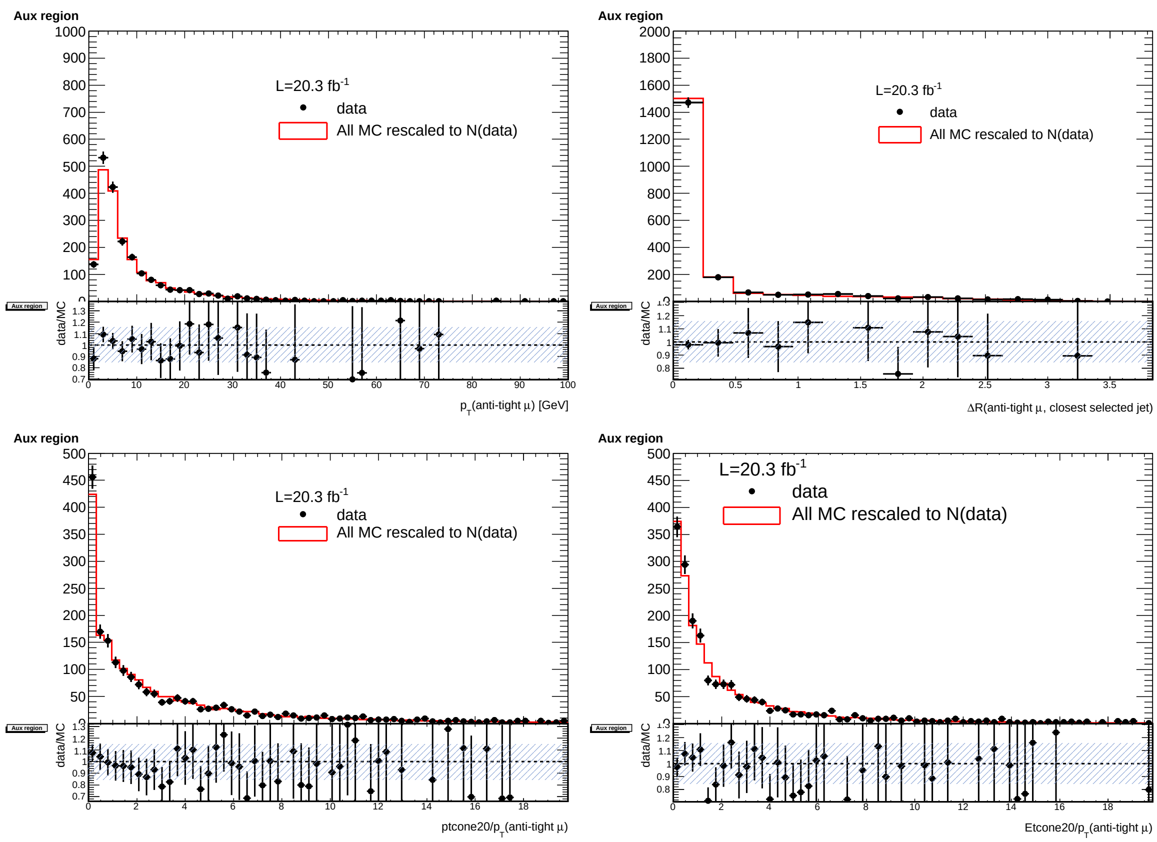


Figure 8.12: Distributions of the properties of the anti-tight muons in data (dots), compared with the total simulation (red line), rescaled to the integral of the data for a shape comparison. The uncertainty on the data distribution is statistical. The number of events for each of them is also presented in the legend. The variables probed are, top: p_T and $\Delta R(\mu, ,closest\,selected\,jet)$; bottom: ptcone20/$p_T$ and Etcone20/p_T. The selection is the signal region event selection with one anti-tight muon (failing at least one of the isolation, muon-jet overlap, or p_T selection criteria). A ratio plot is containing the 20% area around 1, is displayed, demonstrating that a 20% data-MC comparison systematic is sufficient.

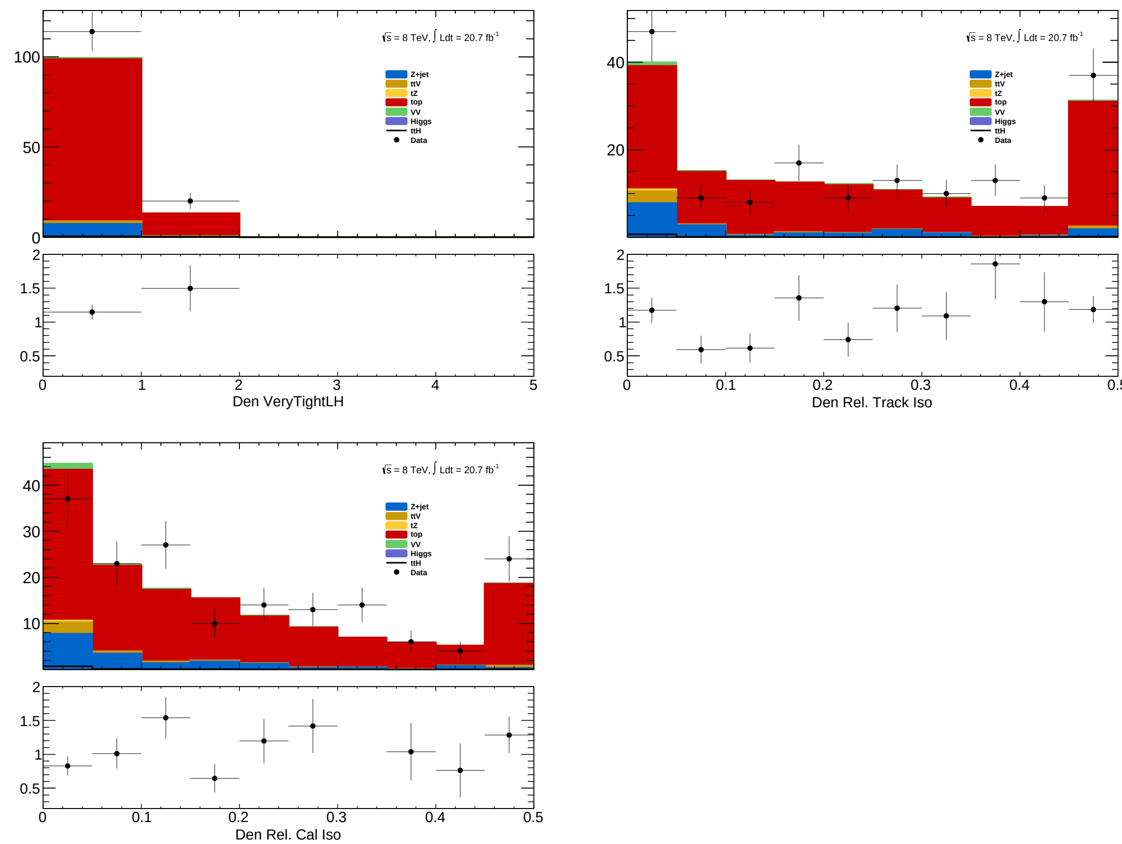


Figure 8.13: Distributions of anti-tight electron variables. The variables presented are, from top left to bottom right, p_T, η, VERYTIGHT Likelihood value, ptcone20/p_T, Etcone20/p_T. The plotted regions have the same cuts as the signal region, except the anti-tight electron must fail isolation for the plot of the VERYTIGHT identification word or fail the VERYTIGHT identification word for the plots of the isolation. Data (dots) are compared with a stacked histogram of the various simulated samples: top in red, V+jets (blue), VV (purple) and $t\bar{t}V$ (yellow). The uncertainty on the data distribution is statistical.

