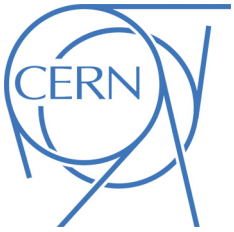
ATLAS NOTE

ATLAS-CONF-2013-018

March 6, 2013



Search for heavy top-like quarks decaying to a Higgs boson and a top quark in the lepton plus jets final state in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector

The ATLAS Collaboration

Abstract

A search is presented for production of a heavy up-type quark (t') together with its antiparticle, assuming a significant branching ratio for subsequent t' decay into a Standard Model Higgs boson and a top quark, as predicted by vector-like quark models. The search is based on 14.3 fb^{-1} of pp collisions at $\sqrt{s} = 8$ TeV recorded in 2012 with the ATLAS detector at the CERN Large Hadron Collider. Data are analysed in the lepton+jets final state, characterised by an isolated electron or muon with moderately high transverse momentum, significant missing transverse momentum, and at least six jets. The search exploits the high total transverse momenta of all final state objects and the high multiplicity of b jets characteristic of signal events with at least one Higgs boson decaying into $b\bar{b}$, to discriminate against the dominant background from top quark pair production. No significant excess of events above the Standard Model expectation is observed, and upper limits are derived for vector-like quarks of various masses in the two-dimensional plane of $BR(t' \to Wb)$ versus $BR(t' \to Ht)$, where H is the Standard Model Higgs boson, assumed to have a mass of 125 GeV. Under the branching ratio assumptions corresponding to a weak-isospin doublet (singlet) scenario, a t' quark with mass lower than 790 (640) GeV is excluded at the 95% confidence level.

ATLAS-CONF-2013-018
07 March 2013

1 Introduction

Since the discovery of the top quark [1,2], which completed the third generation of fundamental fermions in the quark sector of the Standard Model (SM) of particle physics, searches for heavier quarks have been of particular interest in high-energy physics research. These quarks are predicted by many new physics models aimed at solving some of the limitations of the SM.

A compelling possibility is the addition of weak-isospin singlets, doublets or triplets of vector-like quarks [3], defined as quarks for which both chiralities have the same transformation properties under the electroweak group $SU(2) \times U(1)$. Vector-like quarks appear in many extensions of the SM such as little Higgs [4–6] or extra-dimensional models [7,8]. In these models, a top-partner quark, for simplicity referred to here as t', often plays a key role in cancelling the quadratic divergences in the Higgs boson mass induced by radiative corrections involving the top quark.

The large centre-of-mass energy ($\sqrt{s}$) and integrated luminosity in proton-proton (pp) collisions produced at the CERN Large Hadron Collider (LHC) offer a unique opportunity to probe these models. At the LHC, vector-like quarks would be produced predominantly in pairs via the strong interaction for masses below O(1 TeV) [3], with sizable cross sections and clean experimental signatures. For higher masses, single production mediated by the electroweak interaction can potentially dominate, depending on the strength of the interaction between the new quarks and the weak gauge bosons.

Vector-like quarks can couple preferentially with third-generation quarks, as the mixing is proportional to the mass of the SM quark [9], and thus present a rich phenomenology. In particular, a vector-like t' quark has *a priori* three possible decay modes, $t' \to Wb$, $t' \to Zt$, and $t' \to Ht$, with branching ratios that vary as a function of $m_{t'}$ and depend on the weak-isospin quantum number of the t' quark. While all three decay modes can be sizable for a weak-isospin singlet, decays to only Zt and Ht are most natural for a doublet [3]. In the case of a triplet, the t' quark can decay either as a singlet or a doublet depending on its hypercharge. The recent observation of a Higgs-like boson with a mass of ~125 GeV by the ATLAS [10] and CMS [11] Collaborations raises the level of interest for vector-like quark searches, as $t' \to Ht$ and $b' \to Hb$ decays have completely specified final states for a SM-like Higgs boson.

Most searches for $t'\bar{t}'$ production by the ATLAS and CMS Collaborations so far have focused on the $t' \to Wb$ decay mode, exploiting both the lepton+jets signature [12, 13], where one of the W bosons decays leptonically and the other decays hadronically, and the dilepton signature [14, 15], where both W bosons decay leptonically. These searches make the assumption that $BR(t' \to Wb) = 1$, which is reasonable for a chiral fourth generation t' quark. Under this assumption, the most restrictive lower limit obtained on the mass of a t' quark is $m_{t'} > 656$ GeV at 95% confidence level (CL) [16]. A recent search by the CMS Collaboration using the lepton+jets signature [17] has focused instead on the $t' \to Zt$ decay mode assuming that $BR(t' \to Zt) = 1$. Unfortunately, the limits derived by these searches can not easily be applied to other branching ratio values, due to the potentially large signal contamination from mixed decay modes. A consistent treatment of those additional signal contributions is thus necessary to set quasi-model independent limits on the plane of $BR(t' \to Ht)$ vs $BR(t' \to Wb)$ as a function of $m_{t'}$. Such an analysis was done recently by the ATLAS Collaboration [16], using a search for $t'\bar{t}' \to W^+bW^-\bar{b}$ to set constraints in that branching ratio plane: e.g. a t' quark with a mass of 550 GeV and $BR(t' \to Wb) > 0.63$ is excluded at 95% CL, regardless of the value of its branching ratios to Ht and Zt.

The search presented in this note is focused on the lepton+jets signature and requires events with high multiplicities of jets (≥ 6) and b-tagged jets (≥ 2), making it particularly sensitive to decay modes such as $t'\bar{t}' \to HtH\bar{t}$, $ZtHt$ and $WbHt$ [1], with $H \to b\bar{b}$, and thus to the low $BR(t' \to Wb)$ region. This search assumes a SM Higgs boson with mass of 125 GeV. A simple kinematic variable, chosen to be rather insensitive to the signal decay mode in order to ensure good sensitivity over most of the branching ratio plane, is used to discriminate between signal and background, which is dominated by $t\bar{t}$+jets. This

[1] Complex-conjugate decay modes are implicit: $ZtHt$ includes $ZtH\bar{t}$ and $Ht\bar{Z}t$, and $WbHt$ includes $W^+bH\bar{t}$ and $HtW^-\bar{b}$.

represents the first ATLAS search for a t' quark decaying into a Higgs boson.

This note is organised as follows. After a brief overview of the ATLAS detector in Sect. 2, the main reconstructed physics objects used in this search, as well as the dataset and event preselection requirements made are discussed in Sects. 3 and 4, respectively. Section 5 summarises the simulated samples used for signal and backgrounds, while Sect. 6 is devoted to the background predictions based on data-driven techniques. The final event selection and discriminating variable are presented in Sect. 7. A detailed discussion of the systematic uncertainties considered in this search is given in Sect. 8. The statistical analysis and results obtained are presented in Sects. 9 and 10, respectively. Finally, a summary is given in Sect. 11.

2 ATLAS Detector

The ATLAS detector [18] consists of four main subsystems: an inner tracking system surrounded by a superconducting solenoid, electromagnetic and hadronic calorimeters, and a muon spectrometer. The inner detector provides tracking information from pixel and silicon microstrip detectors in the pseudorapidity[2] range $|\eta| < 2.5$ and from a transition radiation tracker covering $|\eta| < 2.0$, all immersed in a 2 T magnetic field provided by a superconducting solenoid. The electromagnetic (EM) sampling calorimeter uses lead and liquid-argon (LAr) and is divided into a barrel region ($|\eta| < 1.475$) and an end-cap region ($1.375 < |\eta| < 3.2$). Hadron calorimetry is based on two different detector technologies, with scintillator tiles or LAr as active media, and with either steel, copper, or tungsten as the absorber material. The calorimeters cover $|\eta| < 4.9$. The muon spectrometer measures the deflection of muon tracks within $|\eta| < 2.7$ using multiple layers of high-precision tracking chambers located in a toroidal field of approximately 0.5 T and 1 T in the central and end-cap regions of ATLAS, respectively. The muon spectrometer is also instrumented with separate trigger chambers covering $|\eta| < 2.4$.

3 Object Reconstruction

The main physics objects considered in this search are electrons, muons, jets, b jets and missing transverse momentum. A brief summary of the main reconstruction and identification criteria applied for each of these physics objects is given below.

Electron candidates [19] are reconstructed from energy deposits (clusters) in the EM calorimeter that are associated to reconstructed tracks in the inner detector. They are required to have a transverse energy, E_{T}, greater than 25 GeV and $|\eta_{\mathrm{cluster}}| < 2.47$ (where $|\eta_{\mathrm{cluster}}|$ is the pseudorapidity of the calorimeter cluster associated with the electron candidate). Candidates in the calorimetry transition region $1.37 < |\eta_{\mathrm{cluster}}| < 1.52$ are excluded. To reduce the background from non-prompt electrons, i.e. from decays of hadrons (including heavy flavour) produced in jets, electron candidates are also required to be isolated. An η-dependent 90% efficient isolation cut, based on the energy sum of cells around the direction of each candidate, is made for a cone of radius $\Delta R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2} = 0.2$. This energy sum excludes cells associated with the electron cluster and is corrected for leakage from the electron cluster itself. A further 90% efficient isolation cut is made on the track transverse momentum (p_{T}) sum around the electron in a cone of radius $\Delta R = 0.3$. The longitudinal impact parameter of the electron track with respect to the selected event primary vertex (see Sect. 4), z_0, is required to be less than 2 mm.

[2] ATLAS uses a right-handed coordinate system with its origin at the nominal interaction point (IP) in the centre of the detector and the z-axis coinciding with the axis of the beam pipe. The x-axis points from the IP to the centre of the LHC ring, and the y-axis points upward. Cylindrical coordinates (r,ϕ) are used in the transverse plane, ϕ being the azimuthal angle around the beam pipe. The pseudorapidity is defined in terms of the polar angle θ as $\eta = -\ln\tan(\theta/2)$. For the purpose of the fiducial selection, this is calculated relative to the geometric centre of the detector; otherwise, it is relative to the reconstructed primary vertex of each event.

