A search for $t \bar{t}$ resonances using lepton plus jets events in proton-proton collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector

The ATLAS Collaboration

Abstract

A search for new particles that decay into top quark pairs is reported. The search is performed with the ATLAS experiment at the LHC using an integrated luminosity of 20.3 fb$^{-1}$ of proton-proton collision data collected at the centre-of-mass energy $\sqrt{s} = 8$ TeV. The lepton plus jets final state is used, where the top-pair decays to $W^+bW^-\bar{b}$, with one $W$ boson decaying leptonically and the other hadronically. The top quark pair invariant mass spectrum is examined for local excesses or deficits that are inconsistent with the Standard Model predictions. No evidence for a top quark pair resonance is found, and 95% confidence level limits on the production rate are determined for massive states in benchmark models. The upper limits on the cross-section times branching ratio of a narrow $Z'$ boson decaying to top pairs range from 4.2 pb to 0.03 pb for resonance masses from 0.4 TeV to 3.0 TeV. A narrow leptophobic topcolour $Z'$ boson with mass below 1.8 TeV is excluded. Upper limits are set on the cross-section times branching ratio for a broad colour-octet resonance with $\Gamma/m = 15\%$ decaying to $t\bar{t}$. These range from 4.8 pb to 0.09 pb for masses from 0.4 TeV to 3.0 TeV. A Kaluza-Klein excitation of the gluon in a Randall–Sundrum model is excluded for masses below 2.2 TeV.

© 2015 CERN for the benefit of the ATLAS Collaboration.
Reproduction of this article or parts of it is allowed as specified in the CC-BY-3.0 license.
1 Introduction

Many models of physics beyond the Standard Model (SM) predict production at the LHC of additional particles with masses about the TeV scale. This paper presents a search for such heavy particles decaying to top-antitop pairs ($t\bar{t}$) using data in proton-proton collisions collected at a centre-of-mass energy $\sqrt{s} = 8$ TeV with the ATLAS detector [1], corresponding to an integrated luminosity of 20.3 fb$^{-1}$.

Searches for production of heavy particles that decay to $t\bar{t}$ are of high interest at the LHC due to the role that the top quark plays in many models of physics beyond the SM (BSM). The top quark is the most massive of the fundamental particles in the SM, and it can have a large coupling to heavy Higgs-boson partners. Thus, heavy Higgs boson-like particles in two-Higgs-doublet models [2, 3] can have a large branching ratio to $t\bar{t}$ final states. Furthermore, many models that propose alternate mechanisms for electroweak symmetry breaking (EWSB) incorporate new heavy particles with a larger coupling to $t\bar{t}$ than to lighter quarks. Examples include strong EWSB models such as topcolour-assisted technicolour [4] (TC2) and Composite Higgs [5–12] scenarios. Models with warped extra dimensions [13–16] form an additional class of models that predict heavy particles that decay to $t\bar{t}$ pairs. In such models, the heavy particles are the counterparts of the gluon and graviton.

The search starts by selecting events with one isolated charged lepton (electron or muon), missing transverse momentum (whose magnitude is denoted as $E_T^{\text{miss}}$) and hadronic jets, compatible with $t\bar{t}\rightarrow W^+bW^-\bar{b}$, with one $W$ boson decaying leptonically and the other hadronically. At least one of the hadronic jets is required to be consistent with having originated from a $b$ quark. An estimator of the $t\bar{t}$ invariant mass ($m_{\text{reco}}^{t\bar{t}}$) is constructed with the events divided into two orthogonal classes by topology: the boosted topology, where the decay products of the hadronically decaying top quark are expected to be fully enclosed within one large-radius jet, and the resolved topology where four small-radius jets are reconstructed and attributed to the $b\bar{b}q\bar{q}'$ quarks. The $m_{\text{reco}}^{t\bar{t}}$ spectrum is scanned for localised excesses or deficits relative to the expectations from the background processes. The compatibility of the data and expectations is assessed, with limits on the production cross-section for new particles are set.

This search is designed to be sensitive to the production of any new particle that decays to $t\bar{t}$. Nonetheless the selection efficiency and acceptance can differ between particular model choices. Hence the sensitivity to a variety of different new particles was evaluated to quantify the performance of the search. The benchmark models adopted in this search include colour singlet and colour octet bosons with spin 0, 1 and 2 and masses from 0.4 to 3 TeV. The resonance width for the specific models varies from very narrow (1%) to a size similar to that of the experimental resolution (15%). Furthermore, the dependence of the limits on the resonance width is explored for heavy gluons up to a width of 40%. With these results, it is possible to interpret the cross section limits in the context of a search for other new particles with the same production modes.

2 Models tested

The details of the benchmark models considered in this search are reviewed below. The interferences of these processes with the SM $t\bar{t}$ production are not considered in this search.
2.1 Spin-1 colour singlet

The first class of models explored produces spin-1 colour singlet vector bosons, $Z'$. This search uses topcolour-assisted technicolour $Z_{TC2}^{'}$ [4, 17, 18] as a benchmark. This is a leptophobic boson, with couplings only to first and third generation quarks, referred to as Model IV in Ref. [17]. The properties of the boson are controlled by three parameters: $\cot \theta_H$, which controls the width and the production cross-section, and $f_1$ and $f_2$, which are related to the coupling to up-type and down-type quarks respectively. Here $f_1 = 1$ and $f_2 = 0$, which maximises the fraction of $Z_{TC2}^{'}$ that decay to $t\bar{t}$. The parameter $\cot \theta_H$ is tuned for each mass point such that the resonance has a width of 1.2% of its mass. To account for higher-order contributions to the cross-section the leading-order calculation is multiplied by a factor of 1.3 based on calculations performed at next-to-leading order (NLO) in QCD [19, 20].

Constraints on $Z_{TC2}^{'}$ have been set by the CDF [21–24] and D0 [25, 26] collaborations using data from proton-antiproton collisions at the Tevatron. Previous constraints at the LHC were set using proton-proton collisions at $\sqrt{s} = 7$ TeV with an integrated luminosity of 5 fb$^{-1}$ by the ATLAS [27, 28] and CMS [29–31] collaborations, and using 20 fb$^{-1}$ of $\sqrt{s} = 8$ TeV data by the CMS Collaboration [32]. For narrow (wide) $Z_{TC2}^{'}$ of width 1.2% (10%) the strongest lower bound on the allowed mass is 2.1 TeV (2.7 TeV) from the $\sqrt{s} = 8$ TeV CMS search.

2.2 Spin-1 colour octet

The second class of models considered produces spin-1 colour-octet vector bosons. Specifically, heavy Kaluza-Klein gluons, $g_{KK}$, as produced in Randall-Sundrum (RS) models with a single warped extra dimension [33] are used as a benchmark in this search. In this model, the $g_{KK}$ has a nominal width of 15.3% of its mass. Previous searches using $\sqrt{s} = 7$ TeV ATLAS data [28] exclude a $g_{KK}$ with a mass less than 2.1 TeV. The CMS Collaboration searched for similar resonances [32], using a slightly different benchmark model [34]. The CMS choice leads to a larger natural width of 20% and a larger production cross-section, and, for such a scenario, they exclude the existence of a $g_{KK}$ with mass less than 2.5 TeV. In the analysis presented here, the sensitivity to the width of the colour octet is also tested for widths from 10% to 40% of the resonance mass.

2.3 Spin-2 colour singlet

The third class of models explored in this search produces spin-2 colour singlets, such as Kaluza-Klein excitations of the graviton, $G_{KK}$. The search uses a Randall-Sundrum model with extra dimensions where the SM fields are in the warped bulk and the fermions are localised appropriately to explain flavour [13, 35, 36]. This kind of graviton is commonly referred to as a “Bulk” RS graviton and is characterised by a dimensionless coupling constant $k/M_{Pl} \sim 1$, where $k$ is the curvature of the warped extra dimension and $M_{Pl} = M_{Pl}/\sqrt{8\pi}$ is the reduced Planck mass. For such gravitons, decays to light fermions are suppressed, and the branching ratio to photons is negligible. The branching ratios to $t\bar{t}$, WW, ZZ and HH are significant. In the model used, $k/M_{Pl}$ is chosen to be 1, the $G_{KK}$ width varies from 3% to 6% in the mass range 400–2000 GeV. The branching ratio of $G_{KK}$ decay into $t\bar{t}$ pair rapidly increases from 18% to 50% between 400 and 600 GeV, plateauing at 68% for masses larger than 1 TeV. There have been no previous direct searches for such gravitons in the $t\bar{t}$ decay channel. The ATLAS Collaboration used the same model to explore the $G_{KK} \rightarrow ZZ$ channel [37] and excluded Bulk RS $G_{KK}$ with mass less
than 740 GeV. The CMS Collaboration performed searches in the $G_{KK} \to ZZ$ and $G_{KK} \to WW$ decay channels \cite{38} but did not consider the case of Bulk RS gravitons with $k/M_{Pl} > 0.5$.