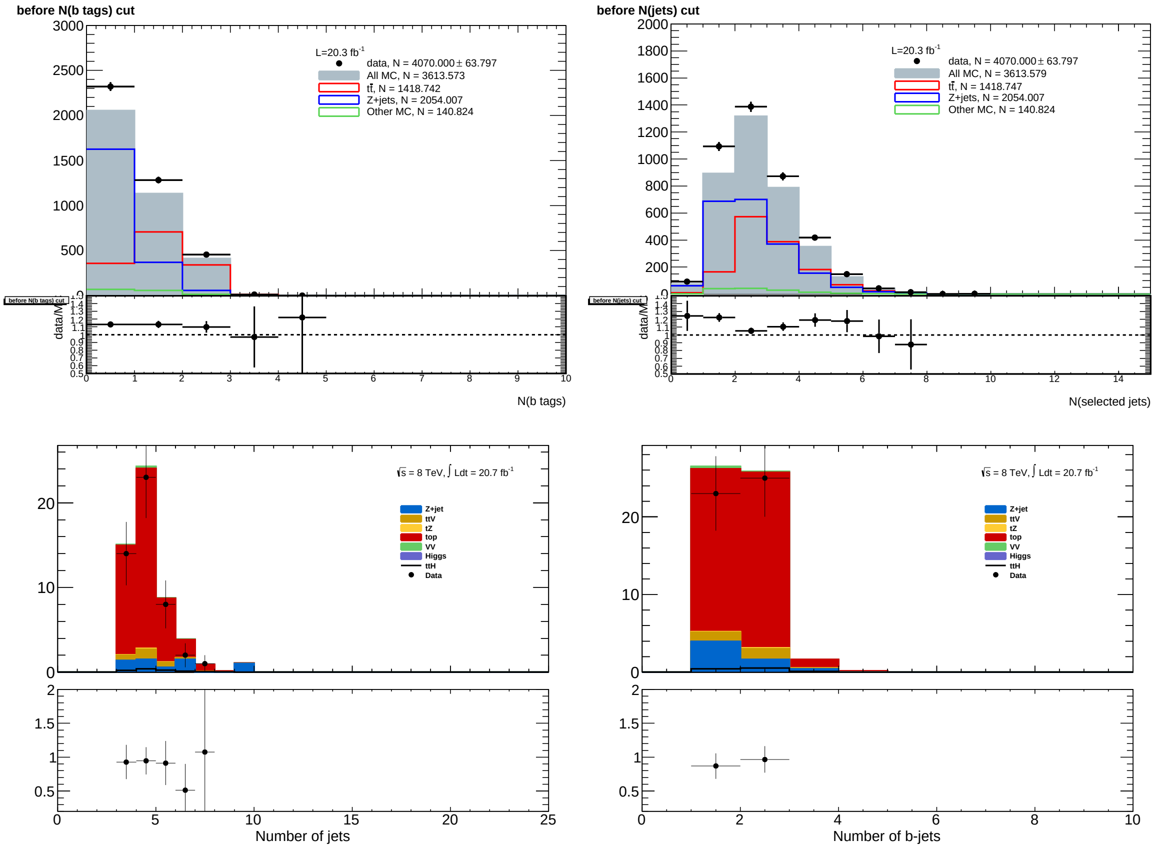


Figure 8.14: Muon-xxp (above) and electron-xxd (below) anti-tight control regions: jet variables. Note: the $t\bar{t}$ and single top MC in the plots is used only as comparison, but is not included in the fake measurement

8.4.3 4ℓ Fakes

We will not discuss the 4ℓ fakes in depth, as it is a very small background - at the % level and will have almost no impact on the final result. The fake method used in the the 4ℓ case is similar to the 2ℓ and 3ℓ cases discussed above. All fakes arise from $t\bar{t}$ and single-top events, where *two* jets are misidentified as leptons. To measure the contribution of this background, control regions with 2 fully identified and 2 anti-identified leptons are created. These control regions do not have a number of jets requirement in order to increase statistics. From these control regions, two extrapolations are made. First, a transfer factor is applied to extrapolate from the anti-tight to tight regions for electrons and muons. The regions are defined with identical object identification selection and reversal as the 3ℓ case, and the same transfer factors can be used. They must be used twice however, because there are two anti-identified leptons in each event. Second, the jet inclusive regions are extrapolated into the 2-jet signal

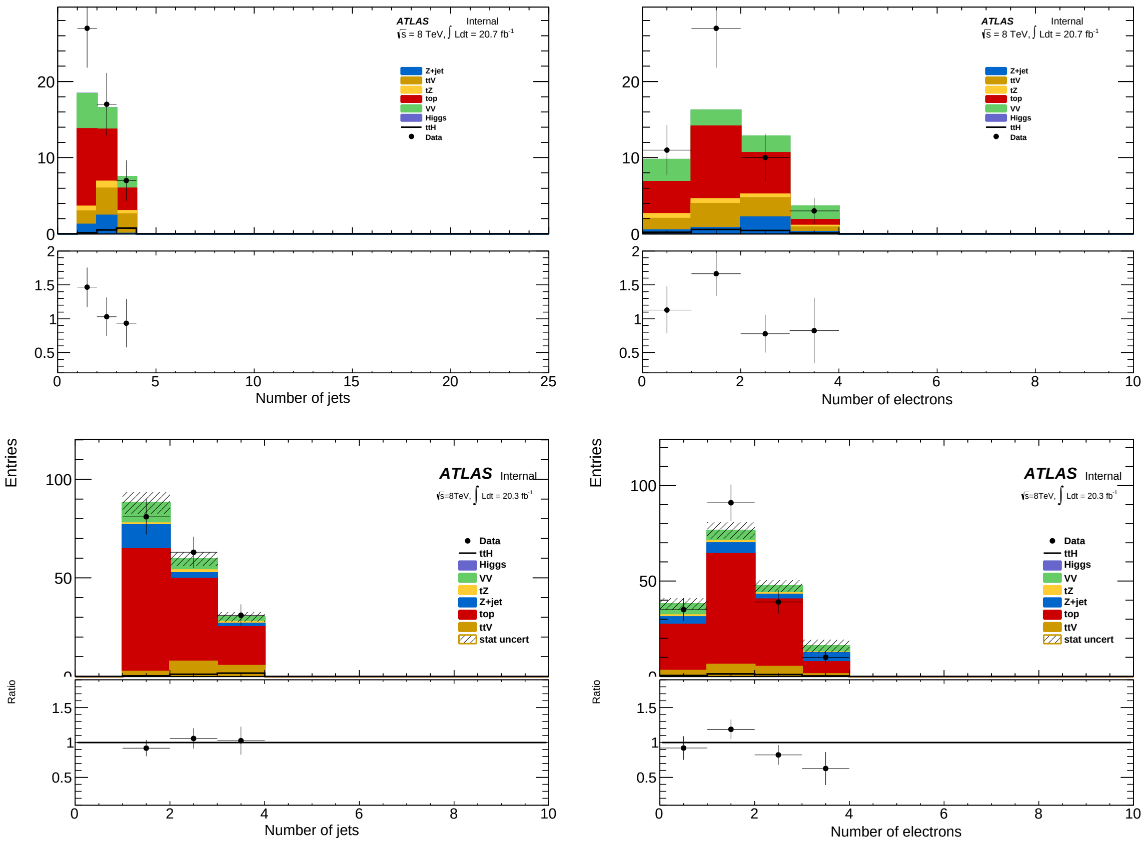


Figure 8.15: 3ℓ fake validation regions for nominal p_{T} selection (above) and relaxed p_{T} selection, $>$ 10 GeV, (below). Plotted are the number of jets and the number of electrons in each event. The data and MC ratio below each plot agree with 1 within the statistics of the region and the overall systematic assigned for the fake component (red)

region, using a second extrapolation factor derived from $t\bar{t}$ events. Since, the majority of fake leptons arise from b-quark initiated jets, the jet spectrum $t\bar{t}$ events with the additional requirement of 2-b-tagged jets from data are used as a model for the jet extrapolation. The overall systematic uncertainty on this measurement arises from the statistics in the control regions and MC based assessments of non-closure and are 35%-50% depending on the sub-channel.

Chapter 9

Summary of Systematic Uncertainties

This chapter summarizes the systematic uncertainties that enter the measurement of the limit of $t\bar{t}H$ multi-lepton analysis. The systematic uncertainties arise from three main sources. The first are the normalization uncertainties on the background process estimation methods, which are discussed in depth in Chapter . The second source is the theoretical uncertainties on the $t\bar{t}H$ production cross-section and acceptance. The final source are the experimental and detector related systematic uncertainties related to event selection efficiencies and measurements and identification of the objects. They affect only the non-data driven backgrounds and the $t\bar{t}H$ signal, as simulation is used to model their acceptance and efficiency for the analysis selection.