Muon candidates are reconstructed from track segments in the various layers of the muon spectrometer, and matched with tracks found in the inner detector. The final candidates are refitted using the complete track information from both detector systems, and required to satisfy $p_{\mathrm{T}} > 25$ GeV and $|\eta| < 2.5$. Additionally, muons are required to be separated by $\Delta R > 0.4$ from any selected jet (see below). Muons are required to satisfy a p_{T}-dependent track-based isolation requirement that has good performance even under high pileup conditions or in boosted configurations where the muon is close to a jet: the scalar sum of the track p_{T} in a cone of variable radius $\Delta R < 10\,\mathrm{GeV}/p_{\mathrm{T}}^{\mu}$ around the muon (excluding the muon track itself) must be less than 5% of the muon p_{T}. Muons are required to have a hit pattern in the inner detector consistent with a well-reconstructed track. Analogously to the electrons, the muon track longitudinal impact parameter with respect to the primary vertex, z_0, is required to be less than 2 mm.

Jets are reconstructed with the anti-k_t algorithm [20–22] with a radius parameter $R = 0.4$ from calibrated topological clusters [18] built from energy deposits in the calorimeters. Prior to jet finding, a local cluster calibration scheme [23, 24] is applied to correct the topological cluster energies for the effects of non-compensation, dead material and out-of-cluster leakage. The corrections are obtained from simulations of charged and neutral particles. After energy calibration [25] jets are required to have $p_{\mathrm{T}} > 25$ GeV and $|\eta| < 2.5$.

To avoid selecting jets from secondary pp interactions, a selection on the so-called "jet vertex fraction" (JVF) variable above 0.5 is applied, representing a requirement that at least 50% of the sum of the p_{T} of tracks with $p_{\mathrm{T}} > 1$ GeV associated with a jet comes from tracks compatible with originating from the primary vertex. During jet reconstruction, no distinction is made between identified electrons and jet energy deposits. Therefore, if any of the jets lie within ΔR of 0.2 of a selected electron, the single closest jet is discarded in order to avoid double-counting of electrons as jets. After this, electrons which are within ΔR of 0.4 of a remaining jet are removed.

Jets are identified as originating from the hadronisation of a b quark (b tagging) via an algorithm [26] using multivariate techniques to combine information from the impact parameters of displaced tracks as well as topological properties of secondary and tertiary decay vertices reconstructed within the jet. The working point used for this search corresponds to 70% efficiency to tag a b-quark jet, with a light-jet rejection factor of ∼130 and a charm jet rejection factor of 5, as determined for b-tagged jets with $p_{\mathrm{T}} > 20$ GeV and $|\eta| < 2.5$ in simulated $t\bar{t}$ events.

The missing transverse momentum ($E_{\mathrm{T}}^{\mathrm{miss}}$) is used to estimate the transverse momentum of an assumed neutrino originating from the decay of one of the W bosons in the $t'\bar{t}'$ final state. $E_{\mathrm{T}}^{\mathrm{miss}}$ is itself reconstructed by first matching each calorimeter energy cluster with either a reconstructed lepton or jet. Failing this, the cluster is left unassociated. The remaining unassociated clusters are then calibrated for energy losses in un-instrumented regions and for different responses of the calorimeters to electromagnetic and hadronic shower components. This calibration scheme is similar to that described in Ref. [27]. $E_{\mathrm{T}}^{\mathrm{miss}}$ is calculated from a vector sum of the calibrated cluster momenta, together with a term associated with muon momenta.

4 Event Preselection

This search is based on 14.3 fb^{-1} of data collected by the ATLAS experiment between April and October 2012 in pp collisions at $\sqrt{s} = 8$ TeV. Only events collected using a single electron or muon trigger under stable beam conditions and for which all detector subsystems were operational are considered. Triggers with different p_{T} thresholds are combined in a logical OR in order to improve the overall efficiency. The p_{T} thresholds are 24 or 60 GeV for electrons and 24 or 36 GeV for muons. The triggers with the lower p_{T} threshold include isolation requirements on the candidate lepton, resulting in inefficiencies at high p_{T} that are recovered by the triggers with higher p_{T} threshold. The triggers use similar but looser selection

criteria than the final reconstruction requirements.

Events accepted by the trigger are required to have at least one reconstructed vertex with at least five associated tracks, consistent with the beam collision region in the $x - y$ plane. If more than one vertex is found, the primary vertex is taken to be the one which has the largest sum of the squared transverse momenta of its associated tracks. Events are discarded if any jet with $p_{\mathrm{T}} > 20$ GeV is independently identified as out-of-time activity from a previous pp collision or as calorimeter noise [28].

Events are required to have exactly one reconstructed electron or muon and at least four jets satisfying the quality and kinematic criteria discussed in Sect. 3. For both electron and muon channels, the selected lepton is required to match ($\Delta R < 0.15$) the lepton reconstructed by the high-level trigger. The background from multijet production is suppressed by a requirement on $E_{\mathrm{T}}^{\mathrm{miss}}$ as well as on the transverse mass of the lepton and $E_{\mathrm{T}}^{\mathrm{miss}}$ (m_{T}) [3]. For both electron and muon channels the requirements are $E_{\mathrm{T}}^{\mathrm{miss}} > 20$ GeV and $E_{\mathrm{T}}^{\mathrm{miss}} + m_{\mathrm{T}} > 60$ GeV.

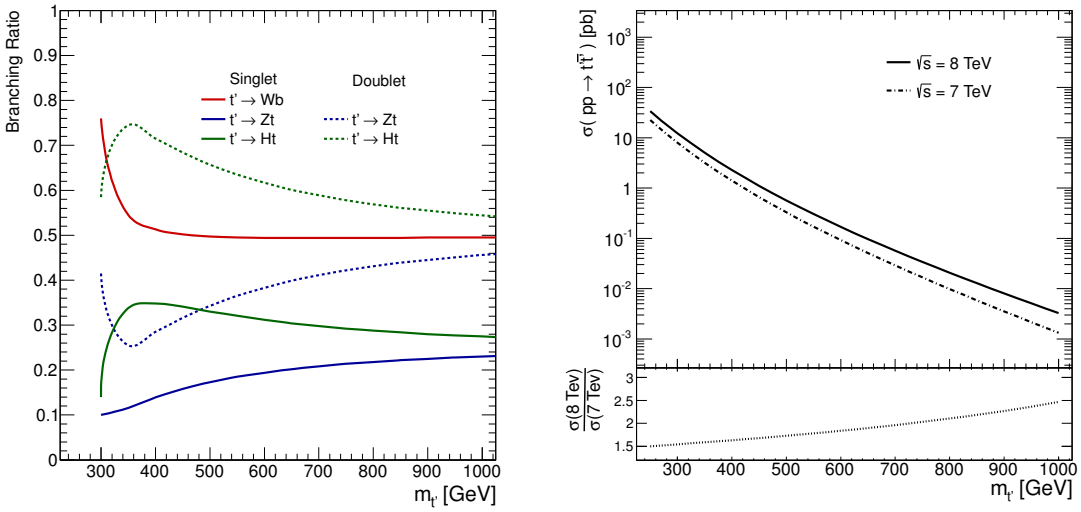
5 Background and Signal Modelling

After event preselection the main background is $t\bar{t}$+jets production, with the production of a W boson in association with jets (W+jets) and multijet events contributing to a lesser extent. Small contributions arise from single top quark, Z+jets and diboson (WW, WZ, ZZ) production, as well as from the associated production of a vector boson and a $t\bar{t}$ pair. Multijet events contribute to the selected sample via the misidentification of a jet or a photon as an electron or the presence of a non-prompt lepton, e.g. from a semileptonic b- or c-hadron decay, and the corresponding yield is estimated via data-driven methods [29]. For the W+jets background, the shape is obtained from the simulation but the normalisation is determined from the data, using the predicted asymmetry between W^++jets and W^-+jets production in pp collisions [30], and separating the events into categories based on the multiplicity of b and c jets. Details on the estimation of the multijet and W+jets backgrounds are given in Sect. 6. The rest of the backgrounds, as well as the signal, are estimated from the simulation and normalised to their theoretical cross sections. In the case of the $t\bar{t}$+jets background prediction, further corrections to match the data are applied, as discussed in Sect. 7.

Samples of $t\bar{t}$+jets and W/Z+jets events are generated using the Alpgen v2.13 [31] leading-order (LO) generator and the CTEQ6L1 PDF set [32]. Parton shower and fragmentation are modelled with Herwig v6.520 [33]. To avoid double-counting of partonic configurations generated by both the matrix-element calculation and the parton-shower evolution, a parton-jet matching scheme ("MLM matching") [34] is employed. The $t\bar{t}$+jets samples are generated separately for $t\bar{t}$+light jets with up to three additional light partons (u, d, s quarks or gluons), and for $t\bar{t}$+heavy-flavour jets, including $t\bar{t}b\bar{b}$ and $t\bar{t}c\bar{c}$. The overlap between $t\bar{t}Q\bar{Q}$ ($Q = b, c$) events generated from the matrix element calculation and those generated from parton-shower evolution in the $t\bar{t}$+light jet samples is avoided via an algorithm based on the angular separation between the extra heavy quarks: if $\Delta R(Q, \bar{Q}) > 0.4$, the matrix-element prediction is used, otherwise the parton-shower prediction is used. These $t\bar{t}$+jets samples are generated assuming a top quark mass of 172.5 GeV and are normalised to the approximate next-to-next-to-LO (NNLO) theoretical cross section calculated with the HATHOR tool [35] using the MSTW2008 NNLO PDF set [36], yielding 238^{+22}_{-24} pb. The W+jets samples are generated with up to five additional partons, separately for W+light jets, $Wb\bar{b}$+jets, $Wc\bar{c}$+jets, and Wc+jets, and their fractions are normalised to measurements in W+1 jet and W+2 jets data control samples [37]. A similar heavy-quark overlap removal prescription as used for $t\bar{t}$+jets is applied. The Z+jets background is normalised to the inclusive NNLO theoretical cross section [38].

[3] $m_{\mathrm{T}} = \sqrt{2p_{\mathrm{T}}^{\ell}E_{\mathrm{T}}^{\mathrm{miss}}(1 - \cos\Delta\phi)}$, where p_{T}^{ℓ} is the transverse momentum (energy) of the muon (electron) and $\Delta\phi$ is the azimuthal angle separation between the lepton and the direction of the missing transverse momentum.