2.4 Spin-0 colour singlet

The last class of models examined here produces colour-singlet scalar particles via gluon-gluon fusion that decay to $t\bar{t}$. The approach previously adopted by the CMS Collaboration \cite{32} is followed by generating narrow scalar resonances and neglecting interference with the SM $t\bar{t}$ production. The CMS Collaboration excluded such resonances with production cross-sections greater than 0.8 pb and 0.3 pb for masses of 500 and 750 GeV, respectively. The specific case of this spin-0 model can not be representative of any meaningful theoretical model existing at the moment.

3 The ATLAS detector

The ATLAS experiment \cite{1} is a multipurpose particle physics detector with forward-backward symmetric cylindrical geometry\footnote{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 along 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 ($r, \phi$) are used in the transverse plane, \phi being the azimuthal angle around the beam pipe. The pseudorapidity is defined in terms of the polar angle $\theta$ as $\eta = -\ln \tan(\theta/2)$. The distance in $\eta$-$\phi$ space is commonly referred to as $\Delta R = \sqrt{(\Delta \phi)^2 + (\Delta \eta)^2}$.}. The inner detector (ID) consists of multiple layers of silicon pixel and strip detectors and a straw-tube transition radiation tracker and covers a pseudorapidity range of $|\eta| < 2.5$. The ID is surrounded by a superconducting solenoid that provides a 2 T magnetic field. The calorimeter system, surrounding the ID and the solenoid, covers the pseudorapidity range $|\eta| < 4.9$. It consists of high-granularity lead and liquid-argon (LAr) electromagnetic calorimeters, a steel and scintillating-tile hadronic calorimeter within $|\eta| = 1.7$ and two copper and LAr hadronic endcap calorimeters covering $1.5 < |\eta| < 3.2$. Forward copper/LAr and tungsten/LAr calorimeter modules complete the solid-angle coverage out to $|\eta| = 4.9$. The muon spectrometer (MS) resides outside the calorimeters. It consists of multiple layers of trigger and tracking chambers within a system of air-core toroids, which enables an independent, precise measurement of muon track momenta for $|\eta| < 2.7$. The muon trigger chambers cover $|\eta| < 2.4$.

4 Data and Monte Carlo samples

This search is performed in proton-proton collision data at $\sqrt{s} = 8$ TeV collected with the ATLAS detector in 2012. The data were only used if they were recorded during stable beam conditions and with all relevant subdetector systems operational. Lepton plus jets events were collected using single-electron and single-muon triggers with thresholds chosen in each case such that the efficiency is uniform for leptons passing offline selections including transverse momentum $p_T > 25$ GeV. The ATLAS muon trigger system suffers from a 20% inefficiency, relative to the offline event selection used in this analysis, largely due to a lack of geometrical coverage of muon chambers owing to support structures in those regions \cite{39}. To mitigate this loss of efficiency, a large-radius-jet ($R = 1.0$) trigger was also used to collect muon plus jets events which failed the muon trigger. This large-radius-jet trigger has recorded 17.4 fb$^{-1}$ data. The chosen
trigger threshold yields a uniform efficiency as a function of reconstructed jet transverse momentum, above 99%, for events containing a large-radius jet with $p_T > 380$ GeV. For $t\bar{t}$ events with invariant masses above 1.5 TeV, this addition increased the overall trigger efficiency in the muon channel to 96%.

Simulated Monte Carlo (MC) samples were used for signal processes, as well as background processes producing jets and prompt leptons. The MC samples are employed to develop the event selection, provide SM background estimates, and evaluate signal efficiencies. Background contributions from processes in which there are no genuine prompt isolated leptons were estimated directly from the data as described in Section 7. The MC samples were processed through the full ATLAS detector simulation [40] based on GEANT4 [41] or through a faster simulation making use of parametrised showers in the calorimeters [42]. Additional simulated proton-proton collisions generated using PYTHIA v8.1 [43] were overlaid to simulate the effects of additional collisions from the same and nearby bunch crossings (pile-up). All simulated events were then processed using the same reconstruction algorithms and analysis chain as the data. The simulated trigger and selection efficiencies were corrected to agree with the performance observed in data.

Production of Bulk RS gluon and graviton signals was modelled using MADGRAPH5 [44] interfaced with PYTHIA v8.1. For the gluon, the MSTW2008LO parton distribution function (PDF) set [45] was used, while for the graviton, the CTEQ6L1 [46] PDF set was used. The $Z'$ signal was modelled using PYTHIA v8.1 with the MSTW2008LO PDF set. Heavy scalar signal samples were generated using MADGRAPH_AMC@NLO [47] with LO matrix elements and the CTEQ6L1 PDF set.

Pair production of top quarks is the dominant background in this search. It was simulated using the POWHEG-BOX [48–51] generator r2330.3 interfaced with PYTHIA v6.427 [52] with the Perugia 2011C [53] tune and the CT10 [54] next-to-leading order PDF set. In the MC generation of POWHEG-BOX events, the parameter $hdamp$ was set to the top quark mass in order to achieve good agreement with the differential cross-section measurements [55]. Alternative samples of $t\bar{t}$ events, used to evaluate uncertainties on $t\bar{t}$ modelling, were generated using the POWHEG-Box and MC@NLO v4.1 [56–58] generators interfaced with HERWIG v6.5 [59,60] with JIMMY [61] for the modelling of the underlying event. For these samples the CT10 PDF set and the ATLAS-AUET2 [62] tune were used. In all cases, the top quark mass used was 172.5 GeV. The cross-sections for these samples were normalised to the calculation from Top++ v2.0 [63] at next-to-next-to-leading order (NNLO) accuracy in the strong coupling constant $\alpha_s$, including resummation of next-to-next-to-leading logarithmic (NNLL) soft gluon terms [64–69]. The top quark kinematics in all SM $t\bar{t}$ samples were corrected to account for electroweak higher-order effects [70].

Production of $W$ bosons in association with jets ($W$+jets) is also a significant background process. Samples of $W$+jets events were generated using the ALPGEN generator [71] interfaced with PYTHIA v6.246, including up to 5 extra partons in the matrix element. Configurations with additional heavy quarks (a single $c$-quark, a $c\bar{c}$ pair or a $b\bar{b}$ pair in the hard process) were included, with the masses of heavy quarks taken into account. The CTEQ6L1 PDF set and the Perugia 2011C tune were used. The samples were normalised using data as described in Section 7. Additional samples were generated with different choice of ALPGEN matching parameters in order to estimate modelling uncertainties on the production of $W$+jets events.

Production of single top quarks can yield events that pass the analysis event selection. The POWHEG-Box generator interfaced with PYTHIA v6.246 was used to estimate the $s$- and $Wt$-channels [72–74] with the same configuration as for the $t\bar{t}$ samples. The $t$-channel was also generated with POWHEG-Box but in a four-flavour scheme, hence the CT10 NLO four-flavour PDF set was used. Overlap between the $Wt$ sample and $t\bar{t}$ samples was handled using the diagram removal scheme [75]. These samples were normalised to approximate NNLO cross-sections [76–78].
Other minor background processes producing prompt isolated leptons include heavy diboson production, production of Z bosons in association with jets (Z+jets) and production of heavy gauge bosons in association with t\bar{t} (t\bar{t}V). Production of Z+jets was modelled using ALPGEN interfaced with PYTHIA v6.426, in the same configuration used for the W+jets samples described above. The samples were normalised to the inclusive Z boson production cross-section calculated at NNLO in QCD using FEWZ [79]. Diboson production was modelled using the SHERPA [80–83] generator, with up to three extra partons in the matrix element and taking into account the mass of the b- and c-quarks. The diboson samples were normalised to calculations at NLO in QCD performed using MCFM [84]. The t\bar{t}V production was modelled using MADGRAPH5 interfaced with PYTHIA v6.426 and normalised to NLO cross-section predictions [85].