These systematic uncertainties, the estimated background and signal event counts in each of the signal regions, and the observed data in each signal region are combined in a statistical fit to an analysis model to extract the measurement of interest. We measure per-channel and combined ratios of the observed production rate to the theoretically predicted production rate of $t\bar{t}H$, a parameter called μ. In the absence of a statistically significant observation, this measurement is translated into a upper confidence limit on μ. The details of this procedure are discussed in the following sections and the results with the observed data are discussed in Chapter 10

9.1 Systematic Uncertainties on Signal Cross-section and Acceptance

The $t\bar{t}H$ signal is simulated with matrix elements at NLO in QCD with Powhel. The simulation details are discussed in Chapter 6. The production cross section and the Higgs boson decay branching fractions together with their theoretical uncertainties from the QCD scale and PDF choice are taken from the NLO theoretical calculations reported in Ref. [66]. The uncertainty from the QCD scale estimated by varying μ_0 by a factor of 2 from the nominal value is $^{+3.8\%}_{-9.3\%}$, while the uncertainty from

Table 9.1: Theoretical uncertainties of the signal event yields in the signal regions due to the impact of QCD scale uncertainties on the event selection.

QCD scale [%]	2ℓ4jets	$2\ell\geq$5jets	3ℓ	4ℓ
Static	$^{+0.6}_{-0.0}$	$^{+2.7}_{-1.3}$	$^{+2.3}_{-0.8}$	$^{+0.9}_{-0.2}$
Dynamic	$^{+1.7}_{-0.8}$	$^{+2.0}_{-2.6}$	$^{+1.7}_{-1.1}$	$^{+0.5}_{-0.0}$

the PDF set and the value of α_S is $\pm 8.1\%$.

The impact of the choice of the QCD scale on the simulation of the $t\bar{t}H$ event selection efficiency is estimated in two independent ways.

First, the factorization and renormalisation scales μ_0 are varied by a factor of 2, as $\mu = 2\mu_0$ and $\mu = \mu_0/2$. The effects of these new scales are estimated via the application of event re-weighting procedures on the nominal simulation using kinematic distributions at parton level. The weights used are dependent on the transverse momenta of both the $t\bar{t}H$ system and of the top quark, as described in Ref. [81].

Second, the choice of the factorization and renormalisation scales, dependent on fixed ("static") parameters in the nominal simulation, is tested comparing its prediction with an alternative ("dynamic"), but still physics motivated choice $\mu_0 = (m_T^t m_T^{\bar{t}} m_T^H)^{\frac{1}{3}}$, which depends on kinematic variables. This comparison is performed via event re-weighting of the nominal static simulation based on weights derived as a function of the $t\bar{t}H$ transverse momentum [81]. In order to take the difference between the choices of scale as systematic uncertainties, a symmetric envelope around the nominal simulation is built applying the weights and also their inverses.

In order to not double-count the variations on the total cross section the predictions from the different QCD scales are normalized to the same total cross section. That means that the observed differences are only coming from the event selection. Significant variations on the jet multiplicities can be seen and these translate into different predictions on the signal event yields in the signal regions. Such differences, listed in Table 9.1, are taken as theoretical systematic uncertainties in addition to the ones affecting the total $t\bar{t}H$ production cross section. The "Static" uncertainties come from the variations by a factor of 2 from the nominal scale and they are correlated with the uncertainties on the total cross section, which are estimated with the same procedure. The "Dynamic" uncertainties come from the difference between the nominal and the alternative dynamic scale and are treated as an independent source of theoretical uncertainty.

The uncertainty of the $t\bar{t}H$ event selection due to the PDF sets is estimated comparing the predictions with three different PDF sets, varying each set within errors and taking the width of the envelope

Table 9.2: Uncertainties on $t\bar{t}H$ acceptance in signal regions due to PDF variation.

Sample	2ℓ 4j	2ℓ 5j	3ℓ	4ℓ
$t\bar{t}H$	0.3%	1.0%	0.5%	1.4%

as systematic uncertainty. The recommended sets are `CT10`, `MSTW2008nlo68cl` and `NNPDF21_100`. We determine the change in the acceptance due to the PDF sets via the formula 9.1 to disentangle it from the change in production cross section.

$$\left(\frac{\text{Reweighed yield in SR}}{\text{Re-weighted total number of events}}\right)\left(\frac{\text{Original yield in SR}}{\text{Original total number of events}}\right)^{-1} - 1 \qquad (9.1)$$

Fig. 9.1 shows the estimated PDF systematic uncertainties as a function of the jet multiplicity in $t\bar{t}H$ events with at least two leptons. The uncertainties are compatible with the uncertainty on the production cross section estimated in Ref. [66] and indicated by the dashed red lines in the lower panel. Table 9.2 shows the half-width of the envelope of the acceptance under all eigenvector variations of the three PDF sets. No significant dependence on the event topology is observed, so that the PDF systematic uncertainty on the $t\bar{t}H$ event selection is neglected.

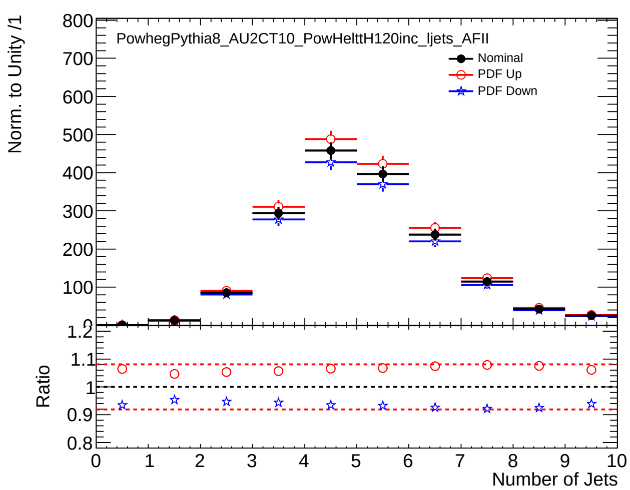


Figure 9.1: PDF systematic uncertainty on the jet multiplicities in $t\bar{t}H$ events with at least 2 leptons. The dashed red lines in the lower panel indicate the systematic uncertainty on the $t\bar{t}H$ production cross section.

9.2 Experimental and Detector Systematic Uncertainties

Experimental and detector systematic uncertainties affect the efficiencies of identifying objects and the efficiencies for events to pass our cuts . These uncertaintites affect only MC models of physics processes, $t\bar{t}V$, $t\bar{t}H$, VV and thus alter ther number of expected events from signal and backgroun in our signal regions. Data-driven backgrounds take into account these uncertainties by construction. We consider systematic effects from a number of sources: the lepton and jet energy scale measurements, the lepton identification and isolation selections, the efficiency and mis-identification rate associated with tagging b-quark jets. Effects due to modeling the energy and objects from additional vertices were studied and found to be negligible. The vast majority of the individual detector systematics effects are small. The sum total of the systematic effects are comparable to the overall normalization and cross-section uncertainties on some of the physics processes and is shown in Table 9.3.

9.2.1 Lepton Identification, Energy Scale, and Trigger

The electron[51] and muon identification efficiencies[82] are measured in data using Z boson and J/Ψ control samples. They are shown in Figure 9.2. The uncertainty on the muon efficiencies are measured as functions of η and p_{T} and are generally less 1%. The uncertainty on the electron and muon efficiencies are also measured as functions of η and p_{T} and are at the 1% level for p_{T} above 30 GeV, but become much larger 5-10% for the lower p_{T}regimes. These translate into sub-% level effects on the $t\bar{t}V$ and $t\bar{t}H$ event counts in the signal regions for the muons and 1% level effects for the electrons. The effects become more important with increasing numbers of leptons.

The electron[83] and muon[82] energy scale and resolution are also measured using the Z-boson control samples in data. The uncertainties related to the scale and resolution for the leptons affect the overall event acceptance through the lepton momentum cuts primary and have negligible impact on the event count uncertainties in the signal regions.

The efficiencies for muons and electrons to pass muon[84] and electron triggers[85] have been calculated with respect to the offline identification operating points using the Z boson control samples. They are in the range of 90% for electron triggers and 70% for muon triggers, owing to gaps in muon trigger coverage, and have 1% level uncertainties. When statistically combined for 2ℓ SS, 3ℓ and 4ℓ lepton signal regions, the overall trigger efficiency is high and the uncertainties on the number of expected events is negligible.

