Search for pair production of heavy top-like quarks decaying to a high-pT W boson and a b quark in the lepton plus jets final state at $\sqrt{s} = 7$ TeV with the ATLAS detector


Published in: Physics Letters B

DOI: 10.1016/j.physletb.2012.11.071

Link to publication

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).
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

ATLAS Collaboration

A R T I C L E I N F O

Article history:
Received 19 October 2012
Received in revised form 30 November 2012
Accepted 30 November 2012
Available online 5 December 2012
Editor: H. Weerts

A B S T R A C T

A search is presented for production of a heavy up-type quark ($t'$) together with its antiparticle, assuming a significant branching ratio for subsequent decay into a $W$ boson and a $b$ quark. The search is based on 4.7 fb$^{-1}$ of pp collisions at $\sqrt{s} = 7$ TeV recorded in 2011 with the ATLAS detector at the CERN Large Hadron Collider. Data are analyzed in the lepton + jets final state, characterized by a high-transverse-momentum isolated electron or muon, large missing transverse momentum and at least three jets. The analysis strategy relies on the substantial boost of the $W$ bosons in the $t't'$ signal when $m_{t'} \gtrsim 400$ GeV. No significant excess of events above the Standard Model expectation is observed and the result of the search is interpreted in the context of fourth-generation and vector-like quark models. Under the assumption of a branching ratio $BR(t' \rightarrow Wb) = 1$, a fourth-generation $t'$ quark with mass lower than 656 GeV is excluded at 95% confidence level. In addition, in light of the recent discovery of a new boson of mass $\sim 126$ GeV at the LHC, upper limits are derived in the two-dimensional plane of $BR(t' \rightarrow Wb)$ versus $BR(t' \rightarrow Ht)$, where $H$ is the Standard Model Higgs boson, for vector-like quarks of various masses.

© 2012 CERN. Published by Elsevier B.V. All rights reserved.

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 often present in new physics models aimed at solving some of the limitations of the SM.

One possibility is the addition of a fourth generation of heavy chiral fermions [3,4], which can provide new sources of CP violation that could explain the matter–antimatter asymmetry in the universe. The new weak-isospin doublet contains heavy up-type ($t'$) and down-type ($b'$) quarks that mix with the lighter quarks via an extended CKM matrix. In order to be consistent with precision electroweak data, a relatively small mass splitting between the new quarks is required [5]. Assuming that $m_t - m_{t'} < m_W$, where $m_W$ is the $W$ boson mass, the $t'$ quark decays predominantly to a $W$ boson and a down-type quark ($q = d, s, b$). Based on the mixing pattern of the known quarks, it is natural to expect that this quark would be dominantly a $b$ quark, which has motivated the assumption of $BR(t' \rightarrow Wb) = 1$ in most experimental searches.

Another possibility is the addition of weak-isospin singlets, doublets or triplets of vector-like quarks [6], 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 or extra-dimensional models. In these models, a top-partner quark, for simplicity referred to here as $t'$, often plays a key role in canceling the quadratic divergences in the Higgs boson mass induced by radiative corrections involving the top quark. Vector-like quarks can mix preferentially with third-generation quarks, as the mixing is proportional to the mass of the SM quark [7], and they present a richer phenomenology than chiral quarks in fourth-generation models. In particular, a vector-like $t'$ quark has $a \text{ priori}$ three possible decay modes, $t' \rightarrow Wb$, $t' \rightarrow Zt$, and $t' \rightarrow Ht$, 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. In the case of a triplet, the $t'$ quark can decay either as a singlet or a doublet depending on its hypercharge.

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, these new heavy quarks would be produced predominantly in pairs via the strong interaction for masses below $O(1$ TeV) [6], 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 $t'$ quark and the weak gauge bosons.

Recent results of SM Higgs boson searches at the LHC have significantly impacted the prospects and focus of heavy-quark searches. In particular, the observation of a new boson by the ATLAS [8] and CMS [9] Collaborations with a mass of $\sim 126$ GeV and couplings close to those expected for the SM Higgs boson disfavors [5,10] fourth-generation models. These models predict a large increase in the production rate for $gg \rightarrow Wb$, which is in tension with searches in the $H \rightarrow WW^{(*)}$ and $H \rightarrow ZZ^{(*)}$ decay channels [11,12]. These results severely constrain perturbative fourth-generation models, although they may not completely exclude them yet. For example, it has been pointed out that a fourth family of fermions can substantially modify the Higgs boson partial decay widths [13] and various scenarios may still remain viable [5,14]. At the same time, the observation of this new boson raises the level of interest for vector-like quark searches, as $t' \rightarrow Ht$ and $b' \rightarrow Hb$ decays now have completely specified final states which offer an exciting opportunity for discovery of new heavy quarks.