Samples of single top quark backgrounds corresponding to the s-channel and Wt production mechanisms are generated with MC@NLO v4.01 [39–41] using the CT10 PDF set [42]. In the case of t-channel single top quark production, the AcerMC v3.8 LO generator [43] with the MRST LO** PDF set is used. Samples of $t\bar{t}V$ ($V = W, Z$) are generated with the Madgraph v5 LO generator [44] and the CTEQ6L1 PDF set. Samples of $t\bar{t}H$ are generated with the Pythia 6.425 [45] LO generator and the MRST LO** PDF set [46], assuming a Higgs boson mass of 125 GeV and considering the $H \to b\bar{b}$, $c\bar{c}$, gg, and W^+W^- decay modes. Parton shower and fragmentation are modelled with Herwig v6.520 [33] in the case of MC@NLO, with Pythia 6.421 in the case of AcerMC, and with Pythia 6.425 in the case of Madgraph. These samples are generated assuming a top quark mass of 172.5 GeV. The single top quark samples are normalised to the approximate NNLO theoretical cross sections [47, 48] using the MSTW2008 NNLO PDF set, while the $t\bar{t}V$ samples are normalised to the NLO cross section predictions [49, 50]. The $t\bar{t}H$ sample is normalised using the NLO theoretical cross section and branching ratio predictions [51]. Finally, the diboson backgrounds are modelled using Herwig with the MRST LO** PDF set, and are normalised to their NLO theoretical cross sections [52].



(a) (b)

Figure 1: (a) Branching ratios for t' decay as a function of $m_{t'}$ as computed with Protos in the weak-isospin singlet and doublet scenarios. (b) Theoretical cross sections at NNLO for $t'\bar{t}'$ production in pp collisions at two center-of-mass energies, $\sqrt{s} = 7$ TeV and $\sqrt{s} = 8$ TeV, as a function of $m_{t'}$ as computed by Hathor. Also shown is the ratio of production cross sections between both center-of-mass energies.

For vector-like t' signals, samples corresponding to a singlet t' quark decaying to Wb, Zt and Ht are generated with the Protos v2.2 LO generator [3, 53] using the CTEQ6L1 PDF set, and interfaced to Pythia for the parton shower and fragmentation. The predicted branching ratios in the weak-isospin singlet and doublet scenarios as a function of $m_{t'}$ are shown in Fig. 1a. The $m_{t'}$ values considered range from 350 GeV to 850 GeV in steps of 50 GeV, and the Higgs boson mass is assumed to be 125 GeV. All Higgs boson decay modes are considered, with branching ratios as predicted by hdecay [54]. Signal samples are normalized to the approximate NNLO theoretical cross sections [35] using the MSTW2008 NNLO PDF set. The theoretical cross sections as a function of $m_{t'}$ are shown in Fig. 1b.

All event generators using Herwig are also interfaced to Jimmy v4.31 [55] to simulate the underlying event. All simulated samples utilise Photos 2.15 [56] to simulate photon radiation and Tauola 1.20 [57] to simulate τ decays. Finally, all simulated samples include multiple pp interactions and are processed through a simulation [58] of the detector geometry and response using Geant4 [59], with the exception of

the signal samples, for which a fast simulation of the calorimeter response is used. All simulated samples are processed through the same reconstruction software as the data. Simulated events are corrected so that the object identification efficiencies, energy scales and energy resolutions match those determined in data control samples.

6 Data-Driven Background Estimates

6.1 Multijet Background

Multijet events can enter the selected data sample through several production and mis-reconstruction mechanisms. In the electron channel, the multijet background consists of both non-prompt electrons and "fake" electrons, where the latter include both electrons from photon conversions and mis-identified jets with a high fraction of their energy deposited in the EM calorimeter. In the muon channel, the background contributed by multijet events is predominantly due to final states with non-prompt muons, such as those from semileptonic b- or c-hadron decays.

The multijet background normalisation and shape are estimated directly from data by using the "Matrix Method" (MM) technique [29]. The MM exploits differences in lepton identification-related properties between prompt, isolated leptons from W and Z boson decays (referred to as "real leptons" below) and those where the leptons are either non-isolated or result from the mis-identification of photons or jets. For this purpose, two samples are defined after imposing the final kinematic selection criteria, differing only in the lepton identification criteria: a "tight" sample and a "loose" sample, the former being a subset of the latter. The tight selection employs the final lepton identification criteria used in the analysis. For the loose selection the lepton isolation requirements are omitted. The method assumes that the number of selected events in each sample (N^{loose} and N^{tight}) can be expressed as a linear combination of the numbers of events with real and fake leptons, in such a way that the following system of equations can be defined:

$$\begin{aligned} N^{\text{loose}} &= N^{\text{loose}}_{\text{real}} + N^{\text{loose}}_{\text{fake}}, \\ N^{\text{tight}} &= \epsilon_{\text{real}} N^{\text{loose}}_{\text{real}} + \epsilon_{\text{fake}} N^{\text{loose}}_{\text{fake}}, \end{aligned} \tag{1}$$

where ϵ_{real} (ϵ_{fake}) represents the probability for a real (fake) lepton that satisfies the loose criteria to also satisfy the tight ones, and both are measured in data control samples. To measure ϵ_{real} samples enriched in real leptons from W bosons decays are selected by requiring high $E_{\text{T}}^{\text{miss}}$ or m_{T}. The average ϵ_{real} is ~0.75 (~0.98) in the electron (muon) channel. To measure ϵ_{fake} samples enriched in multijet background are selected by requiring either low $E_{\text{T}}^{\text{miss}}$ (electron channel) or high impact parameter significance for the lepton track (muon channel). The average ϵ_{fake} value is ~0.35 (~0.20) in the electron (muon) channel. Dependences of ϵ_{real} and ϵ_{fake} on quantities such as lepton p_{T} and η, ΔR between the lepton and the closest jet, or number of b-tagged jets, are parameterized in order to obtain a more accurate estimate.

6.2 W+jets Background

The estimate of the W+jets background is based on data for its overall normalisation and on the simulation for its shape. In proton-proton collisions W+jets production is charge asymmetric. The total number of W+jets events in data, $N_W = N_{W^+} + N_{W^-}$, can be estimated based on the measured difference between the number of positively- and negatively-charged W bosons, $(N_{W^+} - N_{W^-})_{\text{meas}}$, and the ratio of W^+- to W^--boson production, r_{MC}, determined from the simulation:

$$N_W = \left(\frac{r_{\text{MC}} + 1}{r_{\text{MC}} - 1} \right) (N_{W^+} - N_{W^-})_{\text{meas}} \, . \tag{2}$$

It has been shown that the simulation overestimates the number of W+jets events compared to the data-driven determination by up to ~20%, depending on the jet multiplicity. Corresponding scale factors to correct the prediction from the simulation have been derived [60] for $W+\geq 4$ jets events and are used in this analysis to calibrate the W+jets background to data before b tagging. The W+jets prediction is scaled by a factor of 0.83 ± 0.10 (0.94 ± 0.10) in the electron (muon) channel. The fractions of $Wb\bar{b}$+jets, $Wc\bar{c}$+jets and Wc+jets are determined from data [60]. As a result, the fractions of $Wb\bar{b}$+jets and $Wc\bar{c}$+jets events in the simulation are scaled by a factor of 1.41 ± 0.35 (1.24 ± 0.34) in the electron (muon) channel. The fraction of Wc+jets events in the simulation is scaled by a factor of 0.73 ± 0.37 (0.98 ± 0.34) in the electron (muon) channel. The fraction of W+light jets events is scaled accordingly in order to preserve the overall normalisation of the W+jets background before b tagging, which is obtained from data using Eq. (2).

7 Final Event Selection

This search is focused on $t'\bar{t}'$ production where at least one of the t' quarks decays into a top quark and a Higgs boson. The resulting possible final states are $t'\bar{t}' \to HtH\bar{t}$, $ZtHt$ and $WbHt$. For the dominant $H \to b\bar{b}$ decay mode, the resulting final state is characterized by high jet and b-tag multiplicities, which provide a powerful experimental handle to suppress the backgrounds. Similarly, this search is also sensitive to $t'\bar{t}' \to ZtZ\bar{t}$ and $WbZt$, with $Z \to b\bar{b}$. Therefore, after preselection, the final selection requirements are ≥ 6 jets of which ≥ 2 jets are b tagged, leaving a sample completely dominated by $t\bar{t}$+jets background.

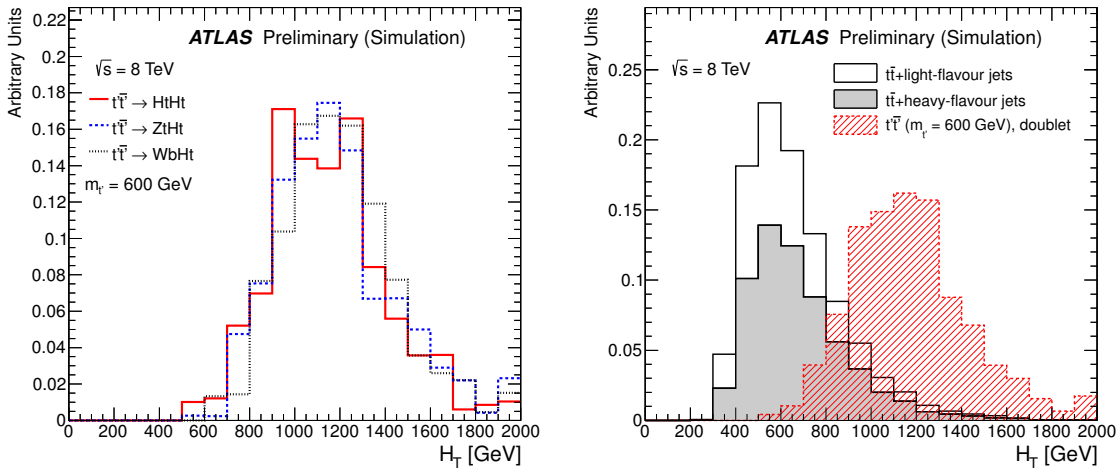
In order to optimize the sensitivity of the search, the selected events are split in three different channels depending on the number of b-tagged jets: 2, 3 and ≥ 4. The channel with ≥ 4 b-tagged jets has the largest signal-to-background ratio and therefore drives the sensitivity of the search. The channels with 2 and 3 b-tagged jets are depleted in signal but they are particularly useful to calibrate the $t\bar{t}$+jets background prediction and constrain the related systematic uncertainties, which in the ≥ 4 b-tags channel are dominated by b tagging, jet energy calibration and physics modelling, including the $t\bar{t}$+heavy-flavour content. A detailed discussion of the systematic uncertainties considered is given in Sect. 8. In the case of the channel with exactly 2 b-tagged jets, an additional requirement of $H_{\mathrm{T}} < 700$ GeV is made, with H_{T} defined as the scalar sum of the p_{T} of the charged lepton, $E_{\mathrm{T}}^{\mathrm{miss}}$ and the p_{T} of the jets . This ensures the orthogonality of this search with another search for vector-like quarks dominantly decaying into Wb with which it will be combined.