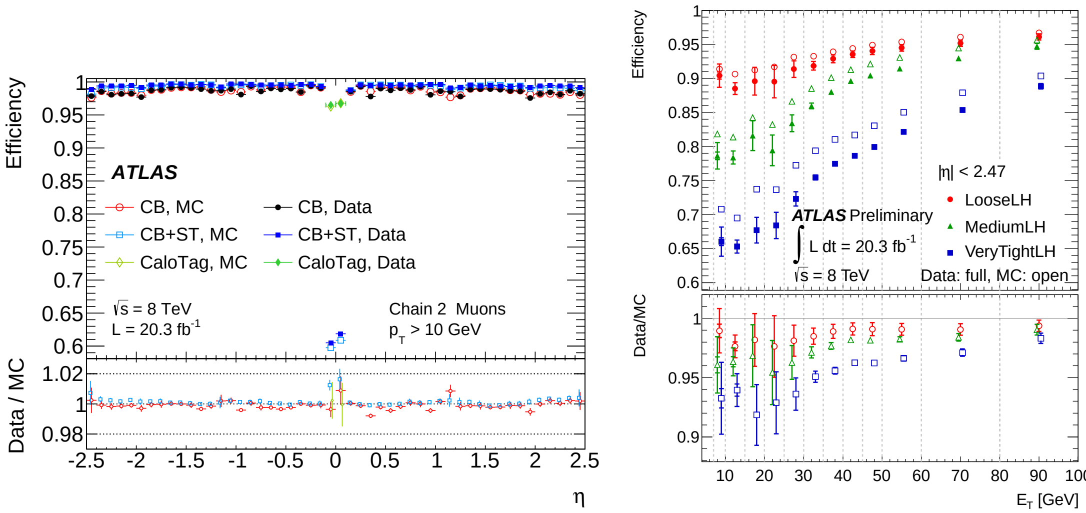


Figure 9.2: Muon (left) and electron(right) identification efficiencies in Data and MC as a function of η and p_{T}respectively. For electrons, the verytight likelihood operating point is used and for muons the CB+ST (combined+segment tagged) operating point is used

9.2.2 Lepton Isolation and Impact Parameter

The isolation and impact parameter selections are specific to this analysis and are discussed in depth in Chapter 7. We calculated their combined efficiency with respect to the full lepton identification selection using the Z boson control samples and define data-MC scale factors to correct the efficiency in the simulation. Background are subtracted using shape templates in the di-lepton invariant mass spectrum. The Z-event template is derived from MC, while the background template is derived from the same-sign control region. We measure the efficiency scale-factors in bins of lepton momentum. Uncertainties are assigned per-bin to account for the level of statistics and variations caused by the fit parameters. An additional 1% uncertainty envelope is added to both the electron and muon measurements to account for trends observed in the dependence of the data-MC efficiency scale-factor as a function of the number of jets. Stability of the efficiency scale-factor as a function of the number of jets is important for this analysis, because event activity in the low jet Z sample where the efficiency is measured is much different from the high jet signal regions where the efficiencies are applied. The dependence of the scale-factor on the number of jets can be seen Figure 9.3. The isolation scale-factor uncertainties are around 1-3% depending on the particle momentum, but these uncertainties propagate to 2-5 % (some of the largest) effects in the event counts in the signal regions. The uncertainties are more important in the regions with more leptons.

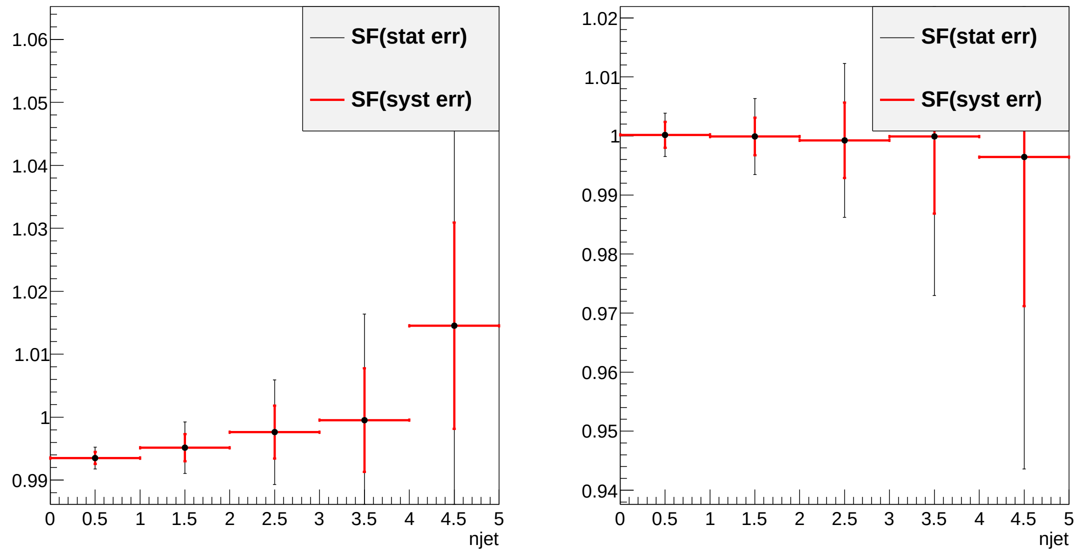


Figure 9.3: Muon (left) and electron(right) isolation efficiency scale-factors from the Z control sample as a function of the number of jets in the event. An additional systematic uncertainty of 1% is added to encompass the variation in the number of jets variable

9.2.3 Jet Energy

The jet energy scale (JES) is calculated using a combination of data-based in-situ techniques, where jet transverse momentum is balanced with respect to a reference photon or a Z boson, as well as single particle test-stand studies[86]. Additional smaller effects are taken into account including the b-quark jet specific response, near-by jets, the effects of pile-up and an inter-calibration of similar η regions using di-jet events. These effects are measured in 2012 data. The JES systematic errors arises from numerous sources that are diagonalized into eigenvectors so that they can be combined in an uncorrelated way. The combined uncertainty is plotted in Figure 9.4 as a function of jet η and p_{T} and is the range 2-4% for jets used in this analysis. The jet energy resolution is calculated in a similar way with slightly larger errors, 10% [87]. The combined scale and resolution systematics are of non-negligible effects 6-7% on the signal and background event counts in the signal regions.

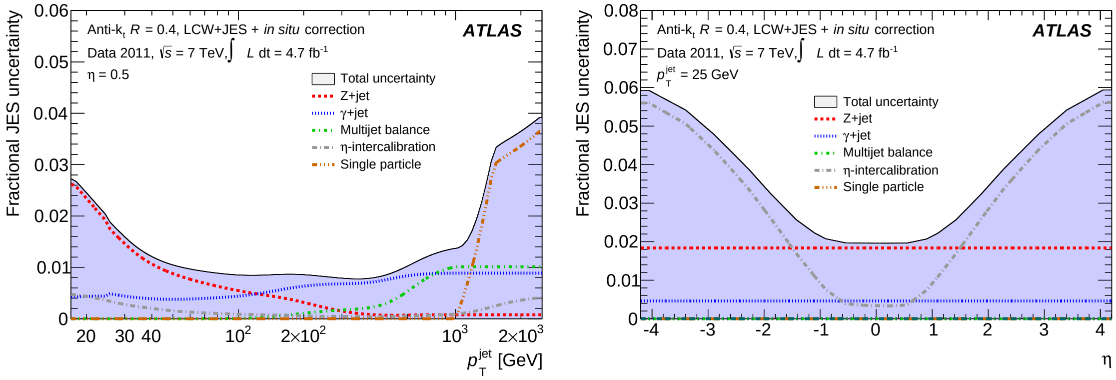


Figure 9.4: JES systematic uncertainties as a function of jet η (for jets $p_{\mathrm{T}}>$ 25 GeV) and p_{T}(for jets $|\eta| < 0.4$). The combined systematic uncertainty is shown with contributions from the largest sources

9.2.4 B-Tagged Jet Efficiency

The b-quark tagging efficiency must be calculated separately for charm, light and b-quarks. ATLAS uses three data based control regions: an inclusive jet sample for mis-tagged light quarks[88], the $t\bar{t}$ sample for b-quarks[89], and a sample of D^* mesons for charm quarks[90]. These efficiencies and rates are well-measured in MC and the data-based corrections are small. The data-MC efficiency scale-factor shown in Figure 9.5 is close to 1 and has an overall systematic uncertainty of around 5%. The uncertainties are applied to the analysis via a number of eigenvectors. Together these uncertainties have a 4 % effect in the event expectation in the signal regions.