5 Event selection

Events consistent with t\bar{t} decaying to a single charged lepton together with hadronic jets and missing transverse momentum are selected. Electron candidates are required to have a transverse energy \( E_T > 25 \) GeV and \(|\eta_{\text{cluster}}| < 2.47\), where \(|\eta_{\text{cluster}}|\) is the pseudorapidity of the cluster of energy deposited in the electromagnetic calorimeter, computed with respect to the center of ATLAS detector and matched to the candidate [86]. Electron candidates in the calorimeter transition region 1.37 < |\eta_{\text{cluster}}| < 1.52 are excluded.

In order to reduce backgrounds from non-prompt sources and hadrons showering with a high electromagnetic energy fraction in the calorimeter, electron isolation is imposed using a variable \( MI_{10} \) defined as the \( \sum_{\text{tracks}} p_T^{\text{track}} \) for all tracks (except the matched lepton track) with \( p_T > 1 \) GeV passing quality selection criteria, and within a cone of radius \( \Delta R < 10 \text{ GeV}/E_T [28,87] \) centred on the electron cluster. Electrons are defined to be isolated if \( MI_{10}/E_T < 0.05 \). The \( E_T \) here is the transverse energy of the reconstructed electron. The isolation variable \( MI_{10} \) is particularly useful in the case of boosted top quark decays since the \( p_T \)-dependent cone size reflects the \( p_T \) dependence of the separation between objects with a common boosted parent. The association of the electron to the collision vertex [88] is imposed by requiring that the longitudinal impact parameter relative to the collision vertex be less than 2 mm.

Muon candidates are required to have \( p_T > 25 \) GeV and \(|\eta| < 2.5\). The association of the muon to the collision vertex is imposed by the requirements that the longitudinal impact parameter relative to the collision vertex be less than 2 mm and that the transverse impact parameter relative to the collision vertex divided by its uncertainty, \(|d_0/\sigma_{d_0}|\), be less than 3. Muons are also required to satisfy the same \( MI_{10} \) requirement as electrons, with the cone centred on the inner-detector track associated with the muon.

Jets are reconstructed using the anti-\( k_T \) algorithm [89] applied to clusters of calorimeter cells that are topologically connected and calibrated to the hadronic energy scale [90] using a local calibration scheme [91]. Both small-radius (\( R = 0.4 \)) and large-radius jets (\( R = 1.0 \)) are used. Small-radius jets are required to satisfy \( p_T > 25 \) GeV and \(|\eta| < 2.5\), while large-radius jets are required to satisfy \( p_T > 300 \) GeV and \(|\eta| < 2.0\). Low-\( p_T \) central small-radius jets (\( p_T < 50 \text{ GeV}, |\eta| < 2.4\)) are required to have a jet vertex fraction [92] greater than 0.5. The jet vertex fraction is defined as the fraction of total transverse momentum (using a scalar sum) of tracks in the jet that are associated with the primary vertex. This variable suppresses jets arising from pileup effects. Large-radius jets have jet trimming [93] applied. In trimming, subjets are formed by applying a jet algorithm with smaller radius parameter, \( R_{\text{sub}} \), and then soft subjets with less than a certain fraction, \( f_{\text{cut}} \), of the original jet \( p_T \) are removed. The properties of the trimmed jet are then calculated using the surviving subjets. This procedure mitigates the effect of pile-up [94].
The trimming parameters used in this search are $f_{\text{cut}} = 0.05$ and $R_{\text{sub}} = 0.3$, and the inclusive $k_t$ [95] algorithm is used to form the subjets.

Only small-radius jets are considered for $b$-tagging. The $b$-tagging algorithm uses a multivariate approach with inputs taken from the results of separate impact parameter, secondary vertex and decay topology algorithms [96]. The operating point of the algorithm is chosen such that the $b$-tagging efficiency for simulated $t\bar{t}$ events is 70%. In MC simulation, factors are applied to correct for the differences between the $b$-tagging efficiency in simulated events and that measured in data. The factors are adapted to dense environments for which the $b$-tagging efficiencies are lower.

The $E_T^{\text{miss}}$ is calculated from the vector sum of the transverse energy of topological clusters in the calorimeter [97]. The clusters associated with the reconstructed electrons and small-radius jets are replaced by the calibrated energies of these objects. Muon transverse momenta determined from the ID and the MS are also included in the calculation.

Overlap in identification of the relevant physics objects is possible and a procedure is implemented to remove duplication. Electrons and small-radius jets are considered for overlap removal if the electron cluster is within $\Delta R = 0.4$ of the nearest jet. In such case, the jets have their four-momentum and jet vertex fraction recalculated by subtracting the electron four-momentum and then are removed if the recalculated values fail the original jet selection criteria. If the distance $\Delta R$ between the electron and the recalculated jet is < 0.2, the electron calorimeter cluster has a high chance to be from the hadronic jet. Therefore the electron is removed from the electron candidate list and its four-momentum is added to that of the recalculated jet. Muons are removed from the muon candidate list if the distance $\Delta R$ between the muon and small-radius jet is less than 0.04 + 10 GeV/$p_T$. This criteria exploits the anti-correlation between the muon $p_T$ and its angular distance from the $b$-quark, in a similar approach as the isolation variable. The parameters are tuned based on signal MC in order to provide constantly high efficiency as a function of resonance mass.

A set of common preselection criteria is used for events to be considered for the boosted and resolved topologies. Events are required to have exactly one lepton (electron or muon) plus multiple jets. Events recorded by the lepton triggers are required to have lepton-trigger objects that match the selected lepton. Additionally, events must have $E_T^{\text{miss}} > 20$ GeV and $E_T^{\text{miss}} + m_T > 60$ GeV, where $m_T$ is the transverse mass calculated as $m_T = \sqrt{2p_T E_T^{\text{miss}}(1 - \cos \phi_{lT})}$ where $p_T$ is the transverse momentum of the lepton and $\phi_{lT}$ is the angle between the $p_T$ and $E_T^{\text{miss}}$ vectors.

Events are next checked against the boosted-topology selection. The selected lepton is required to have at least one small-radius jet within a distance of $\Delta R(l, j) < 1.5$ and, of these, the jet with highest-$p_T$ is termed $j_{\text{sel}}$. Boosted-topology events must have at least one large-radius jet with $p_T > 300$ GeV (380 GeV for the muon plus jets events selected by the large-radius-jet trigger), $|\eta| < 2.0$, mass $m_{\text{jet}} > 100$ GeV, first $k_t$ splitting scale [95] $\sqrt{d_{l2}} > 40$ GeV, $\Delta R$ between the large-radius jet and $j_{\text{sel}} > 1.5$, and $\Delta\phi$ between the large-radius jet and lepton greater than 2.3. The jet mass is calculated using the four-momenta of its constituent clusters, which are taken as massless. If multiple large-radius jets satisfy these criteria, the highest-$p_T$ jet is chosen as the hadronic top quark candidate. Finally, at least one of the small-radius jets in the event must be $b$-tagged and matched to either top candidates, as will be described in Section 6.

Events failing the boosted-topology selection are then tested against the criteria for the resolved-topology selection. These events are required to have at least four small-radius jets, with at least one of them $b$-tagged. A $\chi^2$ algorithm is used to reconstruct the $t\bar{t}$ system, as will be described in Section 6. The lowest $\chi^2$ value is required to satisfy $\log_{10}(\chi^2) < 0.9$. 