In this Letter a search is presented for $tt'$ production using $pp$ collision data at $\sqrt{s} = 7$ TeV collected with the ATLAS detector. The search is optimized for $t'$ quark decays with large branching ratio to $Wb$. The lepton + jets final state signature, where one of the $W$ bosons decays leptonically and the other hadronically, is considered. The most recent search by the ATLAS Collaboration in this final state [15] was based on 1.04 fb$^{-1}$ of $pp$ collisions at $\sqrt{s} = 7$ TeV and, under the assumption of $BR(t' \rightarrow Wb) = 1$, excluded the existence of a $t'$ quark with a mass below 404 GeV at 95% confidence level (CL). A more stringent lower 95% CL limit of $m_{t'} > 570$ GeV [16] was obtained by the CMS Collaboration using 5.0 fb$^{-1}$ of data at $\sqrt{s} = 7$ TeV. Searches have also been performed exploiting the dilepton signature resulting from the leptonic decay of both $W$ bosons. A search by the ATLAS Collaboration in the dilepton final state using 1.04 fb$^{-1}$ of data at $\sqrt{s} = 7$ TeV obtained a lower 95% CL limit of $m_{t'} > 350$ GeV [17]. This search did not attempt to identify the flavor of the jets, making a more relaxed assumption of $BR(t' \rightarrow Wq) = 1$, where $q$ could be any down-type SM quark. A 95% CL limit of $m_{t'} > 557$ GeV [18], assuming $BR(t' \rightarrow Wb) = 1$, was obtained by the CMS Collaboration using 5.0 fb$^{-1}$ of data at $\sqrt{s} = 7$ TeV.

In comparison with the previous result by the ATLAS Collaboration in the lepton + jets final state [15], the search presented in this Letter uses almost a factor of five more data and has revisited the overall strategy, as advocated in Refs. [19–21], to take advantage of the kinematic differences that exist between top quark and $t'$ quark decays when $m_{t'} \gtrsim 400$ GeV. In particular, the hadronically-decaying $W$ boson can be reconstructed as a single isolated jet when it is sufficiently boosted, leading to a significantly improved sensitivity in comparison to previous searches. In addition, the result of this search is interpreted more generically in the context of vector-like quark models where $BR(t' \rightarrow Wb)$ can be substantially smaller than unity. In this case the additional signals, other than $t't' \rightarrow WbWb$, contribute to the signal acceptance and are accounted for in the analysis.

2. ATLAS detector

The ATLAS detector [22] consists of an inner tracking system surrounded by a superconducting solenoid, electromagnetic and hadronic calorimeters, and a muon spectrometer. The inner tracking system is immersed in a 2 T axial magnetic field and consists of a silicon pixel detector, a silicon microstrip detector, and a transition radiation tracker, providing charged particle identification in the region $|\eta| < 2.5$. The electromagnetic (EM) sampling calorimeter uses lead and liquid argon. The hadron calorimeter is based on two different detector technologies with either scintillator tiles or liquid argon as the active medium. The barrel hadronic calorimeter consists of scintillating tiles with steel plates as the absorber material. The endcap and forward hadronic calorimeters both use liquid argon, and copper or tungsten as the absorber, respectively. The calorimeters provide coverage up to $|\eta| = 4.9$. The muon spectrometer consists of superconducting air-core toroids, a system of trigger chambers covering the range $|\eta| < 2.4$, and high-precision tracking chambers allowing muon momentum measurements in the range $|\eta| < 2.7$.

3. Data sample and event preselection

The data used in this analysis correspond to the full dataset recorded in 2011, and were acquired using single-electron and single-muon triggers. The corresponding integrated luminosity is 4.7 fb$^{-1}$.