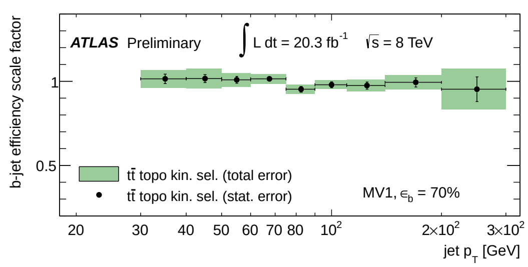


Figure 9.5: b-Tagging data-MC efficiency scale-factors versus jet p_{T} calculated in the $t\bar{t}$ sample from 2012 data. The uncertainties are combined statistical and systematic.

9.2.5 Summary

The combined effect of these detector and experimental systematics on the $t\bar{t}V$ and $t\bar{t}H$ is provided in Table 9.3. The effects are smaller than the normalization uncertainties on some of the backgrounds.

Total Systematic Uncertainty	2ee4j Down-Up		2ee5jincl Down-Up		2em4j Down-Up		2em5jincl Down-Up	
ttH	-4.68	5.84	-8.24	6.14	-5.10	3.50	-5.52	6.40
ttW	-7.20	5.45	-8.72	11.30	-3.63	6.22	-9.72	7.95
ttZ	-9.68	5.07	-5.87	10.98	-4.07	6.16	-8.37	4.99
Total Systematic Uncertainty	2mm4j Down-Up		2mm5jincl Down-Up		3ℓ Down-Up		4ℓ Down-Up	
ttH	-5.20	7.51	-7.28	6.75	-5.84	5.59	-6.54	6.54
ttW	-4.54	5.23	-8.63	6.88	6.36	8.16	—	—
ttZ	-5.24	8.69	-9.73	8.18	-6.14	6.66	-9.58	6.94

Table 9.3: Sum in quadrature of all the systematic uncertainties on the number of event yields per channel.

They are dominated by the lepton isolation scale-factor measurements and the electron identification with smaller contributions from the JES and b-tagging efficiencies. These detector systematic uncertainties enter the fit individually and their ranking of influence on the overall measurement uncertainty can be seen in Figure **??**.

9.3 Summary of Background and Signal Normalization Uncertainties

Tab.9.4 gives the summary of the systematic uncertainties that are included in the analysis for the normalization and acceptance of each process. The relative importance of these uncertainties to the final fit can be seen in Figure.

Type	Description	Name	Uncertainty
		Signal (ttH)	
QCD Scale	Cross Section (Dynamic Scale)	QCDscale_XS_ttH	$+3.8\%$ -9.3% (Section 9.1)
	Analyses Acceptance	QCDscale_Acceptance_ttH	0.-2.6%
PDF+α_S	Cross Section	pdf_ttH	$\pm$ 8.1%
	Analyses Acceptance	pdf_Acceptance_ttH	Negligible
		ttW (Irreducible background)	
QCD Scale	Cross Section (Dynamic Scale)	QCDscale_XS_ttW	$\pm$15% (Section 8.1)
	Analyses Acceptance	QCDscale_Acceptance_ttW	0.4-3.5%
PDF+α_S	Cross Section	pdf_ttW	$\pm$ 13%
	Analyses Acceptance	PDF_Acceptance_ttW	1.1-4.8%
		ttZ (Irreducible background)	
QCD Scale	Cross Section (Dynamic Scale)	QCDscale_XS_ttZ	$\pm$12% (Section 8.1)
	Analyses Acceptance	QCDscale_Acceptance_ttZ	0.1-3.1%
PDF+α_S	Cross Section	pdf_ttZ	$\pm$ 9%
	Analyses Acceptance	PDF_Acceptance_ttZ	0.9-2.7%
		VV Backgrounds	
Normalization Uncertaintiy	WZ,ZZ Processes	WZxsec,ZZxsec	$\pm$ 50% (Section 8.2)
		Data-Driven Backgrounds	
Normalization Uncertainty		Jet Fakes	$\pm$ 30-50% (Section 8.4)
Normalization Uncertainty		Charge MisID	$\pm$ 20-30% (Section 8.3)

Table 9.4: Summary of systematic uncertainties for processes present in the signal regions in the analysis, with their type, description, name, values and uncertainties, and status of inclusion in the final results.

CHAPTER 10

Results and Statistical Model

10.1 Results in Signal Regions

10.1.1 2ℓ SS

10.1.2 3ℓ

10.1.3 4ℓ

10.2 Statistical Model

We use the above results to make two sets of measurements: an upper confidence limit on μ, the signal strength parameter, and a measurement of μ. These measurements are done for each channel individually and then combined. The interpretation of the results in the form of a statistical model follow the procedure, discussed here [91]. We interpret the results as counting experiments in each signal region. Therefore agreement in kinematic shapes do not affect the statistical procedure.

10.2.1 The Likelihood

The observed and expected event yields in the signal regions are analyzed using a binned likelihood function ($\mathcal{L}$), built from product of Poisson models of expected event counts for each bin, where the bins are the separate signal regions:

$$\mathcal{L} \ \alpha \ \prod_{i=0}^{N_{SR}} P(N^i_{obs}|\mu \cdot s^i_{exp} + b^i_{exp}) \tag{10.1}$$

where s^i_{exp} is the SM signal expectation in the signal region, b^i_{exp} are the background expectations, i counts over the signal regions, and P is the Poisson distribution. The signal strength parameter is the parameter of interest in the model (POI) and acts as a simple scale-factor to the SM $t\bar{t}H$production rate and is common to all channels. Setting μ to 0 corresponds to the background only scenario. The background parameter, b, is a sum over all background processes.

The signal and background expectations, s and b, depend on systematic errors. These are included in the likelihood function in the form of a vector nuisance parameters, $\vec{\theta}$, which are constrained to fluctuate within Gaussian distributions. These fluctuations affect the background and signal expectations by response functions, $\nu(\vec{\theta})$, set by systematic uncertainties measured in the previous section. For instance, the $W^{\pm}Z$ normalization uncertainty is 50% from Section 8.2 and is included in the fit as its own unit gaussian,$G(\theta|0,1)$. The fluctuations of the gaussian, θ_{WZ} scale the background contribution via the form, $0.5 \cdot (1+\theta_{WZ}) \cdot b_{WZ}$. For many of the detector systematics, the uncertainties are two sided and are included as piecewise Gaussians. We add correlations to various uncertainties by hand, when appropriate. With these nuisance parameters, the likelihood takes this form:

$$\mathcal{L}(\mu,\vec{\theta}) \ = \left(\prod_{i=0}^{Nch} P(N^i_{obs};\mu\cdot\nu_s(\vec{\theta})\cdot s^i_{exp} + \nu_b(\vec{\theta})\cdot b^i_{exp})\right) \times \prod_j^{N_\theta} G(\theta_j;1,0) \tag{10.2}$$

10.2.2 Test Statistic and Profile Likelihood

Values of μ are tested with the negative log quantity, $q_\mu = -2\ln(\lambda(\mu))$, where $\lambda(\mu)$ is the test statistic. $\lambda(\mu)$ is defined as:

$$\lambda(\mu) \equiv \frac{\mathcal{L}(\mu,\hat{\hat{\vec{\theta}}}_\mu)}{\mathcal{L}(\hat{\mu},\hat{\vec{\theta}})} \tag{10.3}$$

where $\hat{\hat{\vec{\theta}}}_\mu$ are values of the nuisance parameter vector that maximize the likelihood for a given value of μ and $\hat{\mu}$ and $\hat{\vec{\theta}}$ are the fitted values of signal strength and nuisance parameters that maximize the likelihood overall. μ is constrained to be positive.