7
The efficiencies of the boosted and resolved selections for MC simulated signal events are given in Figure 1, for different models of interest. The branching ratio for $t\bar{t}$ to electron or muon plus jets is included, which is 17% for each lepton flavour respectively taking into account $\tau$ decays to $e/\mu$ plus neutral particles [98]. There are efficiency losses from both the large-radius jet requirements and $b$-tagging requirement for the boosted-topology selection. Above resonance masses of 1.5 TeV, the efficiencies of the resolved selections are relatively insignificant due to the $\chi^2$ requirements and the veto of boosted selections. It can also be seen that efficiency times acceptance is smaller for isolated electrons than isolated muons above the same resonance mass point, due to the inefficiency of electron identification and overlap removal in the boosted environment.

Figure 1: Selection efficiency times acceptance times branching ratio as a function of the true $m_{t\bar{t}}$ for the different signals in the models considered: (a) $Z'$, (b) $g_{KK}$, (c) $G_{KK}$, (d) scalar resonance. The dashed lines show the boosted-topology selection and the unbroken lines show the combined selection. The branching ratio for $t\bar{t} \rightarrow e/\mu +$ jets is 17%.

6 Event reconstruction

Signal $t\bar{t}$ resonances should appear in the $m_{t\bar{t}}^{\text{reco}}$ spectrum as an excess of events over the SM expectation clustered around the resonance mass. Events are reconstructed assuming the final state originated from a
Reconstructible events are those where there is a reconstructed object within $\Delta R = 0.4$ of each visible parton from top decays, and the $\Delta \phi$ between the neutrino and the $E_{T}^{\text{miss}}$ is smaller than 1. Among events passing the final selection, 70% (55%) are reconstructible for $Z'$ mass of 750 GeV (400 GeV).

1. $t\bar{t}$ decay. To calculate $m_{t\bar{t}}^{\text{reco}}$, the neutrino four-momentum must be determined. The neutrino transverse momentum is taken to be the $E_{T}^{\text{miss}}$ vector. The longitudinal component of the neutrino momentum, $p_{\nu}$, is calculated by constraining the lepton plus missing transverse momentum system to have the $W$ boson mass and solving the resulting quadratic equation in the neutrino’s longitudinal momentum $p_{\nu}$ [99, 100]. If no real solution exists, the vector is varied by the minimal amount required to produce exactly one real solution. If two real solutions are found, the one with the smallest $|p_{\nu}|$ is used for the boosted-topology reconstruction, while the choice is made by the $\chi^2$ algorithm described below for the resolved-topology.

For the boosted $m_{t\bar{t}}^{\text{reco}}$ reconstruction, $m_{t\bar{t}}^{\text{reco}}$ is computed from the four-momenta of the neutrino, lepton, the previously selected small-radius jet, $j_{\text{sel}}$, and the large-radius jet. In this case the assignment of jets to the semileptonically decaying top quark and hadronically decaying top quark is unambiguous.

For the resolved $m_{t\bar{t}}^{\text{reco}}$ reconstruction, a $\chi^2$ algorithm is employed to find the best assignment of jets to the semileptonically and hadronically decaying top quarks. Using the four-momenta of the neutrino, lepton and jets, a $\chi^2$ is defined using the expected top quark and $W$ boson masses to be

$$
\chi^2 = \frac{m_{jj} - m_{W}}{\sigma_{W}}^2 + \frac{m_{jjb} - m_{jj} - m_{t_{h}} - m_{W}}{\sigma_{t_{h} - W}}^2 + \frac{m_{j\ell\nu} - m_{t_{\ell}}}{\sigma_{t_{\ell}}}^2 + \frac{(p_{T,jj} - p_{T,j\ell\nu}) - (p_{T,t_{h}} - p_{T,t_{\ell}})}{\sigma_{\text{diff}p_{T}}}^2. 
$$

The first term is a constraint using the hadronically decaying $W$ boson mass. The second term is a constraint using the mass difference between the hadronically decaying top quark and the hadronically decaying $W$ boson. Since the hadronically decaying $W$ boson mass, $m_{jj}$, and the hadronically decaying top quark mass, $m_{jjb}$, are highly correlated, the mass of the hadronically decaying $W$ boson is subtracted from the second term so as to decouple it from the first term. The third term is a constraint using the semileptonically decaying top quark mass. The last term arises as a constraint on the expected transverse momentum balance between the two decaying top quarks. In the $\chi^2$ definition above, $t_{h}$ and $t_{\ell}$ refer to the hadronically and semileptonically decaying top quarks. The values of the $\chi^2$ central parameters $m_{W}$, $m_{t_{h} - W}$, $m_{t_{\ell}}$ and $p_{T,t_{h}} - p_{T,t_{\ell}}$, and the values of the width parameters $\sigma_{W}$, $\sigma_{t_{h} - W}$, $\sigma_{t_{\ell}}$ and $\sigma_{\text{diff}p_{T}}$ are found through Gaussian fits to the distributions of relevant reconstructed variables in MC simulations of $Z'$ production with masses from 0.5 to 2.0 TeV in which reconstructed objects can be matched with the partons from the hard process.

In determining the lowest $\chi^2$, all possible jet permutations are tried as well as all neutrino $p_{\nu}$ solutions. The number of jet permutations is reduced by assigning any $b$-tagged jets to the $b$ partons in the top quark decays. Based on studies using MC simulations of $Z' \to t\bar{t}$ events, the permutation with the lowest $\chi^2$ results in the correct assignment of reconstructed objects to the hadronically and semileptonically decaying top quark in 80% of reconstructible events.

The resulting $m_{t\bar{t}}^{\text{reco}}$ distributions for several signal masses from the resolved- and boosted-topology reconstruction are shown in Figures 2 and 3. For these figures, all events satisfying the resolved- or boosted-topology selection criteria are used. The low mass tails arise from two effects: firstly, extra radiation from the $t\bar{t}$ system that is not included in the reconstruction can shift the reconstructed mass to lower values; secondly, before reconstruction the Breit-Wigner signal shape in $m_{T}$ has a tail at lower values due to the steep fall in parton luminosity with increasing partonic centre-of-mass energy. The former is
particularly true for high mass resonances, while the latter has larger effect with broad resonances. The experimental resolution for the invariant mass of the $t\bar{t}$ system is 8% (6%) for the resolved-topology selection at a resonance mass of 400 GeV (1 TeV), and 6% in the boosted-topology selection.

With hadronically and semileptonically decaying top quarks identified for both boosted- and resolved-topology selections, three categories of $b$-tagged events are defined: those in which both decaying top quark candidates have a matching $b$-jet, those in which only the hadronically decaying top quark candidate has a matching $b$-jet and those in which only the semileptonically decaying top quark candidate has a matching $b$-jet. In the boosted-topology selection, the match is defined as either the selected jet, $j_{\text{sel}}$, or a small-radius jet within $\Delta R < 1$ of the large-radius jet as being a $b$-tagged jet. In the resolved-topology selection, the matching is determined by the $\chi^2$ algorithm.

Figure 2: $m_{ij}$ for the different signal models for events satisfying the resolved-topology selection and using the resolved-topology reconstruction: (a) $Z'$, (b) $g_{KK}$, width 15.3%, (c) $G_{KK}$, (d) scalar resonance. The label ‘inclusive’ indicates that for this figure events that pass the resolved-topology selection are included regardless of whether they pass the boosted-topology selection.

---

3 The experimental resolution is extracted from a Gaussian fit of the relative difference $m_{ij}^{\text{reco,true-matched}} - m_{ij}^{\text{reco}}$ from reconstructible events; $m_{ij}^{\text{reco,true-matched}}$ is the reconstructed $m_{ij}^{\text{reco}}$ computed with the right combination of jets identified with the parton level information and $m_{ij}$ being the true mass of the $t\bar{t}$ system.
Figure 3: Reconstructed $t\bar{t}$ invariant mass for the different signal models for events satisfying the boosted-topology selection and using the boosted-topology reconstruction: (a) $Z'$, (b) $g_{KK}$, width 15.3%, (c) $G_{KK}$, (d) scalar.