The event preselection criteria closely follow those used in recent ATLAS top quark studies [23] and require exactly one isolated electron or muon with large transverse momentum $(p_T > 25$ GeV and $|\eta| < 2.47$, excluding the transition region $(1.37 < |\eta| < 1.52)$ between the barrel and endcap EM calorimeters. Muon candidates are required to satisfy $p_T > 20$ GeV and $|\eta| < 2.5$. For leptons satisfying these $p_T$ requirements the efficiencies of the relevant single-lepton triggers have reached their plateau values. To reduce background from non-prompt leptons produced in semileptonic $b$- or $c$-hadron decays, or in $\pi^{\pm}/K^{\pm}$ decays, the selected leptons are required to be isolated, i.e. to have little calorimetric energy or track transverse momentum around them [24]. In this analysis $\tau$ leptons are not explicitly reconstructed. Because of the high-$p_T$ threshold requirements, only a small fraction of $\tau$ leptons decaying leptonically are reconstructed as electrons or muons, while the majority of $\tau$ leptons decaying hadronically are reconstructed as jets.

Jets are reconstructed with the anti-$k_t$ algorithm [25] with radius parameter $R = 0.4$, from topological clusters [26] of energy deposits in the calorimeters, calibrated at the EM scale. These jets are then calibrated to the particle (truth) level [27] using $p_T$- and $\eta$-dependent correction factors derived from a combination of data and simulation. Jets are required to have $p_T > 25$ GeV and $|\eta| < 2.5$. To avoid selecting jets from other $pp$ interactions in the same bunch crossing, at least 75% of the sum of the $p_T$ of tracks associated with a jet is required to come from tracks compatible with originating from the identified hard-scatter primary vertex. This primary vertex is chosen among the reconstructed candidates as the one with the highest $\sum p_T^2$ of associated tracks and is required to have at least three tracks with $p_T > 0.4$ GeV.

To identify jets as originating from the hadronization of a $b$ quark ($b$ tagging), a continuous discriminant is produced by an algorithm [28] using multivariate techniques to combine information from the impact parameter of displaced tracks, as well as topological properties of secondary and tertiary decay vertices reconstructed within the jet. In the preselection, at least one jet is

\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 $(x, y)$ plane, $\phi$ being the azimuthal angle around the beam pipe. The pseudorapidity is defined in terms of the polar angle $\theta$ as $\eta = -\ln \tan(\theta/2)$.}
required to have a discriminant value larger than the point corresponding to an average efficiency in simulated \(t\bar{t}\) events of \(\sim 70\%\) for b-quark jets, of \(\sim 20\%\) for c-quark jets and of \(\sim 0.7\%\) for jets originating from light quarks (u, d, s) or gluons.

The \(E_T^{\text{miss}}\) is constructed \([29]\) from the vector sum of all calorimeter energy deposits\(^{2}\) contained in topological clusters, calibrated at the energy scale of the associated high-p_T object (e.g. jet or electron), and including contributions from selected muons. Background from multi-jet production is suppressed by the requirement \(E_T^{\text{miss}} > 35 (20)\) GeV in the electron (muon) channel, and \(E_T^{\text{miss}} + m_t > 60\) GeV, where \(m_t\) is the transverse mass\(^{3}\) of the lepton and \(E_T^{\text{miss}}\).

### 4. Background and signal modeling

After event preselection the main background is \(t\bar{t}\) production, with lesser contributions from the production of a W boson in association with jets (W + jets) and multi-jet events. Small contributions arise from single top-quark, Z + jets and diboson production. Multi-jet events contribute to the selected sample mostly via the misidentification of a jet or a photon as an electron, or via the presence of a non-prompt lepton, e.g. from a semileptonic b- or c-hadron decay. The corresponding yield is estimated via a data-driven method \([30]\), which compares the number of events obtained with either standard or relaxed criteria for the selection of leptons. For the W + jets background, the shape of the distributions of kinematic variables is estimated from simulation but the normalization is estimated from data using the predicted asymmetry between \(W^+ + \text{jets}\) and \(W^- + \text{jets}\) production in pp collisions \([31]\). All other backgrounds, including the dominant \(t\bar{t}\) background, and the signal, are estimated from simulation and normalized to their theoretical cross sections.