10.2.3 CL$_s$ Method

Exclusions limits on the signal strength are calculated with the test statistic using a modified frequentist method, called the CL$_s$ method[92]. CL$_s$ is defined as a ratio of two frequentist quantities. The numerator quantifies the probability of finding the observed data given the signal + background hypothesis. The denominator quantifies the probability of the data given the background only hypothesis.

Using the numerator alone has the undesirable property that, if the data fluctuates below the expectation, an exclusion limit can be reached that is far beyond the sensitivity of the experiment. Normalizing to the background only hypothesis penalizes these low sensitivity cases.

The probability of obtaining an observation as extreme as the data given a particular signal + background hypothesis is given by the p-value,i p_μ defined as:

$$p_\mu = \int_{q_\mu^{obs}}^{\infty} f(q_\mu) dq_\mu \tag{10.4}$$

and the probability of obtaining an observation as extreme as the data given the background hypothesis, p_b is:

$$p_b = \int_{q_{\mu=0}^{obs}}^{\infty} f(q_{\mu=0}) dq_{\mu=0} \tag{10.5}$$

where $f(q_\mu)$ is the distribution of q_μ for all possible observations for a given μ and q is defined above. Therefore,

$$CL_s = \frac{p_\mu}{1 - p_b} \tag{10.6}$$

. A μ value is considered excluded at 95% confidence when CL$_s$ is less than 5%.

10.2.4 Exclusion Limits

Table **??** shows *expected* exclusion limits for all channels, including the analysis uncertainties cumulatively. It shows the relative importance of the statistical and systematic uncertainties to the analysis sensitivity. The *observed* limits using observed data and predictions can be seen in Figures **??**-**??** for splitting and combining the sub-channels and in Table XX by numbers. We expect a combined limit of XXX (background only) and XXX (signal injected) and see a limit of XXX. The channel sensitivity is dominated by the 2ℓ and 3ℓ channels.

Channels		Stat	+Fakes Unc.	+Theory	+ Experimental
2ℓ	2ℓee	7.44	8.52	8.82	8.94
	2ℓem	3.46	3.81	4.07	4.18
	2ℓmm	4.03	4.14	4.47	4.57
	2ℓtau	8.08	8.92	10.00	10.03
	All	2.16	2.44	2.81	2.90
3ℓ		3.40	3.43	3.59	3.66
4ℓ		15.16	15.16	15.44	15.55
1l2tau		10.41	13.84	14.20	14.22
All		1.68	1.85	2.14	2.22

Table 10.1: 95%CL limits on μ for all channels and combination with cumulative uncertainties.

10.2.5 μ Measurements

In addition to setting a limit on the signal strength, we also fit the best value of the signal strength for μ. We do this by minimizing the negative log likelihood value, q_μ or conversely maximizing the likelihood. The 1-σ error band is set via a profile likelihood scan, where the value q_μ is scanned as a function of μ. Values of μ that increase q_μ^{min} by 1 form the edges of the error band. The fitted values of μ with errors are provided in Table XXX for each sub-channel fit as well as the combined fit.

10.2.6 Nuisance Parameter Impact on the Signal Strength

Finally, we examine the post-fit impact of the various nuisance parameters on the final fit. We expect to have measured the various analysis uncertainties well and do not expect the fit to have much further constraint. For that reason, we expect the pulls of the nuisance parameters to be close to 0 and the measured uncertainties on those parameters to be consistent with the input uncertainties. Figures XXXX show.

Chapter 11

Conclusions

11.1 Higgs Results in Review

11.2 Prospects for Future

Bibliography

[1] S. L. Glashow, *Partial-symmetries of weak interactions*, Nucl. Phys. **2**2 (1961) no. 4, 579. 2.1.1

[2] S. Weinberg, *A Model of Leptons*, Phys. Rev. Lett. **1**9 (1967) 1264. 2.1.1

[3] A. Salam and J. C. Ward, *Gauge theory of elementary interactions*, Phys. Rev. **1**36 (1964) 763–768. 2.1.1

[4] S. Weinberg, *Non-abelian gauge theories of the strong interactions*, Phys. Rev. Lett. **3**1 (1973) 494–497. 2.1.1

[5] G. 't Hooft and M. Veltman, *Regularization and renormalization of gauge fields*, Nuclear Physics B **44** (1972) 189 – 213. 2.1.1

[6] D. J. Gross and F. Wilczek, *Ultraviolet behavior of non-abelian gauge theories*, Phys. Rev. Lett. **30** (1973) 1343–1346. 2.1.1

[7] P. W. Higgs, *Broken symmetries and the masses of gauge bosons*, Phys. Rev. Lett. **1**3 (1964) 508. 2.1.2

[8] P. W. Higgs, *Spontaneous symmetry breakdown without massless bosons*, Phys. Rev. **1**45 (1966) 1156. 2.1.2

[9] F. Englert and R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons*, Phys. Rev. Lett. **13** (1964) 321–322. 2.1.2

[10] The ALEPH, CDF, DØ, DELPHI, L3, OPAL, SLD Collaborations, the LEP Electroweak Working Group, the Tevatron Electroweak Working Group, and the SLD electroweak and heavy flavour groups, *Precision Electroweak Measurements and Constraints on the Standard Model*, CERN-PH-EP-2010-095 (2010) , arXiv:1012.2367 [hep-ex]. 2.1.3, 2.2

[11] M. Baak, M. Goebel, J. Haller, A. Hoecker, D. Kennedy, R. Kogler, K. Mnig, M. Schott, and J. Stelzer, *The electroweak fit of the standard model after the discovery of a new boson at the LHC*, The European Physical Journal C **7**2 (2012) no. 11, . `http://dx.doi.org/10.1140/epjc/s10052-012-2205-9`. 2.1.3

[12] J. C. Collins, D. E. Soper, and G. Sterman, *Factorization for short distance hadron-hadron scattering*, Nuclear Physics B **261** (1985) 104 – 142. 2.2

[13] CERN, . CERN, Geneva, 1984. 2.2

[14] LHC Higgs Cross Section Working Group, S. Dittmaier, C. Mariotti, G. Passarino, and R. Tanaka (Eds.), *Handbook of LHC Higgs Cross Sections: 2. Differential Distributions*, CERN-2012-002 (CERN, Geneva, 2012) , arXiv:1201.3084 [hep-ph]. 2.2

[15] *Updated coupling measurements of the Higgs boson with the ATLAS detector using up to 25 fb^{-1} of proton-proton collision data*, Tech. Rep. ATLAS-CONF-2014-009, CERN, Geneva, Mar, 2014. 2.2.1

[16] CMS Collaboration Collaboration, *Precise determination of the mass of the Higgs boson and studies of the compatibility of its couplings with the standard model*, Tech. Rep. CMS-PAS-HIG-14-009, CERN, Geneva, 2014. 2.2.1

[17] ATLAS Collaboration Collaboration, G. Aad et al., *Measurement of the Higgs boson mass from the $H \to \gamma\gamma$ and $H \to ZZ^* \to 4\ell$ channels with the ATLAS detector using 25 fb^{-1} of pp collision data*, arXiv:1406.3827 [hep-ex]. 2.2.1

[18] *Evidence for the spin-0 nature of the Higgs boson using {ATLAS} data*, Physics Letters B **7**26 (2013) no. 13, 120 – 144. http://www.sciencedirect.com/science/article/pii/S0370269313006527. 2.2.1

[19] S. Dawson, A. Gritsan, H. Logan, J. Qian, C. Tully, et al., *Working Group Report: Higgs Boson*, arXiv:1310.8361 [hep-ex]. 2.2.2

[20] O. Eberhardt, G. Herbert, H. Lacker, A. Lenz, A. Menzel, et al., *Impact of a Higgs boson at a mass of 126 GeV on the standard model with three and four fermion generations*, Phys.Rev.Lett. **1**09 (2012) 241802, arXiv:1209.1101 [hep-ph]. 2.2.2

[21] M. Carena, S. Gori, N. R. Shah, C. E. Wagner, and L.-T. Wang, *Light Stops, Light Staus and the 125 GeV Higgs*, JHEP **1**308 (2013) 087, arXiv:1303.4414. 2.2.2

[22] N. Arkani-Hamed, K. Blum, R. T. D'Agnolo, and J. Fan, *2:1 for Naturalness at the LHC?*, JHEP **1**301 (2013) 149, arXiv:1207.4482 [hep-ph]. 2.2.2