7 Background contributions estimated from data

Data are used to estimate the magnitudes and uncertainties of two important background contributions: $W$ + jets and multi-jet production. For the former, data are used to derive scale factors that are applied to correct the normalisation and flavour fractions of the $W$+jets background given by 

\begin{align}
N_{W^+} + N_{W^-} = \left( \rho_{MC} + 1 \right) \left( D_{corr^+} - D_{corr^-} \right),
\end{align}

where $\rho_{MC}$ is the ratio of the measured $W$ boson charge asymmetry in data to that predicted by Alpgen MC simulation.
where $r_{MC}$ is the ratio given by MC simulation of the number of $W$+jets events with a positively-charged lepton to that with a negatively-charged lepton and $D_{corr(+)}$ is the number of observed events with a positively-charged (negatively-charged) lepton. Contributions to $D_{corr(+)}$ from charge-asymmetric processes such as single top, diboson and $t\bar{t}+W$ boson production are estimated from MC simulation and are subtracted. Contributions from charge-symmetric processes such as $t\bar{t}$ production cancel in the difference on the right-hand side of Equation 2. The scale factor, $C_{\lambda}$, applied to the MC simulated samples of $W$ + jets events is then calculated as the ratio of the $N_{W^+} + N_{W^-}$ evaluated from data to that predicted from MC simulation.

The value and statistical uncertainty obtained for $C_{\lambda}$ in the electron (muon) channel are $1.026 \pm 0.011$ ($0.978 \pm 0.010$) with the resolved selection, and $0.89 \pm 0.06$ ($0.81 \pm 0.05$) with the boosted selection. While the $C_{\lambda}$ scale factors determined are close to unity within statistical uncertainties, this method reduces the systematic uncertainties (from the jet energy scale, $b$-tagging and other uncertainties) compared to using $W$+jets MC simulation alone. Systematic uncertainties in the $W$+jets normalisation and flavour fraction corrections are determined by rederiving these scale factors when a given systematic effect is applied. The new scale factors are then used in producing the $m_{rec}$ mass spectrum for that particular systematic uncertainty.

Scale factors for the relative fraction of heavy-flavour contributions from $W + b\bar{b}$, $W + c\bar{c}$, $W + c$ and $W$+light quark flavours are first determined from MC simulation for this sample. The relative fraction of the $W + b\bar{b}$ to $W + c\bar{c}$ contribution is taken from MC simulation and fixed at that value. A system of three equations is used to fit to the two-jet data sample with at least one $b$-tagged jet in order to determine correction factors for each of the flavour fractions determined from MC simulation:

$$\begin{align*}
\left( C_{\lambda} \cdot \left( N_{b\bar{b}}^{W^+} + N_{c\bar{c}}^{W^-} \right) + C_{e} \cdot N_{c}^{W^-} + C_{light} \cdot N_{light}^{W^-} \right) & = \left( f_{b\bar{b}} + f_{c\bar{c}} \right) N_{MC}^{W^-} \\
\left( C_{\lambda} \cdot \left( N_{b\bar{b}}^{W^+} + N_{c\bar{c}}^{W^-} \right) + C_{e} \cdot N_{c}^{W^+} + C_{light} \cdot N_{light}^{W^+} \right) & = \left( f_{b\bar{b}} + f_{c\bar{c}} \right) N_{MC}^{W^+} \\
\left( C_{\lambda} \cdot \left( N_{b\bar{b}}^{W^+} + N_{c\bar{c}}^{W^-} \right) + C_{e} \cdot N_{c}^{W^+} + C_{light} \cdot N_{light}^{W^+} \right) & = \left( f_{b\bar{b}} + f_{c\bar{c}} \right) N_{MC}^{W^+}
\end{align*}$$

where $D_{W^+}$ is the expected number of $W$+jets events with a positively-charged or negatively-charged lepton in the data. The flavour fraction is $f_{flavour}$, and the correction factor for a given flavour component is $K_{flavour}$. The different flavour labels are $b\bar{b}$, $c\bar{c}$, $c$, and $light$ corresponding to $W + b\bar{b}$, $W + c\bar{c}$, $W + c$, and $W$+light jets respectively. The numbers of positively-charged and negatively-charged leptons in the data are found by subtracting all non-$W$-jets contributions, which are determined from MC simulations as 35% (15%) of the selected events for the electron (muon) channel. An iterative process is used to find the $K_{flavour}$ factors, which are then used to correct the corresponding flavour fractions $f_{flavour}$ that are applied during the $C_{\lambda}$ factor calculation. In this interactive process, only the $K_{flavour}$ and $C_{\lambda}$ factors are allowed to vary. The $K_{flavour}$ factors are initially set to unity, thus altering the $C_{\lambda}$ factor calculation. New correction factors $K_{flavour}$ are calculated by inverting the above equation, and then the process is repeated ten times, with each repetition using the correction factors determined from the previous one. It was checked that using more than ten iterations produced negligible changes in the extracted correction factors.

The correction factors found for the two-jet sample are extrapolated to higher jet multiplicities by keeping the same relative ratio between them while conserving the normalisation in each jet multiplicity bin. The $K_{flavour}$ factors thus obtained are different from unity. For events containing electrons (muons), the
extracted values and statistical uncertainties of $K_{b\bar{b}}$ and $K_{c\bar{c}}$ are $1.36 \pm 0.07 (1.51 \pm 0.08)$, $K_c$ is $0.71 \pm 0.03 (0.66 \pm 0.026)$, and $K_{light}$ is $0.934 \pm 0.005 (0.873 \pm 0.004)$.

The multi-jet background in events passing the resolved or boosted selection criteria consists of events with a jet that is misreconstructed as a lepton or with non-prompt leptons that pass the identification criteria. The normalisation, $m_{T\bar{T}}^{\text{reco}}$ shape, and statistical and systematic uncertainties associated with the multi-jet background are estimated from data using a matrix method \cite{103,105}.

The multi-jet background in events passing the resolved or boosted selection criteria is estimated using data events that pass all selection criteria, except that the loose lepton identification criteria are used. This sample contains both prompt and non-prompt leptons.

The number of events with leptons passing the loose identification criteria, $N_L$ is defined as

$$N_L = N_{\text{prompt}} + N_{\text{multi-jet}}$$

where $N_{\text{prompt}}$ is the number of events with prompt leptons passing the loose identification criteria and $N_{\text{multi-jet}}$ is the number of events passing the loose identification criteria with leptons from other sources. The number of events passing the tight identification criteria, $N_T$ is then

$$N_T = \epsilon \times N_{\text{prompt}} + f \times N_{\text{multi-jet}}.$$  

Solving these two equations for $N_{\text{prompt}}$ and $N_{\text{multi-jet}}$ gives the multi-jet contribution from events passing all the selection criteria.

Good shape modelling of the $m_{T\bar{T}}^{\text{reco}}$ distributions has been achieved by parameterising the efficiencies as functions of relevant kinematic variables, and validated in the multi-jet control regions. Systematic uncertainties are evaluated using several different definitions of multi-jet control regions that result in slightly different $f$ estimations. The impact from systematic uncertainties associated with object reconstruction and MC simulation are also considered, resulting in a total normalisation uncertainty of 20%.

8 Systematic uncertainties

The systematic uncertainties can be broadly divided into two categories: uncertainties that affect reconstructed objects (such as jets) and uncertainties that affect the modelling of certain background or signal processes. Some of the uncertainties affect both the shape and the normalisation of the $m_{T\bar{T}}^{\text{reco}}$ spectra, while others affect the normalisation only. In Table 1, an overview of the effects of the dominant systematic uncertainties on the background and signal yields is given. Only the impact on the overall normalisation
The average impact of the dominant systematic uncertainties on the total background yield and on the estimated yield for a $Z'$ sample with $m = 1.75$ TeV. The electron and muon channel spectra are added. The shift is given in percent of the nominal value. Certain systematic uncertainties are not applicable to the $Z'$ samples, which is indicated with a bar (−) in the table.

Table 1: Average impact of the dominant systematic uncertainties on the total background yield and on the estimated yield for a $Z'$ sample with $m = 1.75$ TeV. The electron and muon channel spectra are added. The shift is given in percent of the nominal value. Certain systematic uncertainties are not applicable to the $Z'$ samples, which is indicated with a bar (−) in the table.