Simulated samples of \(t\bar{t}\) and single top-quark backgrounds (in the s-channel and for the associated production with a W boson) are generated with Mc@NLO v4.01 \([32-34]\) using the CT10 set of parton distribution functions (PDFs) \([35]\). In the case of t-channel single top-quark production, the AcesMC v3.8 leading-order (LO) generator \([36]\) with the MRST LO** PDF set \([37]\) is used. These samples are generated assuming a top quark mass of 172.5 GeV and are normalized to approximate next-to-next-to-LO (NNLO) theoretical cross sections \([36-38]\) using the MSTW2008 NNLO PDF set \([41]\). Samples of W/Z + jets events are generated with up to five additional partons using the ALPGEN v2.13 \([42]\) LO generator and the CTEQ6L1 PDF set \([43]\). The parton-shower and fragmentation steps are performed with HERWIG v6.520 \([44]\) in the case of MC@NLO and ALPGEN, and with PYTHIA 6.421 \([45]\) in the case of AcesMC. To avoid double-counting of partonic configurations in W/Z + jets events generated both by the matrix-element calculation and the parton shower, a matching scheme \([46]\) is employed. The W + jets samples are generated separately for W + light jets, Wbb + jets, Wc\(\bar{c}\) + jets, and Wc + jets, and their relative contributions are normalized using the fraction of b-tagged jets in W + 1-jet and W + 2-jets data control samples \([47]\). The Z + jets background is normalized to the inclusive NNLO theoretical cross section \([48]\). The diboson backgrounds are modeled using HERWIG with the MRST LO** PDF set, and are normalized to their NLO theoretical cross sections \([49]\). In all cases where HERWIG is used, the underlying event is simulated with JIMMY v4.31 \([50]\).

For four-generation t’ quark signals, samples are generated with PYTHIA using the CTEQ6.6 PDF set \([43]\) for a range of masses, \(m_{t’}\), from 400 GeV to 750 GeV in steps of 50 GeV. For vector-like t’ signals, samples corresponding to a singlet t’ quark decaying to Wb, Zt and Ht are generated with the Procos v2.2 LO generator \([6,51]\) using the CTEQ6L1 PDF set, and interfaced to PYTHIA for the parton shower and fragmentation. The \(m_{t’}\) values considered range from 400 GeV to 600 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 \([52]\). For both types of signal, the samples are normalized to the approximate NNLO theoretical cross sections \([38]\) using the MSTW2008 NNLO PDF set.

All simulated samples include multiple pp interactions and simulated events are weighted such that the distribution of the average number of interactions per bunch crossing agrees with data. The simulated samples are processed through a simulation \([53]\) of the detector geometry and response using GEANT \([54]\) and the same reconstruction software as the data. Simulated events are corrected so that the physics object identification efficiencies, energy scales and energy resolutions match those determined in data control samples, enriched in the physics objects of interest.

### 5. Final selection

After preselection, further background suppression is achieved by applying requirements aimed at exploiting the distinct kinematic features of the signal. The large t’ quark mass results in energetic W bosons and b quarks in the final state with large angular separation between them, while the decay products from the boosted W bosons have small angular separation. The combination of these properties is very effective in suppressing the dominant \(t\bar{t}\) background since \(t\bar{t}\) events with boosted W boson configurations are rare, and are typically characterized by a small angular separation between the W boson and b quark from the top quark decay.

To take advantage of these properties, it is necessary to identify the hadronically-decaying W boson (\(W_{\text{had}}\)) as well as the b jets in the event. The candidate b jets are defined as the two jets with the highest b-tag discriminant (although only one of them is explicitly required to be b tagged in the event selection). Two types of \(W_{\text{had}}\) candidates are defined, \(W_{\text{typeI had}}\) and \(W_{\text{typeII had}}\), depending on the angular separation between their decay products. \(W_{\text{typeI had}}\) is defined as a single jet with \(p_T > 250\) GeV and mass in the range of 60–110 GeV. The mass distribution for \(W_{\text{typeI had}}\) candidates, prior to the jet mass requirement itself, is shown in Fig. 1(a). \(W_{\text{typeII had}}\) is defined as a dijet system with \(p_T > 150\) GeV, angular separation\(^{4}\) \(\Delta R(j, j) < 0.8\) and mass within the range of 60–110 GeV. If multiple pairs satisfy the above requirements, the one with mass closest to the nominal W boson mass is chosen. The mass distribution for \(W_{\text{typeII had}}\) candidates, prior to the dijet mass requirement, is shown in Fig. 1(b). In the construction of both types of \(W_{\text{had}}\) candidates, all selected jets except for the two candidate b jets are considered. Small discrepancies observed between the data and the background prediction, e.g. at low \(W_{\text{typeII had}}\) candidate invariant mass, are not significant and are covered by the systematic uncertainties.