To further improve the separation between signal and background, the distinct kinematic features of the signal are exploited. In particular, the large t' quark mass results in energetic leptons and jets in the final state and H_{T} provides a suitable discriminating variable between signal and background which is rather independent of the signal decay mode (see Fig. 2a). The H_{T} distribution typically peaks at $2m_{t'}$ for signal events and at lower values for $t\bar{t}$ background, as illustrated in Fig. 2b.

7.1 Control Regions

Dedicated samples ("control regions") are used to improve or validate the background modelling. These samples are selected by making requirements that suppress potentially large signal contributions, including those from t' decay modes different than the ones targeted by this analysis.

By making a requirement of $H_{\mathrm{T}} < 700$ GeV in the 2, 3 and ≥ 4 b-tag channels, signal-depleted samples are obtained which are dominated by the $t\bar{t}$+jets background. In addition, the heavy-flavour content of the $t\bar{t}$+jets sample varies strongly as a function of the b-tagged jet multiplicity. This fact is exploited to simultaneously fit to data scaling factors for the $t\bar{t}$+light jets and $t\bar{t}$+heavy-flavour jets components of the background prediction.



(a) (b)

Figure 2: Comparison of the shape of the H_{T} distribution in simulation for (a) different $t'\bar{t}'$ decay modes, assuming $m_{t'} = 600$ GeV, and (b) between $t\bar{t}$+jets background (with $t\bar{t}$+light jets and $t\bar{t}$+heavy-flavour jets shown stacked) and $t'\bar{t}'$ signal ($m_{t'} = 600$ GeV) in the t' doublet scenario. The selection used corresponds to the combined e+jets and μ+jets channels with ≥ 6 jets and ≥ 4 b tags. The last bin in all figures contains the overflow.

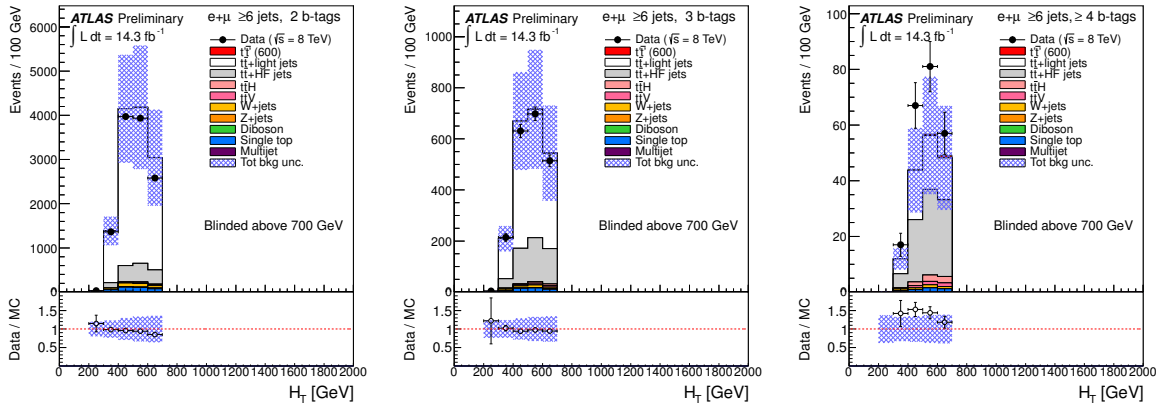
The H_{T} distributions in the 2, 3 and ≥ 4 b-tag channels corresponding to the nominal ALPGEN prediction are shown in Fig. 3. Although reasonable agreement is found within the large assigned systematic uncertainties, in the ≥ 4 b-tag channel the prediction appears systematically below the data. In order to improve the $t\bar{t}$+jets background prediction, a simultaneous fit to the three H_{T} distributions in data is performed where two scaling factors, one for $t\bar{t}$+light jets and another for $t\bar{t}$+heavy-flavour jets, are determined. The measured scaling factors are $0.87 \pm 0.02\,(\mathrm{stat.})$ and $1.35 \pm 0.11\,(\mathrm{stat.})$ for $t\bar{t}$+light jets and $t\bar{t}$+heavy-flavour jets, respectively.

An additional control region is defined by selecting events with at most two jets with $p_{\mathrm{T}} > 60$ GeV and $H_{\mathrm{T}} < 1.2$ TeV, a requirement that effectively suppresses signal, allowing the scrutiny of the H_{T} distribution into the signal region. This control region is only studied in the 2 b-tags and 3 b-tags channels, as the large signal content and low statistics of the ≥ 4 b-tags channel preclude a useful cross-check. Data are found to be in reasonable agreement with the prediction within the assigned systematic uncertainties.

7.2 Signal Region

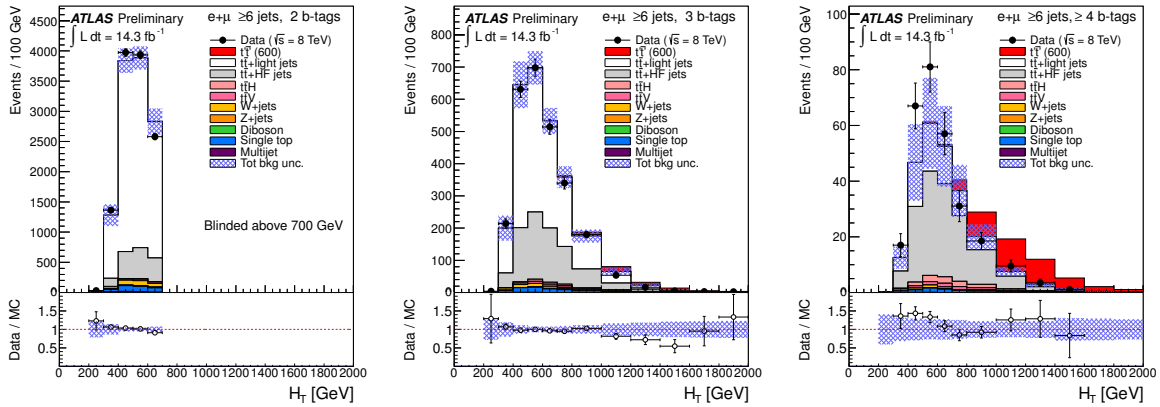
After validation of the background modelling within the assigned systematics uncertainties, the signal is searched for by analyzing the H_{T} spectra in the 2, 3 and ≥ 4 b-tag channels after final selection ("signal region"). The fit to the two scaling factors for $t\bar{t}$+light jets and $t\bar{t}$+heavy-flavour jets is redone, this time considering the full H_{T} spectrum. The measured scaling factors are $0.88 \pm 0.02\,(\mathrm{stat.})$ and $1.21 \pm 0.08\,(\mathrm{stat.})$ for $t\bar{t}$+light jets and $t\bar{t}$+heavy-flavour jets, respectively.

Figure 4 displays the H_{T} distribution in each of the search channels considered, showing the large signal-to-background ratio and good discrimination expected in the sample with ≥ 4 b-tagged jets. This figure displays exactly the same data as in Fig. 3, except that the blinding cut of $H_{\mathrm{T}} < 700$ GeV has been removed in the 3 and ≥ 4 b-tag channels. The data is found to be consistent with the background prediction and no indications of a signal-like excess is observed. Table 1 summarizes the corresponding event yields in each of the analyzed channels.



(a) (b) (c)

Figure 3: Comparison of H_{T} between data and simulation in the combined e+jets and μ+jets channels with ≥ 6 jets and (a) 2 b tags, (b) 3 b tags, and (c) ≥ 4 b tags. A requirement of $H_{\mathrm{T}} < 700$ GeV is made in order to suppress a possible signal contribution. The $t\bar{t}$+jets background is the nominal ALPGEN prediction before the fit to data (see text for details). Also shown is the expected $t'\bar{t}'$ signal corresponding to $m_{t'} = 600$ GeV in the t' doublet scenario. The bottom panel displays the ratio between data and the background prediction. The shaded area represents the total background uncertainty.



(a) (b) (c)

Figure 4: Comparison between data and simulation for H_{T} in the combined e+jets and μ+jets channels with ≥ 6 jets and (a) 2 b tags, (b) 3 b tags, and (c) ≥ 4 b tags. The $t\bar{t}$ background prediction is after fitting to data using the full H_{T} spectrum (see text for details). Also shown is the expected $t'\bar{t}'$ signal corresponding to $m_{t'} = 600$ GeV in the t' doublet scenario. The last bin in all figures contains the overflow. The bottom panel displays the ratio between data and background prediction. The shaded area represents the total post-fit background uncertainty.