<table>
<thead>
<tr>
<th>Systematic Uncertainties</th>
<th>Resolved selection yield impact [%]</th>
<th>Boosted selection yield impact [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>total bkg.</td>
<td>$Z'$</td>
</tr>
<tr>
<td>Luminosity</td>
<td>2.5</td>
<td>2.8</td>
</tr>
<tr>
<td>PDF</td>
<td>2.4</td>
<td>3.6</td>
</tr>
<tr>
<td>ISR/FSR</td>
<td>3.7</td>
<td>−</td>
</tr>
<tr>
<td>Parton shower and fragmentation</td>
<td>4.8</td>
<td>−</td>
</tr>
<tr>
<td>$t\bar{t}$ normalisation</td>
<td>5.3</td>
<td>−</td>
</tr>
<tr>
<td>$t\bar{t}$ EW virtual correction</td>
<td>0.2</td>
<td>−</td>
</tr>
<tr>
<td>$t\bar{t}$ Generator</td>
<td>0.3</td>
<td>−</td>
</tr>
<tr>
<td>$t\bar{t}$ Top quark mass</td>
<td>0.6</td>
<td>−</td>
</tr>
<tr>
<td>W+jets Generator</td>
<td>0.3</td>
<td>−</td>
</tr>
<tr>
<td>Multi-jet norm, $e+\text{jets}$</td>
<td>0.5</td>
<td>−</td>
</tr>
<tr>
<td>Multi-jet norm, $\mu+\text{jets}$</td>
<td>0.1</td>
<td>−</td>
</tr>
<tr>
<td>JES+JMS, large-radius jets</td>
<td>0.1</td>
<td>2.1</td>
</tr>
<tr>
<td>JER+JMR, large-radius jets</td>
<td>&lt; 0.1</td>
<td>0.2</td>
</tr>
<tr>
<td>JES, small-radius jets</td>
<td>5.6</td>
<td>2.6</td>
</tr>
<tr>
<td>JER, small-radius jets</td>
<td>1.8</td>
<td>1.4</td>
</tr>
<tr>
<td>Jet vertex fraction</td>
<td>0.8</td>
<td>0.8</td>
</tr>
<tr>
<td>$b$-tagging $b$-jet efficiency</td>
<td>1.1</td>
<td>2.0</td>
</tr>
<tr>
<td>$b$-tagging $c$-jet efficiency</td>
<td>0.1</td>
<td>0.7</td>
</tr>
<tr>
<td>$b$-tagging light-jet efficiency</td>
<td>&lt; 0.1</td>
<td>&lt; 0.1</td>
</tr>
<tr>
<td>Electron efficiency</td>
<td>0.3</td>
<td>0.6</td>
</tr>
<tr>
<td>Muon efficiency</td>
<td>0.9</td>
<td>1.0</td>
</tr>
<tr>
<td>MC statistical uncertainty</td>
<td>0.4</td>
<td>6.0</td>
</tr>
<tr>
<td>All systematic uncertainties</td>
<td>10.8</td>
<td>8.8</td>
</tr>
</tbody>
</table>

is shown in the table, but some of the systematic uncertainties have a significant dependence on the reconstructed $t\bar{t}$ mass, which is fully taken into account in the analysis.

The dominant uncertainty on the normalisation of the total background estimate is the NNLO+NNLL $t\bar{t}$ cross-section uncertainty of 6.5%. This uncertainty includes renormalisation and factorisation scale uncertainties, combined PDF and strong-coupling uncertainties evaluated following the PDF4LHC [106] recommendations, and uncertainties associated with the value of the top quark mass. The combined PDF and strong-coupling uncertainties are extracted for each of the three PDF sets: MSTW2008 68% confidence level (CL) NNLO [45, 107], CT10 NNLO [54, 108] and NNPDF2.3 NNLO [109]; the total uncertainty associated with the PDF and strong-coupling uncertainties is one half of the size of the envelope of the three resultant error bands, with the central prediction being the midpoint of the envelope. Variations from changing the top quark mass by ±1.0 GeV are added in quadrature to the scale uncertainties, and combined PDF and strong-coupling uncertainties.

The $W+$jets normalisation, determined using data, is separately evaluated for each experimental source of systematic uncertainty. An additional systematic uncertainty on the prediction is evaluated by using the simulated samples generated with varied ALPGEN matching parameters. The normalisation uncertainty on the multi-jet background is 20%, as described in Section 7.
The single top quark background normalisation uncertainty is 7.7% [76–78]. The normalisation uncertainty on the Z+jets sample is 48%, estimated using Berends-Giele-scaling [110]. The diboson normalisation uncertainty is 34%, which is a combination of the NLO PDF and scale uncertainties, and additional uncertainties from the requirements on the jet multiplicity.

The luminosity uncertainty is 2.8% [111]. It is applied to all signal and background samples except multi-jet and W+jets, which are estimated from data.

The PDF uncertainty on all MC samples is estimated by taking the envelope of the MSTW2008NLO, NNPDF2.3 and CT10 PDF set uncertainties at 68% CL following the PDF4LHC recommendation and normalising to the nominal cross-section. The PDF uncertainty on the \( t\bar{t} \) mass spectrum has a much larger effect on the boosted sample than on the resolved sample. The effect on the total background yield is 2.1% (4.2%) after the resolved (boosted) selection. The size of the uncertainty grows with reconstructed mass attaining values of 20% (20–50%) above 2 TeV in the resolved (boosted) selection. The relatively small uncertainty on the resolved selection at high \( m_{t\bar{t}} \) arises because there is a larger contribution than in the boosted selection from events that migrate upwards from low \( m_{t\bar{t}} \), correlated with low incoming fractional momentum carried by the parton where the PDF uncertainty is smaller.

One of the dominant uncertainties affecting reconstructed objects is the jet energy scale (JES) uncertainty, especially for large-radius jets [112,113]. Uncertainties on the jet mass scale (JMS) and the \( k_t \) splitting scales [113] are also important for this analysis. These uncertainties have an impact of 10% on the overall background yield in the boosted selection. The impact on the background estimates falls with increasing \( m_{t\bar{t}}^{\text{reco}} \), varying from 22% at lowest masses to about 7% from 1.5 TeV. The impact is smaller for the resolved selection, since the large-radius jets are only used indirectly there via the vetoing of events that pass the boosted selection. For large-radius jets, uncertainties on the jet energy resolution (JER) and jet mass resolution (JMR) are also considered. These are less significant than the JES. For small-radius jets, the uncertainties on the JES, the jet reconstruction efficiency and the jet energy resolution are considered [90,114]. The small-radius JES uncertainty is one of the most significant systematic uncertainties in the resolved selection, changing the overall expected yield by 6%. The effect of uncertainties associated with the jet vertex fraction is also considered. The \( b \)-tagging uncertainty is modelled through simultaneous variations of the uncertainties on the efficiency and rejection [115,116]. An additional \( b \)-tagging uncertainty is applied for high-momentum jets (\( p_T > 300 \text{ GeV} \)) to account for uncertainties on the modelling of the track reconstruction in high-\( p_T \) environments, which is one of the dominant uncertainties for high mass signals.

For the leptons, the uncertainties on the isolation efficiency, the single lepton trigger and the reconstruction efficiency are estimated using \( Z \to ee \) and \( Z \to \mu\mu \) events. In addition, high-jet-multiplicity \( Z \to \ell\ell \) events are studied, from which extra uncertainties on the isolation efficiency are assigned to account for the difference between \( Z \) and \( t\bar{t} \) events. Uncertainties on the \( E_T^{\text{miss}} \) reconstruction, as well as on the energy scale and resolution of the leptons, are also considered, and generally have a smaller impact on the yield and the search sensitivity than the uncertainties mentioned above.

The uncertainty on the \( t\bar{t} \) background due to uncertainties on the modelling of QCD initial- and final-state radiation (ISR/FSR) is estimated using AcerMC v3.8 [117] plus Pythia v6.426 MC samples by varying the Pythia ISR and FSR parameters within ranges allowed by a previous ATLAS measurement
of $t\bar{t}$ production with a veto on additional central jet activity [118]. The QED ISR/FSR uncertainty is negligible at this level. The dependency of the $m^{\text{reco}}_{t\bar{t}}$ shape on the choice of NLO generator is accounted for by using the difference between samples generated with MC@NLO and Powheg-Box+Herwig as a systematic uncertainty. The parton showering and fragmentation uncertainty on the $t\bar{t}$ background is estimated by comparing the result from a sample generated with Powheg-Box [50], interfaced with Pythia or Herwig. The uncertainty on the $m^{\text{reco}}_{t\bar{t}}$ distribution arising from the uncertainty in the top quark mass is evaluated by comparing the $m^{\text{reco}}_{t\bar{t}}$ spectrum using the nominal sample to those generated with top quark masses of 170 and 175 GeV, and multiplying the difference by 0.4 (to approximate a one standard deviation uncertainty, corresponding to ±1.0 GeV). The uncertainty on the electroweak corrections to top quark pair production is modelled by changing the difference of each correction factor from unity by ±10%.