[23] D. Carmi, A. Falkowski, E. Kuflik, and T. Volansky, *Interpreting LHC Higgs Results from Natural New Physics Perspective*, JHEP **1**207 (2012) 136, arXiv:1202.3144 [hep-ph]. 2.2.2

[24] G. Degrassi, S. Di Vita, J. Elias-Miro, J. R. Espinosa, G. F. Giudice, et al., *Higgs mass and vacuum stability in the Standard Model at NNLO*, JHEP **1**208 (2012) 098, arXiv:1205.6497 [hep-ph]. 2.2.2

[25] L. Evans and P. Bryant, *LHC Machine*, JINST **3** (2008) no. 08, S08001. 3.1

[26] T. S. Pettersson and P. Lefevre, *The Large Hadron Collider: conceptual design.*, Tech. Rep. CERN-AC-95-05 LHC, CERN, Geneva, Oct, 1995. https://cdsweb.cern.ch/record/291782. 3.1

[27] T. Linnecar et al., *Hardware and Initial Beam Commissioning of the LHC RF Systems. oai:cds.cern.ch:1176380*, Tech. Rep. LHC-PROJECT-Report-1172. CERN-LHC-PROJECT-Report-1172, CERN, Geneva, Oct, 2008. https://cdsweb.cern.ch/record/1176380. 3.1

[28] ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, JINST **3** (2008) S08003. 3.1

[29] The CMS Collaboration, *The CMS experiment at the CERN LHC*, Journal of Instrumentation **3** (2008) no. 08, S08004. 3.1

[30] The LHCb Collaboration, *The LHCb Detector at the LHC*, Journal of Instrumentation **3** (2008) no. 08, S08005. 3.1

[31] The ALICE Collaboration, *The ALICE experiment at the CERN LHC*, Journal of Instrumentation **3** (2008) no. 08, S08002. http://stacks.iop.org/1748-0221/3/i=08/a=S08002. 3.1

[32] A. Team, *The four main LHC experiments*, Jun, 1999. 3.1

[33] ATLAS Collaboration, *ATLAS inner detector: Technical Design Report 1*. Technical Design Report ATLAS. CERN, Geneva, 1997. https://cdsweb.cern.ch/record/331063. 3.2

[34] ATLAS Collaboration, *ATLAS inner detector: Technical Design Report, 2*. Technical Design Report ATLAS. CERN, Geneva, 1997. https://cdsweb.cern.ch/record/331064. 3.2

[35] ATLAS Collaboration, *ATLAS magnet system: Technical Design Report, 1*. Technical Design Report ATLAS. CERN, Geneva, 1997. https://cdsweb.cern.ch/record/338080. 3.2

[36] ATLAS Collaboration, *ATLAS pixel detector: Technical Design Report*. Technical Design Report ATLAS. CERN, Geneva, 1998. https://cdsweb.cern.ch/record/381263. 3.2

[37] ATLAS Collaboration, *ATLAS pixel detector electronics and sensors*, JINST **3** (2008) P07007. 3.2

[38] ATLAS Collaboration, *The barrel modules of the ATLAS semiconductor tracker*, Nucl.Instrum.Meth. **A**568 (2006) 642–671. 3.2

[39] ATLAS Collaboration, *The ATLAS semiconductor tracker end-cap module*, Nucl.Instrum.Meth. **A**575 (2007) 353–389. 3.2

[40] The ATLAS TRT Collaboration, *The ATLAS Transition Radiation Tracker (TRT) proportional drift tube: Design and performance*, JINST **3** (2008) P02013. 3.2

[41] The ATLAS TRT Collaboration, *The ATLAS TRT barrel detector*, JINST **3** (2008) P02014. 3.2

[42] E. Abat et al., *The ATLAS TRT electronics*, J. Instrum. **3** (2008) P06007. 3.2

[43] T. A. Collaboration, *ATLAS liquid argon calorimeter: Technical design report*, . CERN-LHCC-96-41. 3.2

[44] ATLAS Collaboration, *ATLAS tile calorimeter: Technical Design Report*. Technical Design Report ATLAS. CERN, Geneva, 1996. https://cdsweb.cern.ch/record/331062. 3.2

[45] ATLAS Collaboration, *ATLAS muon spectrometer: Technical Design Report*. Technical Design Report ATLAS. CERN, Geneva, 1997. https://cdsweb.cern.ch/record/331068. 3.2

[46] G. Aielli, A. Aloisio, M. Alviggi, V. Aprodu, V. Bocci, et al., *The RPC first level muon trigger in the barrel of the ATLAS experiment*, Nucl.Phys.Proc.Suppl. **158** (2006) 11–15. 3.2

[47] F. Bauer, U. Bratzler, H. Dietl, H. Kroha, T. Lagouri, et al., *Construction and test of MDT chambers for the ATLAS muon spectrometer*, Nucl.Instrum.Meth. **A**461 (2001) 17–20. 3.2

[48] T. Argyropoulos, K. A. Assamagan, B. H. Benedict, V. Chernyatin, E. Cheu, et al., *Cathode strip chambers in ATLAS: Installation, commissioning and in situ performance*, IEEE Trans.Nucl.Sci. **56** (2009) 1568–1574. 3.2

[49] ATLAS Collaboration, *ATLAS level-1 trigger: Technical Design Report*. Technical Design Report ATLAS. CERN, Geneva, 1998. https://cdsweb.cern.ch/record/381429. 3.2

[50] P. Jenni, M. Nessi, M. Nordberg, and K. Smith, *ATLAS high-level trigger, data-acquisition and controls: Technical Design Report*. Technical Design Report ATLAS. CERN, Geneva, 2003. https://cdsweb.cern.ch/record/616089. 3.2

[51] *Electron efficiency measurements with the ATLAS detector using the 2012 LHC proton-proton collision data*, Tech. Rep. ATLAS-CONF-2014-032, CERN, Geneva, Jun, 2014. 3.2.6.2, 7.4, 9.2.1

[52] *Preliminary results on the muon reconstruction efficiency, momentum resolution, and momentum scale in ATLAS 2012 pp collsion data*, Tech. Rep. ATLAS-CONF-2013-088, CERN, Geneva, February, 2013. 3.2.6.3, 7.5

[53] M. Cacciari, G. P. Salam, and G. Soyez, *The Anti-k(t) jet clustering algorithm*, JHEP **0**804 (2008) 063, arXiv:0802.1189 [hep-ph]. 3.2.6.4, 7.6

[54] ATLAS Collaboration, *Commissioning of the ATLAS high-performance b-tagging algorithms in the 7 TeV collision data*, Tech. Rep. ATLAS-CONF-2011-102, CERN, Geneva, Jul, 2011. https://cdsweb.cern.ch/record/1369219. 3.2.6.5, 7.6

[55] ATLAS Collaboration Collaboration, 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*, arXiv:1409.3122 [hep-ex]. 5

[56] *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-CONF-2014-011, CERN, Geneva, Mar, 2014. 5

[57] ATLAS Collaboration, G. Aad et al., *Improved luminosity determination in pp collisions at $\sqrt{s}$ = 7 TeV using the ATLAS detector at the LHC*, Eur.Phys.J. **C**73 (2013) 2518, arXiv:1302.4393 [hep-ex]. 6.1.1

[58] CERN, . CERN, Geneva, 2012. 6.1.1

[59] ATLAS Collaboration Collaboration, G. Aad et al., *The ATLAS Simulation Infrastructure*, Eur.Phys.J. **C**70 (2010) 823–874, arXiv:1005.4568 [physics.ins-det]. 6.2

[60] GEANT4 Collaboration, S. Agostinelli et al., *GEANT4: A Simulation toolkit*, Nucl.Instrum.Meth. **A**506 (2003) 250–303. 6.2

[61] G. Bevilacqua, M. Czakon, M. Garzelli, A. van Hameren, A. Kardos, C. Papadopoulos, R. Pittau, and M. Worek. 6.2.1

[62] P. Nason, *A new method for combining NLO QCD with shower Monte Carlo algorithms*, **11** (2004) 040. 6.2.1

[63] S. Frixione, P. Nason, and C. Oleari **11** (2007) 070, arXiv:0709.2092 [hep-ph]. 6.2.1