8 Systematic Uncertainties

Several sources of systematic uncertainties are considered that can affect the normalisation of signal and background and/or the shape of their corresponding final discriminant distributions. Individual sources of systematic uncertainty are considered uncorrelated. Correlations of a given systematic uncertainty are maintained across processes and channels. Table 2 presents a summary of the systematic uncertainties considered in the analysis indicating whether they are taken to be normalisation-only, or to affect both

	≥ 6 jets, 2 b-tags	≥ 6 jets, 3 b-tags	≥ 6 jets, ≥ 4 b-tags
$t\bar{t}$+heavy-flavour jets	1500 ± 900	900 ± 400	170 ± 70
$t\bar{t}$+light-flavour jets	9600 ± 1000	1900 ± 350	75 ± 22
W+jets	250 ± 130	50 ± 30	5 ± 3
Z+jets	50 ± 40	9 ± 6	0.5 ± 0.9
Single top	300 ± 70	75 ± 18	7 ± 3
Diboson	1.7 ± 0.6	0.3 ± 0.1	0.03 ± 0.03
$t\bar{t}V$	70 ± 20	36 ± 12	7 ± 3
$t\bar{t}H$	28 ± 4	31 ± 6	12 ± 3
Multijet	49 ± 23	1.7 ± 0.8	0.15 ± 0.06
Total background	11860 ± 260	2990 ± 210	270 ± 60
Data	11885	2922	318
Doublet			
$t'\bar{t}'(400)$	550 ± 70	1100 ± 100	790 ± 160
$t'\bar{t}'(600)$	4.3 ± 1.2	94 ± 7	79 ± 18
$t'\bar{t}'(800)$	0.12 ± 0.05	10.7 ± 0.8	9.1 ± 2.1
Singlet			
$t'\bar{t}'(400)$	290 ± 30	650 ± 80	330 ± 70
$t'\bar{t}'(600)$	2.3 ± 0.4	61 ± 7	36 ± 9
$t'\bar{t}'(800)$	0.06 ± 0.01	6.9 ± 0.7	4.2 ± 1.1

Table 1: Predicted and observed yields in the combined e+jets and μ+jets channels with ≥ 6 jets as a function of b-tag multiplicity. The $t\bar{t}$ background prediction is after fitting to data using the full H_{T} spectrum (see text for details). Also shown is the expected $t'\bar{t}'$ signal in both the doublet and singlet scenarios for $m_{t'}$ = 400, 600 and 800 GeV. The uncertainties shown are post-fit and include the effect of statistical and systematic uncertainties. The uncertainty on the total background is smaller than the sum in quadrature of the uncertainties on the individual background sources due to the anti-correlation between the $t\bar{t}$+light jets and $t\bar{t}$+heavy-flavour jets components resulting from the fit.

shape and normalisation. The total prior systematic uncertainty in the background normalisation in the ≥ 4 b-tags channel is ∼42%, with the dominant uncertainties being from b tagging efficiency (16%), c tagging efficiency (11%), jet energy scale (11%), $t\bar{t}$ modelling (11%), $t\bar{t}$+heavy-flavour fractions (32%) and $t\bar{t}$ cross section (10%). As a result of the two-parameter fit, the total background uncertainty is reduced by about 80% in this channel. The total systematic uncertainty in the signal normalisation in the ≥ 4 b-tags channel is ∼21%, completely dominated by the uncertainty in the b tagging efficiency. The following sections describe each of the systematic uncertainties considered in the analysis.

8.1 Luminosity

The luminosity estimate has an uncertainty of 3.6% [61]. This systematic uncertainty is applied to all processes except the multijet background.

8.2 Physics Objects

In this section uncertainties in the reconstruction of leptons, jets, and b-, c-, and light flavour-tagging are considered.

Systematic uncertainty	Type	Components
Luminosity	N	1
Lepton ID+reco+trigger	N	1
Jet vertex fraction efficiency	S	1
Jet energy scale	SN	8
Jet energy resolution	SN	1
b-tagging efficiency	SN	9
c-tagging efficiency	SN	5
Light jet-tagging efficiency	SN	1
$t\bar{t}$ cross section	N	1
$t\bar{t}V$ cross section	N	1
$t\bar{t}H$ cross section	N	1
Single top cross section	N	1
Dibosons cross section	N	1
V+jets normalisation	N	1
Multijet normalisation	N	1
$t\bar{t}$ modelling	SN	3
$t\bar{t}$+heavy-flavour fractions	N	1

Table 2: List of systematic uncertainties considered. A "N" means that the uncertainty is taken as normalisation-only for all processes and channels affected. A "SN" means that the uncertainty is taken as both shape and normalisation, although for small backgrounds only the normalisation uncertainty is considered. Some of the systematic uncertainties are split into several different components for a more accurate treatment.

8.2.1 Lepton Reconstruction, Identification and Trigger

The reconstruction and identification efficiency of electrons and muons, as well as the efficiency of the triggers used to record the events, differ between data and simulation. Scale factors are derived using tag-and-probe techniques on $Z \to \ell^+\ell^-$ ($\ell = e, \mu$) data and simulated samples to correct the simulation for these discrepancies. Since this analysis combines the e+jets and μ+jets channels, a single per-lepton uncertainty of 2.1% is estimated from the quadratic sum of all above contributions on the combined e+jets and μ+jets yields.

8.2.2 Lepton Momentum Scale and Resolution

The accuracy of lepton momentum scale and resolution in simulation is checked using reconstructed distributions of the $Z \to \ell^+\ell^-$ and $J/\psi \to \ell^+\ell^-$ masses. In the case of electrons, E/p studies using $W \to e\nu$ events are also used. Small discrepancies are observed between data and simulation, and corrections for the lepton energy scale and resolution in the latter are applied. In the case of electrons, energy scale corrections need to be applied to data (all regions) and simulation (calorimeter transition region), while energy resolution corrections are applied to the simulation only. In the case of muons, momentum scale and resolution corrections are only applied to the simulation. Uncertainties on both the momentum scale and resolution are considered, and varied separately. The resulting uncertainties on the total yields predicted by the simulation are at the sub-percent level and therefore neglected in the analysis.

8.2.3 Jet Vertex Fraction Efficiency

The per-jet efficiency to satisfy the |JVF| > 0.5 requirement is measured in $Z(\rightarrow \ell^+\ell^-)$+1-jet events in data and simulation, selecting separately events enriched in hard-scatter jets and events enriched in jets from other proton interactions in the same bunch crossing (pileup). Dedicated data/simulation efficiency and inefficiency scale factors are measured separately for both types of jets. The efficiency scale factor for hard-scatter jets decreases from ∼1.03 at p_{T} = 25 GeV to ∼1.01 for p_{T} > 150 GeV. The scale factors for pileup jets are found to be consistent with 1. The product of all per-jet scale factors define a per-event weight used to calibrate the simulation to data. The propagation of the per-jet scale factor uncertainty results in an overall uncertainty on the signal and background acceptance of ∼2.2%.

8.2.4 Jet Energy Scale

The jet energy scale (JES) and its uncertainty have been derived combining information from test-beam data, LHC collision data and simulation [25, 62–65]. The jet energy scale uncertainty is split into 8 uncorrelated sources which can have different jet p_{T} and η dependencies and are treated independently in this analysis. The missing transverse momentum is corrected according to the varied p_{T} of the jets in each event.

8.2.5 Jet Energy Resolution

The jet energy resolution has been measured separately for data and simulation using two *in-situ* techniques [25]. The expected fractional p_{T} resolution for a given jet was measured as a function of its p_{T} and rapidity. A systematic uncertainty is defined as the quadratic difference between the jet energy resolutions for data and simulation. To estimate the corresponding systematic uncertainty in the analysis, the energy of jets in the simulation is smeared by this residual difference, and the changes in the normalisation and shape of the final discriminant are compared to the default prediction. Since jets in the simulation cannot be under-smeared, by definition the resulting uncertainty on the normalisation and shape of the final discriminant is one-sided. This uncertainty is then symmetrised.

8.2.6 Heavy- and Light-Flavour Tagging

The effects of uncertainties in efficiencies for the heavy flavour identification of jets by the b-tagging algorithm have been evaluated. These efficiencies are measured from data and depend on the jet flavour.

Efficiencies for b and c quarks in the simulation have to be corrected by p_{T}-dependent factors of 0.9–1.0 and 1.1–1.2, respectively, whereas the light jet efficiency has to be scaled up by a factor of ∼1.3. These scale factors have an uncertainty between 7% and 13% for b jets, between 15% and 39% for c jets, and ∼25% for light jets. The scale factors and their uncertainties are applied to each jet in the simulation depending on its flavour and p_{T} and η [66–68].

A total of nine and five independent sources of uncertainty are considered for the b-tagging and c-tagging efficiencies, respectively. Each of these uncertainties correspond to a resulting eigenvector after diagonalising the matrix containing the information of total uncertainty per p_{T} bin and the bin-to-bin correlations. These systematic uncertainties are taken as uncorrelated between b, c jets, and light flavour jets. A per-jet weighting procedure is applied to simulated events to propagate the calibration of b tagging and the related uncertainties.

8.3 Normalisations of Data-Driven Backgrounds

In this section the uncertainties in data-driven background estimates are described.

8.3.1 *W/Z*+jets Normalisation

The *W/Z*+jets cross sections from ALPGEN are affected by large uncertainties because they are a leading-order calculation. As discussed in Sect. 6.2, the overall *W*+jets normalisation is obtained via data-driven methods separately for events with exactly 4 and $\geq$ 5 jets in order to ensure the best possible central value for the predicted *W*+jets yield. An additional 24% uncertainty is assigned to the extrapolation of the data-driven estimate to events with $\geq$ 6 jets.

In addition, the following normalisation uncertainties are considered in determining the systematic uncertainty on the estimate of the *W*+jets heavy-flavour content:

- the fractions of $Wb\bar{b}$, $Wc\bar{c}$ and Wc are varied by ±20% in a correlated way; the fraction of *W*+light jets is recomputed to maintain the overall *W*+jets normalisation before any *b*-tagging requirement;
- the fractions of $Wb\bar{b}$, $Wc\bar{c}$ are varied by ±20% in a correlated way, while the fraction of *W*+light jets is left untouched; the fraction of Wc is recomputed to maintain the overall *W*+jets normalisation before any *b*-tagging requirement;
- a theoretical systematic uncertainty on the $Wb\bar{b}$, $Wc\bar{c}$ and Wc fractions from the extrapolation made from 2-jet events, where these fractions are measured in data, to higher jet multiplicity is assessed. This uncertainty is estimated from simulation studies where the *W*+heavy-flavour fractions are studied as a function of variations in the ALPGEN generator parameters. For each jet multiplicity above 2 jets, a 25% uncertainty on these fractions is added in quadrature, resulting in a total uncertainty of 35% for 4 jets, 43% for 5 jets, and 50% for $\geq$ 6 jets. These extrapolation uncertainties on $Wb\bar{b}$ and $Wc\bar{c}$ are treated as fully correlated, and both are treated as uncorrelated with respect to the extrapolation uncertainty on Wc. The fraction of *W*+light jets is recomputed to maintain the overall *W*+jets normalisation before any *b*-tagging requirement.