For the $W+\text{jets}$ background, the uncertainty on the $m^{\text{reco}}_{t\bar{t}}$ distribution is estimated by reweighting the events to the kinematics of MC samples generated with different matching scale or functional form of the factorization scale [71].

9 Comparison of data with expected background contributions

After all event selection criteria are applied, 223,330 events are selected for the resolved topology and 8,206 events remain for the boosted topology. The event yields from data and from expected background processes are listed in Table 2 together with the associated systematic uncertainties.

Good agreement is observed between the data and the total expected background. Figure 4 shows the reconstructed mass of the semileptonically and hadronically decaying top quark candidates, and the mass of the hadronically decaying $W$-boson candidate for the resolved selection. The equivalent distributions for the boosted selection are the mass of the semileptonically decaying top quark candidate and the mass of the large-radius jet; both are shown in Figure 5. Figure 6 shows the distribution of the transverse momentum and first $k_t$-splitting scale of the selected large-radius jets.

The $t\bar{t}$ invariant mass spectra for the resolved and the boosted selections in the electron and muon channels, separated by $b$-tagging category, are shown in Figures 7 and 8. Figure 9 shows the $t\bar{t}$ invariant mass spectrum for all channels added together. The data generally agree with the expected background with slight shape deviations seen especially in the high mass regions; these deviations are consistent with the uncertainties.
Figure 4: Reconstructed mass of the hadronically decaying top quark candidate, $m_{t_h}$, semileptonically decaying top quark candidate, $m_{t_l}$, and hadronically decaying $W$-boson candidate $m_W$ after the resolved-topology selection in the electron and muon channels. The SM background components are shown as stacked histograms. The shaded areas indicate the total systematic uncertainties. Some background sources are too small to be visible in the figure.
Table 2: Data and expected background event yields after the resolved and boosted selections. The quadratic sum of all systematic uncertainties on the expected background yields is also given.

<table>
<thead>
<tr>
<th>Type</th>
<th>Resolved-topology selection</th>
<th>Boosted-topology selection</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>e+jets</td>
<td>μ+jets</td>
</tr>
<tr>
<td>$t\bar{t}$</td>
<td>93,000 ± 11,000</td>
<td>91,000 ± 11,000</td>
</tr>
<tr>
<td>Single top</td>
<td>3,800 ± 400</td>
<td>3,800 ± 400</td>
</tr>
<tr>
<td>$t\bar{t}V$</td>
<td>274 ± 13</td>
<td>267 ± 13</td>
</tr>
<tr>
<td>Multi-jet $e$</td>
<td>5,300 ± 1,100</td>
<td>–</td>
</tr>
<tr>
<td>Multi-jet $\mu$</td>
<td>–</td>
<td>1,050 ± 240</td>
</tr>
<tr>
<td>$W$+jets</td>
<td>6,600 ± 800</td>
<td>7,100 ± 800</td>
</tr>
<tr>
<td>$Z$+jets</td>
<td>1,400 ± 330</td>
<td>650 ± 130</td>
</tr>
<tr>
<td>Di-bosons</td>
<td>320 ± 50</td>
<td>310 ± 50</td>
</tr>
<tr>
<td>Total</td>
<td>110,000 ± 12,000</td>
<td>105,000 ± 12,000</td>
</tr>
<tr>
<td>Data</td>
<td>114,377</td>
<td>108,953</td>
</tr>
<tr>
<td></td>
<td>e+jets</td>
<td>μ+jets</td>
</tr>
<tr>
<td>$t\bar{t}$</td>
<td>4,100 ± 600</td>
<td>4,000 ± 600</td>
</tr>
<tr>
<td>Single top</td>
<td>138 ± 20</td>
<td>154 ± 20</td>
</tr>
<tr>
<td>$t\bar{t}V$</td>
<td>37 ± 4</td>
<td>38 ± 4</td>
</tr>
<tr>
<td>Multi-jet $e$</td>
<td>91 ± 18</td>
<td>–</td>
</tr>
<tr>
<td>Multi-jet $\mu$</td>
<td>–</td>
<td>8.6 ± 1.6</td>
</tr>
<tr>
<td>$W$+jets</td>
<td>260 ± 50</td>
<td>290 ± 50</td>
</tr>
<tr>
<td>$Z$+jets</td>
<td>31 ± 6</td>
<td>17 ± 5</td>
</tr>
<tr>
<td>Di-bosons</td>
<td>21 ± 4</td>
<td>20 ± 4</td>
</tr>
<tr>
<td>Total</td>
<td>4,700 ± 600</td>
<td>4,500 ± 600</td>
</tr>
<tr>
<td>Data</td>
<td>4,148</td>
<td>4,058</td>
</tr>
</tbody>
</table>
Figure 5: The invariant mass of the large-radius jets and the invariant mass of the semileptonically decaying top quark candidate, $m_{t\ell}$, after the boosted selection. The SM background components are shown as stacked histograms. The shaded areas indicate the total systematic uncertainties. Some background sources are too small to be visible in the figure.
Figure 6: The transverse momentum, $p_T$, and first $k_t$ splitting scale, $\sqrt{d_{12}}$, of the large-radius jet after the boosted selection. The SM background components are shown as stacked histograms. The shaded areas indicate the total systematic uncertainties. Some background sources are too small to be visible in the figure.
Figure 7: The $m^{\text{reco}}$ spectrum for the different channels, before any nuisance parameter fit, after the resolved-topology selection. The SM background components are shown as stacked histograms. The shaded areas indicate the total systematic uncertainties. The red line shows the expected distribution for a hypothetical $g_{KK}$ of mass 2.0 TeV, width 15.3%.
The systematic uncertainties. The red line shows the expected distribution for a hypothetical selection. The SM background components are shown as stacked histograms. The shaded areas indicate the total systematic uncertainties. The red line shows the expected distribution for a hypothetical $g_{KK}$ of mass 2.0 TeV, width 15.3%.

Figure 8: The $m_{t\bar{t}}^{reco}$ spectrum for the different channels, before any nuisance parameter fit, after the boosted-topology selection. The SM background components are shown as stacked histograms. The shaded areas indicate the total systematic uncertainties. The red line shows the expected distribution for a hypothetical $g_{KK}$ of mass 2.0 TeV, width 15.3%.

(a) Electron channel, boosted selection, both top candidates have $b$-tagged jets.

(b) Muon channel, boosted selection, both top candidates have $b$-tagged jets.

(c) Electron channel, boosted selection, only the hadronic top candidate has a $b$-tagged jet.

(d) Muon channel, boosted selection, only the hadronic top candidate has a $b$-tagged jet.

(e) Electron channel, boosted selection, only the semileptonic top candidate has a $b$-tagged jet.

(f) Muon channel, boosted selection, only the semileptonic top candidate has a $b$-tagged jet.
Figure 9: The $m_{t\bar{t}}^{\text{reco}}$ distributions, before any nuisance parameter fit, summed over (a) all 6 boosted channels, (b) all 6 resolved channels, and (c) all 12 channels compared with data. The SM background components are shown as stacked histograms. The shaded areas indicate the total systematic uncertainties. The red line shows the expected distribution for a hypothetical $g_{KK}$ of mass 2.0 TeV, width 15.3%.
10 Results

The final discriminating observables are the twelve $t\bar{t}$ invariant mass spectra: three $b$-tag categories for two selections and two decay channels. After the reconstruction of the $t\bar{t}$ mass spectra, the data and expected background distributions are compared using BumpHunter [119], which is a hypothesis testing tool that searches the data for local excesses or deficits compared to the expected background, taking the look-elsewhere effect into account over the full mass spectrum. The search is performed on three combinations of the spectra: the six channels of the resolved selections, the six channels of the boosted selections and the 12 channels. The most significant deviations of each spectra are requested to appear at the same place in the different channels of a combination. After accounting for the systematic uncertainties, no deviation from the total expected background is found. Upper limits are set on the cross-section times branching ratio for the studied signal models using a profile likelihood ratio test. The CL$_S$ prescription [120] is used to derive one-sided 95% confidence level (CL) limits. The results are obtained using the HistFitter [121] framework with all spectra from the 12 channels, excluding low statistics bins with $m_{t\bar{t}}^{\text{reco}}$ below 400 GeV in the boosted channels or above 2 TeV in the resolved channels.