The leptonically-decaying W boson is reconstructed using the lepton and \(E_T^{\text{miss}}\), identified as the neutrino \(p_T\). Requiring that the

---

2 Each calorimeter cluster/cell is considered a massless object and is assigned the four-momentum \((E_{\text{cell}}, p_{\text{cell}})\), where \(E_{\text{cell}}\) is the measured energy and \(p_{\text{cell}}\) is a vector of magnitude \(E_{\text{cell}}\) directed from \((x, y, z) = (0, 0, 0)\) to the center of the cell.

3 The transverse mass is defined by the formula \(m_T = \sqrt{2p_T E_T^{\text{miss}} (1 - \cos \Delta \phi)}\), where \(p_T\) is the \(p_T\) of the lepton and \(\Delta \phi\) is the azimuthal angle separation between the lepton and \(E_T^{\text{miss}}\) directions.

4 The angular separation is defined as \(\Delta R = \sqrt{(\Delta \phi)^2 + (\Delta \eta)^2}\) where \(\phi\) is the azimuthal angle and \(\eta\) the pseudorapidity.
invariant mass of the lepton–neutrino system equals the nominal W boson mass allows reconstruction of the neutrino longitudinal momentum up to a two-fold ambiguity. In case no real solution exists, the neutrino pseudorapidity is set equal to that of the lepton, since in the kinematic regime of interest for this analysis the decay products of the W boson tend to be collinear.

Two final selections, loose and tight, are defined. The loose selection considers events with either $\geq 3$ jets, at least one of which is a $W_{\text{had}}$ candidate, or $\geq 4$ jets, two of which combine to make at least one $W_{\text{had}}$ candidate, and no $W_{\text{had}}$ candidate. The events must satisfy $H_T > 750$ GeV, where $H_T$ is the scalar sum of the lepton $p_T$, $E^{\text{miss}}$, and the $p_T$ of the four (or three if there are only three) highest-$p_T$ jets. The $H_T$ distribution peaks at $\sim 2m_t$ for signal events, which makes the $H_T > 750$ GeV requirement particularly efficient for signal with $m_t \geq 400$ GeV, while rejecting a large fraction of the background. In addition, the highest-$p_T$ b-jet candidate ($b_1$) and the next-to-highest-$p_T$ b-jet candidate ($b_2$) are required to have $p_T > 160$ GeV and $p_T > 60$ GeV, respectively. Finally, the angular separation between the lepton and the reconstructed neutrino is required to satisfy $\Delta R(\ell, \nu) < 1.4$. The tight selection adds the following isolation requirements to the loose selection: $\min(\Delta R(W_{\text{had}}, b_1, b_2)) > 1.4$ and $\min(\Delta R(\ell, b_1, b_2)) > 1.4$, which are particularly effective at suppressing $t\bar{t}$ background. Table 1 presents a summary of the background estimates for the loose and tight selections, as well as a comparison of the total predicted and observed yields. The quoted uncertainties include both statistical and systematic contributions. The latter are discussed in Section 7. The predicted and observed yields are in agreement within these uncertainties.

6. Heavy-quark mass reconstruction

The main discriminant variable used in this search is the reconstructed heavy-quark mass ($m_{\text{reco}}$), built from the $W_{\text{had}}$ candidate and one of the two $b$-jet candidates. The reconstruction of the leptonically-decaying W boson usually yields two solutions, and there are two possible ways to pair the $b$-jet candidates with the W boson candidates to form the heavy quarks. Among the four possible combinations, the one yielding the smallest absolute difference between the two reconstructed heavy quark masses is chosen. The resulting $m_{\text{reco}}$ distributions in Fig. 2 show that the SM background has been effectively suppressed, and that, as is most visible for the loose selection, good discrimination between signal and background is achieved. The small contributions from $W +$ jets, $Z +$ jets, diboson, single-top and multi-jet events are combined into a single background source referred to as non-$t\bar{t}$. It was verified a priori that the tight selection has the better sensitivity, and it is therefore chosen to derive the final result for the search. The loose selection, displaying a significant $t\bar{t}$ background at low $m_{\text{reco}}$, which is in good agreement with the expectation, provides further confidence in the background modeling prior to the application of $b$-jet isolation requirements in the tight selection.