The sum in quadrature of the above contributions result in a total uncertainty of ~50% on the estimated *W*+jets normalisation for events with $\geq$ 6 jets and $\geq$ 2 *b* tagged jets. This uncertainty is also assigned to the *Z*+jets normalisation.

8.3.2 Multijet Normalisation

Systematic uncertainties on the multijet background estimate via the Matrix Method receive contributions from the limited data statistics, particularly at high jet and *b*-tag multiplicities, as well as from the uncertainty on the method, based on the difference between estimates obtained using different control regions and from the calibration of the method using simulated multijet events. The uncertainty due to the method is assessed to be 50%, which is taken as correlated across jet and *b*-tag multiplicity bins.

8.4 Signal and Background Modelling

The following sections describe uncertainties in the rates and shapes of the discriminating variables arising from estimates based on theoretical calculations.

8.4.1 Theoretical Cross-sections

Uncertainties of +10%/-11% are assumed for the inclusive $t\bar{t}$ production cross section evaluated at approximate NNLO using HATHOR [35]. Uncertainties of +5%/-4% and ±5% are assumed for the theoretical cross sections of the single top [47,48] and diboson [52] backgrounds, respectively. Finally, uncertainties of +12%/-17% and ±30% are assumed for the theoretical cross sections of the $t\bar{t}H$ [51] and $t\bar{t}V$ [49, 50] backgrounds.

8.4.2 $t\bar{t}$+jets Modelling

A number of systematic uncertainties affecting the modelling of $t\bar{t}$+jets are considered in this analysis. Systematic uncertainties associated with the choice of factorisation and renormalisation scales in ALPGEN are considered. For the former, two different uncertainties are taken into account.

On the one hand, the factorisation scale for the hard scatter is varied by a factor of two up and down relative to the original scale, $Q^2 = \sum_{\text{partons}}(m^2 + p_{\text{T}}^2)$. Since sometimes both variations can go in the same direction, the largest of the two is taken and symmetrised. On the other hand, the default choice for the dynamic factorisation scale, $Q^2 = \sum_{\text{partons}}(m^2 + p_{\text{T}}^2)$, is compared to an alternate choice, $Q^2 = x_1 x_2 s$. This uncertainty is significantly larger than that obtained by simply scaling the factorization scale up and down by a factor two and is symmetrised to obtain a two-sided uncertainty.

The renormalisation scale associated with the evaluation of α_s at each local vertex in the matrix element calculation is varied by a factor of two up and down relative to the original scale, k_{T}, between two partons. This uncertainty is only applicable for the $t\bar{t}$+light partons sample, since that is the only sample to which the MLM matching prescription [34] is applied. As a result, this uncertainty cannot be applied to the events originating from the dedicated $t\bar{t}b\bar{b}$ and $t\bar{t}c\bar{c}$ simulated samples. However, this uncertainty is applied to the subset of $t\bar{t}b\bar{b}$ and $t\bar{t}c\bar{c}$ events selected from the $t\bar{t}$+light partons MC samples after the heavy-flavour overlap removal procedure.

8.4.3 $t\bar{t}$+jets Heavy-Flavour Content

The fraction of $t\bar{t}Q\bar{Q}$ ($Q = b, c$) events relative to all $t\bar{t}jj$ events, where j denotes any parton, is one of the most important systematic uncertainties in this analysis. Currently there are no available theoretical predictions for the $t\bar{t}$+heavy-flavour fractions in pp collisions at $\sqrt{s} = 8$ TeV at NLO matched to a parton shower. In order to estimate a systematic uncertainty, the dependence of the ratio of cross sections for $t\bar{t}b\bar{b}$ over $t\bar{t}jj$ as a function of the factorisation scale choice is examined in ALPGEN. These cross sections are computed requiring the extra partons to satisfy $p_{\text{T}} > 20$ GeV, $|\eta| < 2.5$ and $\Delta R(\text{j}, \text{j}) > 0.4$, which are similar requirements to those used in this analysis. The ratio of cross sections is computed for the default factorisation scale choice in ALPGEN, $Q^2 = \sum_{\text{partons}}(m^2 + p_{\text{T}}^2)$, which is then scaled up and down by a factor of two in a correlated way for $t\bar{t}b\bar{b}$ and $t\bar{t}jj$. The variation in the ratio of cross sections is found to be $\leq 25\%$. A similar conclusion is reached if a different dynamic scale, $Q^2 = x_1 x_2 s$, is chosen, and then scaled up and down by a factor of two. The systematic uncertainty assigned to the $t\bar{t}$+heavy-flavour fraction is 50%, conservatively doubling the variation found in the generator-level study with ALPGEN.

Therefore, the fraction of $t\bar{t}Q\bar{Q}$ ($Q = b, c$) events relative to all $t\bar{t}$+jets events is varied up and down by $\pm 50\%$ (relative) with respect to the original ALPGEN prediction. This uncertainty is taken to be fully correlated between the $t\bar{t}b\bar{b}$ and $t\bar{t}c\bar{c}$ fractions. The fraction of $t\bar{t}$+light jet events is adjusted accordingly to preserve the total $t\bar{t}$ yield in each jet multiplicity bin prior to any b-tagging requirement.

9 Statistical Analysis

In the absence of any significant data excess, the H_{T} spectra shown in Fig. 4 are used to derive 95% CL upper limits on the $t't'$ production cross section times branching fraction using the CL_{s} method [69, 70]. This method employs a log-likelihood ratio $LLR = -2\log(L_{\text{s+b}}/L_{\text{b}})$ as test-statistic, where $L_{\text{s+b}}$ (L_{b}) is a binned likelihood function (product of Poisson probabilities) to observe the data under the signal-plus-background (background-only) hypothesis. Pseudo-experiments are generated for both hypotheses, taking into account per-bin statistical fluctuations of the total predictions according to Poisson statistics, as well as Gaussian fluctuations describing the effect of systematic uncertainties. The prediction for the $t\bar{t}$ background in the most sensitive search channel (≥ 6 jets/≥ 4 b tags) is affected by large systematic uncertainties originating from b jet identification, jet energy calibration and physics modelling, including

the fraction of $t\bar{t}$+heavy-flavour jets. In order to reduce the degrading impact of systematic uncertainties on the sensitivity of the search, two nuisance parameters corresponding to scaling factors on the overall yields of $t\bar{t}$+light jets and $t\bar{t}$+heavy-flavour jets, respectively, are fitted to data during the statistical analysis, exploiting the constraining power from the background-dominated regions considered.

The fraction of pseudo-experiments for the signal-plus-background (background-only) hypothesis with *LLR* larger than a given threshold defines $CL_{\mathrm{s+b}}$ (CL_{b}). Such threshold is set to the observed (median) *LLR* for the observed (expected) limit. Signal cross sections for which $CL_{\mathrm{s}} = CL_{\mathrm{s+b}}/CL_{\mathrm{b}} < 0.05$ are deemed to be excluded at 95% CL. Dividing by CL_{b} minimizes the possibility of mistakenly excluding a small signal due to a downward fluctuation of the background.

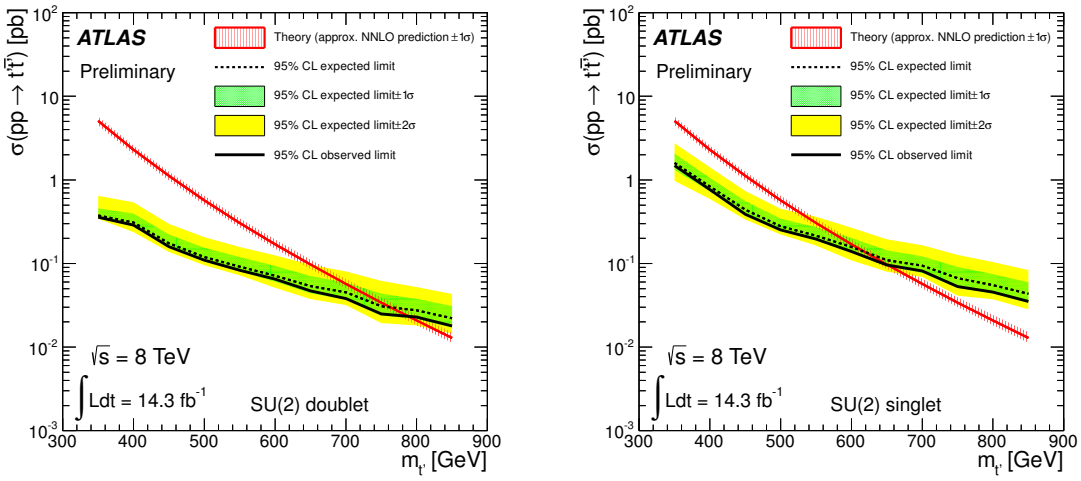
10 Results

The resulting observed and expected upper limits on the $t'\bar{t}'$ production cross section times branching fraction are shown in Fig. 5 as a function of $m_{t'}$, and compared to the theoretical prediction, for two benchmark scenarios, a weak-isospin doublet and singlet t' quark. The total uncertainty on the theoretical cross section [35] includes the contributions from scale variations and PDF uncertainties. For a weak-isospin doublet, an observed (expected) 95% CL limit $m_{t'} > 790\,(745)$ GeV is obtained for the central value of the theoretical cross section. This represents the most stringent limit to date on the mass of a weak-isospin doublet t' quark. For a weak-isospin singlet, the observed (expected) 95% CL limit is $m_{t'} > 640\,(615)$ GeV, also for the central value of the theoretical cross section.