The statistical and systematic uncertainties on the expected distributions are included in this CL$_S$ procedure as nuisance parameters in the likelihood fits. The nuisance parameters for the systematic uncertainties are constrained by a Gaussian probability density function with a width corresponding to the size of the uncertainty considered. Correlations between different channels and bins are taken into account. The product of the various probability density functions forms the likelihood function that is maximised in the fit by adjusting the free parameter (the signal strength) and nuisance parameters. The resulting expected $m_{t\bar{t}}^{\text{reco}}$ distributions are compared to data in Figure 10 for a fit of nuisance parameters with background-only hypothesis. It can be seen that the uncertainties have been reduced compared to Figure 9 and that the procedure is able to produce a good-quality fit to the data.

The expected and observed limits on the different signal models are presented in Figure 11. Here the expected limits are obtained by taking the nominal background estimates as the expected data. For the $Z^{'}_{TC2}$ benchmark, limits on the production cross-sections vary from 4.2 pb to 0.03 pb for masses from 0.4 TeV to 3 TeV. A $Z^{'}_{TC2}$ of width 1.2% is excluded for masses less than 1.8 TeV, while masses below 2.0 TeV are expected to be excluded. The $Z^{''}_{TC2}$ mass limits are stronger for a width of 2% (3%), reaching 2.0 TeV (2.3 TeV). For the $g_{KK}$ benchmark, limits on the production cross-sections vary from 4.8 pb for a mass of 0.4 TeV, to 0.09 pb for a mass of 3 TeV. The 15.3% width $g_{KK}$ is excluded for masses less than 2.2 TeV, while masses below 2.3 TeV are expected to be excluded. The cross-section limits on the $G_{KK}$ model range from 2.5 pb for a mass of 0.4 TeV to 0.03 pb for 2.5 TeV, with no mass range excluded in the benchmark scenario. The cross-section limits on the narrow scalar range from 3.0 pb for a mass of 0.4 TeV to 0.03 pb for 2.5 TeV. The cross-section limits are generally stronger for the latter two benchmark models than those for the spin-1 resonances due to higher acceptance.

The width dependence of the cross-section limits was also evaluated for the $g_{KK}$ models. The results are presented in Figure 12. For a 1 TeV resonance, the limits weaken by approximately a factor of two as the width increases from 10% to 40%. The effect is stronger for 2 and 3 TeV resonances, where the limits weaken by a factor of three over this width range.

The observed upper limits on the cross-section times $t\bar{t}$ branching ratio are larger than the expected limits, especially for $t\bar{t}$ resonance masses greater than 1.8 TeV. This arises from the use of the profile likelihood method which allows the data to constrain the systematic uncertainties using the full $m_{t\bar{t}}^{\text{reco}}$ distribution, thanks to the abundant data. The maximisation of the likelihood can give both new central values and new
Figure 10: The $m_{t\bar{t}}^{\text{reco}}$ distributions, after the nuisance-parameter fit with background-only hypothesis, summed over (a) all 6 boosted channels, (b) all 6 resolved channels, and (c) all 12 channels compared with data. The SM background components are shown as stacked histograms. The shaded areas indicate the total systematic uncertainties. The red line shows the expected distribution for a hypothetical $g_{KK}$ of mass 2.0 TeV.
Obs. 95% CL upper limit
Exp. 95% CL upper limit
Exp. 1 σ uncertainty
Exp. 2 σ uncertainty

(a) $Z'$, resolved and boosted combination.

(b) $g_{KK}$, resolved and boosted combination.

(c) $G_{KK}$, resolved and boosted combination.

(d) Scalar resonance, resolved and boosted combination.

Figure 11: Observed and expected upper limits on the cross-section times the $t\bar{t}$ branching ratio as a function of the mass of (a) Topcolour-assisted-technicolour $Z'_{TC2}$, (b) Bulk RS Kaluza–Klein gluon, (c) Bulk RS Kaluza–Klein graviton, (d) scalar resonance.

widths for the nuisance parameters. In the region of $m_{t\bar{t}}^{reco}$ above 1.5 TeV, the background prediction as seen in Figure 9 is slightly higher than the data, which would naively lead to the anticipation that the observed limits should be slightly better than the expected ones. However, the central values of some nuisance parameters are significantly shifted in the fit, leading to counterintuitive results. In that region, one of the dominant uncertainties is the high-$p_T$ jet b-tagging extrapolation uncertainty, as detailed in Section 8. This uncertainty is reduced to half of its original size by the fit, and the central value is also reduced by approximately one pre-fit standard deviation. As this uncertainty is associated with reconstructed physics objects, it has correlated effects on the predictions of BSM signal and SM background. The $1\sigma$ change in the central value of this nuisance parameter reduces the acceptance of high mass signals in the boosted selection considerably, approximately 25% for $Z'$ with a mass of 2 TeV. As a consequence, the observed upper limits on the cross-section obtained from the fit to the data are larger than the expected ones fitted to the nominal background estimates. Constraints from the fit are also observed in the nuisance parameters associated with other major systematic uncertainties such as the PDF, the shape of $W +$ jets background, the energy and mass scales for both large-radius and small-radius jets, b-tagging efficiencies, and $t\bar{t}$ MC modelling. These constraints have been understood through their impact on the fitted $m_{t\bar{t}}^{reco}$ distributions.
Figure 12: Observed and expected upper limits on the cross-section times the $t\bar{t}$ branching ratio as a function of the width of Kaluza–Klein gluon. The theoretical predictions for the production cross-section times branching ratio at the corresponding widths are also shown.

A few of the nuisance parameters also have their central values changed slightly by the fit, but their impact on the signal acceptance and hence cross-section limits are much smaller than the one detailed above.

Studies have been done to understand the origin of the shifted nuisance parameter associated with the high-$p_T$ jet b-tagging extrapolation uncertainty, in particular to investigate its potential origin as a mis-modelling of the $t\bar{t}$ background. Indeed an equally good fit can be obtained by introducing an additional ad hoc uncertainty in the modelling of the SM $t\bar{t}$ background with a similar dependency on $m_{t\bar{t}}^{reco}$. Such a mis-modelling would not reduce the signal acceptance and results in more stringent limits on the cross-section times $t\bar{t}$ branching ratio. In the end, no method was found to unambiguously ascertain the origin of this mis-modelling. The limits presented here are those with 1σ change in the fitted value of the nuisance parameter associated with the high-$p_T$ jet b-tagging extrapolation uncertainty, which happens to be more conservative results.

11 Summary

A search for heavy particles decaying to $t\bar{t}$ in the lepton plus jets decay channel was carried out with the ATLAS experiment at the LHC. The search uses data corresponding to an integrated luminosity of 20.3 fb$^{-1}$ of proton-proton collisions at a centre-of-mass energy of 8 TeV. No excess of events beyond the
Standard Model predictions is observed in the $t\bar{t}$ invariant mass spectra. Upper limits on the cross-section times branching ratio are set for four different signal models: a narrow ($\leq 3\%$ width) $Z'$ boson from Ref. [4], a broad (15.3\% width) Randall–Sundrum Kaluza–Klein gluon from Ref. [33], a Bulk Randall–Sundrum Kaluza–Klein graviton from Ref. [34], and a narrow scalar resonance. Based on these results, the existence of the narrow leptophobic topcolour $Z'$ in the range $0.4 \text{ TeV} < m_{Z'} < 1.8 \text{ TeV}$ is excluded at 95\% CL. A broad Kaluza–Klein gluon with mass between 0.4 TeV and 2.2 TeV is also excluded at 95\% CL.
References


31