[64] S. Alioli, P. Nason, C. Oleari, and E. Re **06** (2010) 040, arXiv:1002.2581 [hep-ph]. 6.2.1

[65] T. Sjöstrand, S. Mrenna, and P. Skands, *A Brief Introduction to Pythia 8.1*, arXiv:0710.3820 [hep-ph]. 6.2.1

[66] LHC Higgs Cross Section Working Group Collaboration, S. Heinemeyer et al., *Handbook of LHC Higgs Cross Sections: 3. Higgs Properties*, arXiv:1307.1347 [hep-ph]. 6.2.1, 9.1, 9.1

[67] M. L. Mangano, M. Moretti, F. Piccinini, R. Pittau, and A. D. Polosa, *ALPGEN, a generator for hard multiparton processes in hadronic collisions*, JHEP **0307** (2003) 001, arXiv:hep-ph/0206293 [hep-ph]. 6.2.2

[68] F. Maltoni and T. Stelzer, *MadEvent: Automatic event generation with MadGraph*, JHEP **0302** (2003) 027, arXiv:hep-ph/0208156 [hep-ph]. 6.2.2

[69] B. P. Kersevan and E. Richter-Was, *The Monte Carlo event generator AcerMC versions 2.0 to 3.8 with interfaces to PYTHIA 6.4, HERWIG 6.5 and ARIADNE 4.1*, Comput.Phys.Commun. **184** (2013) 919–985, arXiv:hep-ph/0405247 [hep-ph]. 6.2.2

[70] P. M. Nadolsky, H.-L. Lai, Q.-H. Cao, J. Huston, J. Pumplin, et al., *Implications of CTEQ global analysis for collider observables*, Phys.Rev. **D78** (2008) 013004, arXiv:0802.0007 [hep-ph]. 6.2.2

[71] S. Frixione, P. Nason, and C. Oleari, *Matching NLO QCD computations with Parton Shower simulations: the POWHEG method*, JHEP **0711** (2007) 070, arXiv:0709.2092 [hep-ph]. 6.2.2

[72] T. Gleisberg, S. Hoeche, F. Krauss, M. Schonherr, S. Schumann, et al., *Event generation with SHERPA 1.1*, JHEP **0902** (2009) 007, arXiv:0811.4622 [hep-ph]. 6.2.2

[73] A. Martin, W. Stirling, R. Thorne, and G. Watt, *Parton distributions for the LHC*, Eur.Phys.J. **C63** (2009) 189–285, arXiv:0901.0002 [hep-ph]. 6.2.2, 8.1

[74] M. Garzelli, A. Kardos, C. Papadopoulos, and Z. Trocsanyi, *$t\bar{t}W^{\pm}$ and $t\bar{t}Z$ Hadroproduction at NLO accuracy in QCD with Parton Shower and Hadronization effects*, JHEP **1211** (2012) 056, arXiv:1208.2665 [hep-ph]. 8.1

[75] J. M. Campbell and R. K. Ellis, *$t\bar{t}W^{+-}$ production and decay at NLO*, JHEP **1207** (2012) 052, arXiv:1204.5678 [hep-ph]. 8.1

[76] Campbell, John and Ellis, R. Keith and Röntsch, Raoul, *Single top production in association with a Z boson at the LHC*, Phys.Rev. **D87** (2013) 114006, arXiv:1302.3856 [hep-ph]. 8.1

[77] ATLAS Collaboration Collaboration, G. Aad et al., *Measurement of WZ production in proton-proton collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector*, Eur.Phys.J. **C72** (2012) 2173, arXiv:1208.1390 [hep-ex]. 8.2

[78] ATLAS Collaboration Collaboration, G. Aad et al., *Measurement of ZZ production in pp collisions at $\sqrt{s} = 7$ TeV and limits on anomalous ZZZ and $ZZ\gamma$ couplings with the ATLAS detector*, JHEP **1303** (2013) 128, arXiv:1211.6096 [hep-ex]. 8.2

[79] ATLAS Collaboration Collaboration, G. Aad et al., *Measurement of the cross-section for W boson production in association with b-jets in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector*, JHEP **1306** (2013) 084, arXiv:1302.2929 [hep-ex]. 8.2

[80] ATLAS Collaboration Collaboration, G. Aad et al., *Measurement of differential production cross-sections for a Z boson in association with b-jets in 7 TeV proton-proton collisions with the ATLAS detector*, arXiv:1407.3643 [hep-ex]. 8.2

[81] S. Guindon, E. Shabalina, J. Adelman, M. Alhroob, S. Amor dos Santos, A. Basye, J. Bouffard, M. Casolino, I. Connelly, A. Cortes Gonzalez, V. Dao, S. D'Auria, A. Doyle, P. Ferrari, F. Filthaut, R. Goncalo, N. de Groot, S. Henkelmann, V. Jain, A. Juste, G. Kirby, D. Kar, A. Knue, K. Kroeninger, T. Liss, E. Le Menedeu, J. Montejo Berlingen, M. Moreno Llacer, O. Nackenhorst, T. Neep, A. Onofre, M. Owen, M. Pinamonti, Y. Qin, A. Quadt, D. Quilty, C. Schwanenberger, L. Serkin, R. St Denis, J. Thomas-Wilsker, and T. Vazquez-Schroeder, *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. ATL-COM-PHYS-2013-1659, CERN, Geneva, Dec, 2013. The note contains internal documentation of the ttH(bb) analysis approved as a preliminary result (ATLAS-CONF-2014-011). 9.1

[82] ATLAS Collaboration Collaboration, G. Aad et al., *Measurement of the muon reconstruction performance of the ATLAS detector using 2011 and 2012 LHC proton-proton collision data*, arXiv:1407.3935 [hep-ex]. 9.2.1, 9.2.1

[83] ATLAS Collaboration Collaboration, G. Aad et al., *Electron and photon energy calibration with the ATLAS detector using LHC Run 1 data*, Eur.Phys.J. **C**74 (2014) no. 10, 3071, arXiv:1407.5063 [hep-ex]. 9.2.1

[84] ATLAS Collaboration Collaboration, *Performance of the ATLAS muon trigger in 2011*, Tech. Rep. ATLAS-CONF-2012-099, CERN, Geneva, Jul, 2012. 9.2.1

[85] ATLAS Collaboration Collaboration, *Performance of the ATLAS Electron and Photon Trigger in p-p Collisions at sqrts = 7 TeV in 2011*, Tech. Rep. ATLAS-CONF-2012-048, CERN, Geneva, May, 2012. 9.2.1

[86] ATLAS Collaboration Collaboration, G. Aad et al., *Jet energy measurement and its systematic uncertainty in proton-proton collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector*, arXiv:1406.0076 [hep-ex]. 9.2.3

[87] ATLAS Collaboration Collaboration, G. Aad et al., *Jet energy resolution in proton-proton collisions at $\sqrt{s} = 7$ TeV recorded in 2010 with the ATLAS detector*, Eur.Phys.J. **C**73 (2013) 2306, arXiv:1210.6210 [hep-ex]. 9.2.3

[88] ATLAS Collaboration Collaboration, *Measurement of the Mistag Rate with 5 fb^{-1} of Data Collected by the ATLAS Detector*, Tech. Rep. ATLAS-CONF-2012-040, CERN, Geneva, 2012. 9.2.4

[89] ATLAS Collaboration Collaboration, *Measuring the b-tag efficiency in a top-pair sample with 4.7 fb^{-1} of data from the ATLAS detector* , Tech. Rep. ATLAS-CONF-2012-097, CERN, Geneva, 2012. 9.2.4

[90] ATLAS Collaboration Collaboration, *b-jet tagging calibration on c-jets containing D?+ mesons*, Tech. Rep. ATLAS-CONF-2012-039, CERN, Geneva, 2012. 9.2.4

[91] Glen Cowan, Kyle Cranmer, Eilam Gross, Ofer Vitells, *Asymptotic formulae for likelihood-based tests of new physics*, Eur.Phys.J.C **71** (2011) 1554. 10.2

[92] A. L. Read, *Presentation of search results: the CL s technique*, Journal of Physics G: Nuclear and Particle Physics **2**8 (2002) no. 10, 2693.
http://stacks.iop.org/0954-3899/28/i=10/a=313. 10.2.3