The same analysis is used to derive exclusion limits on vector-like t' quark production, for different values of $m_{t'}$ and as a function of the two branching ratios $BR(t' \to Wb)$ and $BR(t' \to Ht)$. The branching ratio $BR(t' \to Zt)$ is fixed to $BR(t' \to Zt) = 1 - BR(t' \to Wb) - BR(t' \to Ht)$. To probe this two-dimensional branching-ratio plane, the signal samples with the original branching ratios as generated by Protos are weighted. The resulting 95% CL exclusion limits are shown in Fig. 6 for different values of $m_{t'}$. For instance, a t' quark with a mass of 600 GeV and $BR(t' \to Ht) > 0.3$ is excluded at $\geq 95\%$ CL, regardless of the value of its branching ratios to Wb and Zt.

11 Summary

A search has been presented for production of a heavy up-type quark (t') together with its antiparticle, assuming a significant branching ratio for subsequent t' decay into a Standard Model Higgs boson and a top quark, as predicted by vector-like quark models. The search is based on 14.3 fb^{-1} of pp collisions at $\sqrt{s} = 8$ TeV recorded in 2012 with the ATLAS detector at the CERN Large Hadron Collider. Data are analysed in the lepton+jets final state, characterised by an isolated electron or muon with moderately high transverse momentum, significant missing transverse momentum, and at least six jets. The search exploits the high total transverse momenta of all final state objects and the high multiplicity of b jets characteristic of signal events with at least one Higgs boson decaying into $b\bar{b}$, to discriminate against the dominant background from top quark pair production. No significant excess of events above the Standard Model expectation is observed, and upper limits are derived for vector-like quarks of various masses in the two-dimensional plane of $BR(t' \to Wb)$ versus $BR(t' \to Ht)$, where H is the Standard Model Higgs boson, assumed to have a mass of 125 GeV. Under the branching ratio assumptions corresponding to a weak-isospin doublet (singlet) scenario, a t' quark with mass lower than 790 (640) GeV is excluded at the 95% confidence level.



(a) (b)

Figure 5: Observed (solid line) and expected (dashed line) 95% CL upper limits on the $t'\bar{t}'$ cross section times branching fraction for a weak-isospin (a) doublet and (b) singlet t' quark as a function of the t' quark mass. The surrounding shaded bands correspond to the ±1 and ±2 standard deviations around the expected limit. The thin red line and band show the theoretical prediction and its ±1 standard deviation uncertainty.

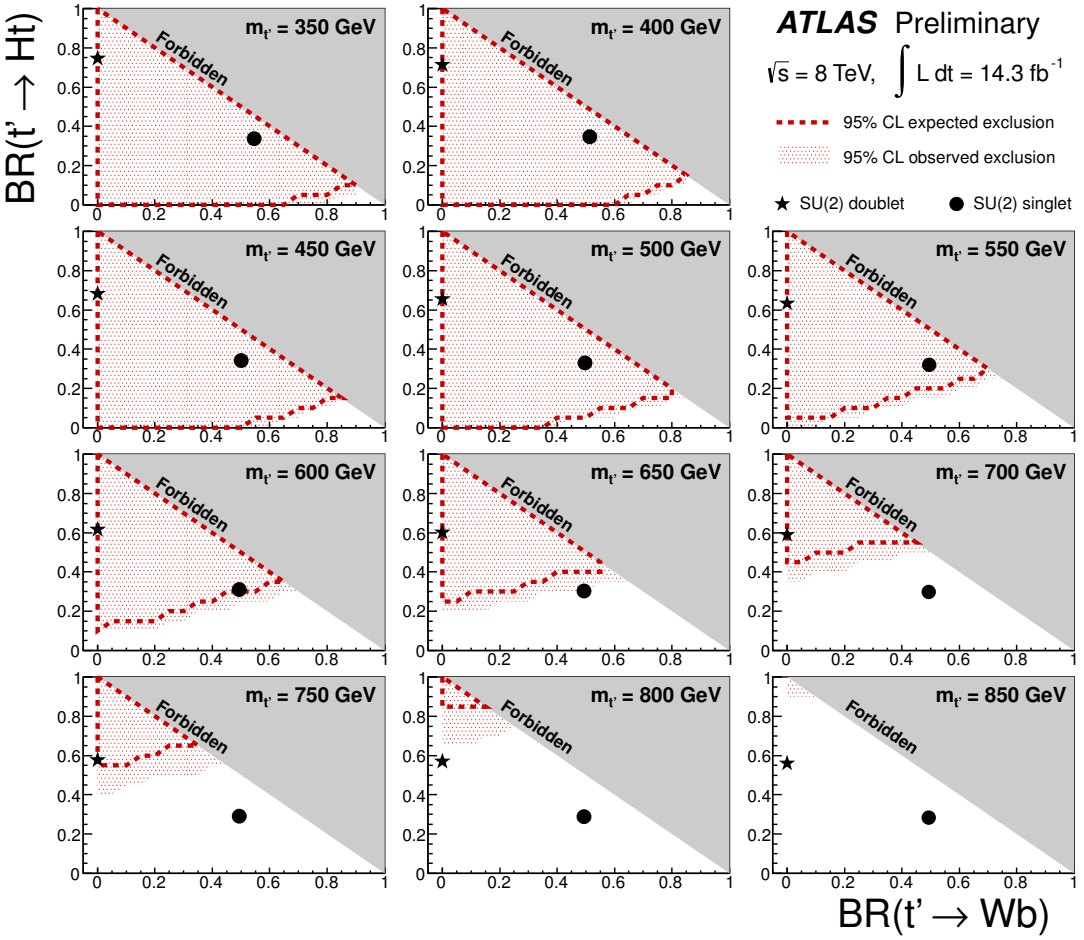


Figure 6: Observed (red filled area) and expected (red dashed line) 95% CL exclusion in the plane of $BR(t' \to Wb)$ versus $BR(t' \to Ht)$, for different values of the vector-like t' quark mass. The grey (dark shaded) area corresponds to the unphysical region where the sum of branching ratios exceeds unity. The default branching ratio values from the Protos event generator for the weak-isospin singlet and doublet cases are shown as plain circle and star symbols, respectively. This result includes both statistical and systematic uncertainties.

References

[1] CDF Collaboration, F. Abe, et al., *Observation of top quark production in $\bar{p}p$ collisions*, Phys. Rev. Lett. **74** (1995) 2626.

[2] D0 Collaboration, S. Abachi, et al., *Observation of the top quark*, Phys. Rev. Lett. **74** (1995) 2632.

[3] J. A. Aguilar-Saavedra, *Identifying top partners at LHC*, JHEP **11** (2009) 030, arXiv:0907.3155 [hep-ph].

[4] N. Arkani-Hamed, A. G. Cohen, and H. Georgi, *Electroweak symmetry breaking from dimensional deconstruction*, Phys. Lett. B **513** (2001) 232, arXiv:0105239 [hep-ph].

[5] N. Arkani-Hamed, A. Cohen, E. Katz, and A. Nelson, *The Littlest Higgs*, JHEP **0207** (2002) 034, arXiv:0206021 [hep-ph].

[6] M. Perelstein, *Little Higgs Models and Their Phenomenology*, Prog. Part. Nucl. Phys. **58** (2007) 247, arXiv:0512128 [hep-ph].

[7] R. Contino, L. D. Rold, and A. Pomarol, *Light custodians in natural composite Higgs models*, Phys. Rev. D **75** (2007) 055014, arXiv:0607106 [hep-ph].

[8] M. Carena, E. Ponton, J. Santiago, and C. E. Wagner, *Light Kaluza-Klein States in Randall-Sundrum Models with Custodial SU(2)*, Nucl. Phys. **B 759** (2006) 202, arXiv:0607106 [hep-ph].

[9] F. del Aguila and M. J. Bowick, *The Possibility of New Fermions with $\Delta I = 0$ Mass*, Nucl. Phys. **B 224** (1983) 107.

[10] ATLAS Collaboration, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Phys. Lett. B **716** (2012) 1, arXiv:1207.7214 [hep-ex].

[11] CMS Collaboration, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, Phys. Lett. B **716** (2012) 30, arXiv:1207.7235 [hep-ex].

[12] ATLAS Collaboration, *Search for pair production of a heavy up-type quark decaying to a W boson and a b quark in the lepton+jets channel with the ATLAS detector*, Phys. Rev. Lett. **108** (2012) 261802, arXiv:1202.3076 [hep-ex].

[13] CMS Collaboration, *Search for pair produced fourth-generation up-type quarks in pp collisions at $\sqrt{s} = 7$ TeV with a lepton in the final state*, Phys. Lett. **B718** (2012) 307, arXiv:1209.0471 [hep-ex].

[14] ATLAS Collaboration, *Search for pair-produced heavy quarks decaying to Wq in the two-lepton channel at $\sqrt{s} = 7$ TeV with the ATLAS detector*, Phys. Rev. D **86** (2012) 012007, arXiv:1202.3389 [hep-ex].

[15] CMS Collaboration, *Search for heavy, top-like quark pair production in the dilepton final state in pp collisions at $\sqrt{s} = 7$ TeV*, Phys. Lett. **B716** (2012) 103, arXiv:1203.5410 [hep-ex].

[16] ATLAS Collaboration, *Search for pair production of heavy top-like quarks decaying to a high-p_{T} W boson and a b quark in the lepton plus jets final state at $\sqrt{s} = 7$ TeV with the ATLAS detector*, Phys. Lett. **B718** (2013) 1284, arXiv:1210.5468 [hep-ex].

[17] CMS Collaboration, *Search for heavy quarks decaying into a top quark and a W or Z boson using lepton + jets events in pp collisions at* $\sqrt{s}$ = 7 TeV, arXiv:1210.7471 [hep-ex].

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

[19] ATLAS Collaboration, *Electron performance measurements with the ATLAS detector using the 2010 LHC proton-proton collision data*, Eur. Phys. J. C **72** (2012) 1909, arXiv:1110.3174 [hep-ex].

[20] M. Cacciari, G. P. Salam, and G. Soyez, *The anti-k_t jet clustering algorithm*, JHEP **04** (2008) 063, arXiv:0802.1189v2 [hep-ph].

[21] M. Cacciari and G. P. Salam, *Dispelling the N^3 myth for the k_t jet-finder*, Phys. Lett. **B641** (2006) 57, arXiv:0512210v2 [hep-ph].

[22] M. Cacciari, G. P. Salam, and G. Soyez, *FastJet User Manual*, Eur. Phys. J. **C72** (2012) 1896, arXiv:1111.6097 [hep-ph]. http://fastjet.fr/.

[23] C. Cojocaru et al., *Hadronic calibration of the ATLAS liquid argon end-cap calorimeter in the pseudorapidity region* $1.6 < |\eta| < 1.8$ *in beam tests*, Nucl. Instr. Meth. A **531** (2004) 481, arXiv:0407009 [physics].

[24] T. Barillari et al., *Local hadronic calibration*, ATL-LARG-PUB-2009-001 (2009) .

[25] ATLAS Collaboration, *Jet energy measurement with the ATLAS detector in proton-proton collisions at* $\sqrt{s}$ = 7 *TeV*, submitted to Eur. Phys. J. , arXiv:1112.6426v1 [hep-ph].

[26] ATLAS Collaboration, *Commissioning of the ATLAS high-performance b-tagging algorithms in the 7 TeV collision data*, ATLAS-CONF-2011-102 (2011) .

[27] ATLAS Collaboration, *Reconstruction and Calibration of Missing Transverse Energy and Performance in Z and W events in ATLAS Proton-Proton Collisions at 7 TeV*, ATLAS-CONF-2011-080 (2011) .

[28] ATLAS Collaboration, *Selection of jets produced in proton-proton collisions with the ATLAS detector using 2011 data*, ATLAS-CONF-2012-020 (2012) .

[29] ATLAS Collaboration, *Measurement of the top quark-pair production cross section with ATLAS in pp collisions at* $\sqrt{s}$ = 7 TeV, Eur. Phys. J. C **71** (2011) 1577, arXiv:1012.1792 [hep-ex].

[30] ATLAS Collaboration, *Measurement of the charge asymmetry in top quark pair production in pp collisions at* $\sqrt{s}$ = 7 TeV *using the ATLAS detector*, Eur. Phys. J. C **72** (2012) 2039, arXiv:1203.4211 [hep-ex].

[31] M. L. Mangano et al., *ALPGEN, a generator for hard multiparton processes in hadronic collisions*, JHEP **07** (2003) 001, arXiv:0206293 [hep-ph].

[32] P. M. Nadolsky et al., *Implications of CTEQ global analysis for collider observables*, Phys. Rev. D **78** (2008) 013004, arXiv:0802.0007 [hep-ph].

[33] G. Corcella et al., *HERWIG 6: an event generator for hadron emission reactions with interfering gluons (including supersymmetric processes)*, JHEP **01** (2001) 010.

[34] M. L. Mangano et al., *Multijet matrix elements and shower evolution in hadronic collisions: $Wb\bar{b}$ + n jets as a case study*, Nucl. Phys. **B 632** (2002) 343, arXiv:0108069 [hep-ph].

[35] M. Aliev et al., *HATHOR: a Hadronic Top and Heavy quarks cross section calculator*, Comput. Phys. Commun. **182** (2011) 1034, arXiv:1007.1327 [hep-ph].

[36] A. D. Martin et al., *Parton distributions for the LHC*, Eur. Phys. J. C **63** (2009) 189, arXiv:0901.0002 [hep-ph].

[37] ATLAS Collaboration, *Measurement of the t-channel single top-quark production cross section in pp collisions at $\sqrt{s}$ = 7 TeV with the ATLAS detector*, Phys. Lett. **B717** (2012) 330, arXiv:1205.3130 [hep-ex].

[38] K. Melnikov and F. Petriello, *Electroweak gauge boson production at hadron colliders through $O(\alpha_s^2)$*, Phys. Rev. D **74** (2006) 114017, arXiv:0609070 [hep-ph].

[39] S. Frixione and B. R. Webber, *Matching NLO QCD computations and parton shower simulations*, JHEP **06** (2002) 029, arXiv:0204244 [hep-ph].

[40] S. Frixione, E. Laenen, P. Motylinski, and B. R. Webber, *Single-top production in MC@NLO*, JHEP **03** (2006) 092, arXiv:0512250 [hep-ph].

[41] S. Frixione, E. Laenen, P. Motylinski, C. White, and B. R. Webber, *Single-top hadroproduction in association with a W boson*, JHEP **07** (2008) 029, arXiv:0805.3067 [hep-ph].

[42] H.-L. Lai et al., *New parton distributions for collider physics*, Phys. Rev. D **82** (2010) 074024, arXiv:1007.2241 [hep-ph].

[43] B. P. Kersevan and E. Richter-Was, *The Monte Carlo Event Generator AcerMC 2.0 with Interfaces to PYTHIA 6.2 and HERWIG 6.5*, arXiv:0405247 [hep-ph].

[44] J. Alwall et al., *MadGraph/MadEvent v4: the new web generation*, JHEP **09** (2007) 028, arXiv:0706.2334 [hep-ph].

[45] T. Sjostrand et al., *High-energy-physics event generation with Pythia 6.1*, Comput. Phys. Commun. **135** (2001) 238, arXiv:0010017 [hep-ph].

[46] A. Sherstnev and R. Thorne, *Parton distributions for LO generators*, Eur. Phys. J. C **55** (2008) 553, arXiv:0711.2473 [hep-ph].

[47] N. Kidonakis, *Next-to-next-to-leading-order collinear and soft gluon corrections for t-channel single top quark production*, Phys. Rev. D **83** (2011) 091503, arXiv:1103.2792 [hep-ph].

[48] N. Kidonakis, *Next-to-next-to-leading logarithm resummation for s-channel single top quark production*, Phys. Rev. D **81** (2010) 054028.

[49] J. M. Campbell and R. K. Ellis, *$t\bar{t}W$ production and decay at NLO*, arXiv:1204.5678 [hep-ph].

[50] M. V. Garzelli, A. Kardos, C. G. Papadopoulos, and Z. Trcsnyi, *$t\bar{t}W$ 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].

[51] LHC Higgs Cross Section Working Group Collaboration, S. Dittmaier et al., *Handbook of LHC Higgs Cross Sections: 1. Inclusive Observables*, arXiv:1101.0593 [hep-ph].

[52] J. Campbell and R. Ellis, *An update on vector boson pair production at hadron colliders*, Phys. Rev. D **60** (1999) 113006, arXiv:9905386 [hep-ph].

[53] J. A. Aguilar-Saavedra, *PROTOS, a Program for Top Simulations*, . http://jaguilar.web.cern.ch/jaguilar/protos/.

[54] A. Djouadi, J. Kalinowski, and M. Spira, *HDECAY: a Program for Higgs Boson Decays in the Standard Model and its Supersymmetric Extension*, Comput. Phys. Commun. **108** (1998) 56, arXiv:9704448 [hep-ph].

[55] J. Butterworth, J. Forshaw, and M. Seymour, *Multiparton interactions in photoproduction at HERA*, Z. Phys. C **72** (1996) 637, arXiv:9601371 [hep-ph].

[56] P. Golonka and Z. Wąs, *PHOTOS Monte Carlo: a precision tool for QED corrections in Z and W decays*, Eur. Phys. J. C **45** (2006) 97, arXiv:0506026 [hep-ph].

[57] S. Jadach, J. H. Kühn, and Z. Waąs, *TAUOLA - a library of Monte Carlo programs to simulate decays of polarized τ leptons*, Comput. Phys. Commun. **64** (1991) 275.

[58] ATLAS Collaboration, *The ATLAS Simulation Infrastructure*, Eur. Phys. J. C **70** (2010) 823, arXiv:1005.4568 [physics.ins-det].

[59] S. Agostinelli et al., *Geant4: a simulation toolkit*, Nucl. Instr. Meth. A **506** (2003) no. 3, 250.

[60] ATLAS Collaboration, *Measurements of top quark pair relative differential cross sections with ATLAS in pp collisions at $\sqrt{s}$ = 7 TeV*, Eur. Phys. J. C **73** (2013) 2261, arXiv:1207.5644 [hep-ex].

[61] ATLAS Collaboration, *Luminosity determination in pp collisions at $\sqrt{s}$ = 7 TeV using the ATLAS detector at the LHC*, Eur. Phys. J. C **71** (2011) 1630, arXiv:1101.2185 [hep-ex].

[62] ATLAS Collaboration, *In-situ jet energy scale and jet shape corrections for multiple interactions in the first ATLAS data at the LHC*, ATLAS-CONF-2011-030 (2011) .

[63] ATLAS Collaboration, *In-situ pseudorapidity intercalibration for evaluation of jet energy scale uncertainty using dijet events in proton-proton collisions at sqrt(s) = 7 TeV*, ATLAS-CONF-2011-014 (2011) .

[64] ATLAS Collaboration, *Jet energy resolution and reconstruction efficiencies from in-situ techniques with the ATLAS Detector Using Proton-Proton Collisions at a Center of Mass Energy sqrts = 7 TeV*, ATLAS-CONF-2010-054 (2011) .

[65] ATLAS Collaboration, *In-situ pseudorapidity intercalibration to evaluate jet energy scale uncertainty and calorimeter performance in the forward region*, ATLAS-CONF-2010-055 (2011) .

[66] ATLAS Collaboration, *Measurement of the b-tag Efficiency in a Sample of Jets Containing Muons with 5 fb^{-1} of Data from the ATLAS Detector*, ATLAS-CONF-2012-043 (2012) .

[67] ATLAS Collaboration, *b-jet tagging calibration on c-jets containing D^* mesons*, ATLAS-CONF-2012-039 (2012) .

[68] ATLAS Collaboration, *Measurement of the Mistag Rate of b-tagging algorithms with 5 fb^{-1} of Data Collected by the ATLAS Detector*, ATLAS-CONF-2012-040 (2012) .

[69] T. Junk, *Confidence level computation for combining searches with small statistics*, Nucl. Instr. Meth. A **434** (1999) 435, arXiv:9902006 [hep-ex].

[70] A. L. Read, *Presentation of search results: the CL_s technique*, J. Phys. G **28** (2002) 2693.