---

Table 1

<table>
<thead>
<tr>
<th></th>
<th>Loose selection</th>
<th>Tight selection</th>
</tr>
</thead>
<tbody>
<tr>
<td>$W +$ jets</td>
<td>$5.4 \pm 4.2$</td>
<td>$2.0 \pm 1.4$</td>
</tr>
<tr>
<td>$Z +$ jets</td>
<td>$0.5 \pm 0.4$</td>
<td>$0.2 \pm 0.2$</td>
</tr>
<tr>
<td>Single top</td>
<td>$7.2 \pm 1.7$</td>
<td>$1.1 \pm 0.5$</td>
</tr>
<tr>
<td>Dibosons</td>
<td>$0.1 \pm 0.1$</td>
<td>$0.04 \pm 0.04$</td>
</tr>
<tr>
<td>Multi-jet</td>
<td>$5.9 \pm 8.4$</td>
<td>$3.8 \pm 3.2$</td>
</tr>
<tr>
<td>Total background</td>
<td>$113 \pm 30$</td>
<td>$11.3 \pm 4.8$</td>
</tr>
<tr>
<td>Data</td>
<td>122</td>
<td>11</td>
</tr>
</tbody>
</table>

$t\bar{t}$ (500 GeV)

$Wb : Zt : Ht = 1.0 : 0.0 : 0.0$  
$47.4 \pm 6.3$  
$28.2 \pm 3.6$

$Wb : Zt : Ht = 0.5 : 0.0 : 0.5$  
$25.4 \pm 3.6$  
$11.2 \pm 1.5$
due to the dominant uncertainties associated, including contributions from uncertainties on the theories (2%), jet identification efficiency (2%), and cross sections for integrated luminosity (3.9%), lepton identification and trigger efficiencies (2%), and the jet energy resolution are 2%, 3% and 3%, respectively. Uncertainties associated with the jet mass scale and resolution, affecting the jet misidentification rate (50%) in the data-driven prediction.

7. Systematic uncertainties

Systematic uncertainties affecting the normalization and shape of the $m_{\text{reco}}$ distribution are estimated taking into account correlations.

Uncertainties affecting only the normalization include the integrated luminosity (3.3%), lepton identification and trigger efficiencies (2%), jet identification efficiency (2%), and cross sections for various background processes. The uncertainties on the theoretical cross sections for $t\bar{t}$, single-top and diboson production are $(+9.9/-10.7)\%$ [38], $(+4.7/-3.7)\%$ [39,40], and $\pm 5\%$ [49], respectively. A total uncertainty on the $W^+\text{jets}$ normalization of 58% is assumed, including contributions from uncertainties on the $W+4$-jets cross section [48] [55], the heavy-flavor content measured in $W+1,2$-jets data samples (23%) [47], as well as its extrapolation to higher jet multiplicities (19%). The latter is estimated from the simulation where the $W+\text{heavy-flavor}$ fractions are studied as a function of variations in the ALPGEN generator parameters. Similarly, the $Z+\text{jets}$ normalization is assigned an uncertainty of 48% due to the dominant $Z+4$-jets contribution after final selection, which is evaluated at LO by ALPGEN. The multi-jet normalization is assigned an uncertainty of 80% including contributions from the limited size of the data sample (64%) as well as the uncertainty on the jet misidentification rate (50%) in the data-driven prediction.

The rest of the systematic uncertainties modify both the normalization and shape of the $m_{\text{reco}}$ distribution. To indicate their magnitudes, their impact on the normalization for the tight selection is discussed in the following. Among the largest uncertainties affecting the $t\bar{t}$ background are those related to modeling, such as (1) the choice of NLO event generator (evaluated by comparing MC@NLO and POWHEG [56]), (2) the modeling of initial- and final-state QCD radiation (evaluated by varying the relevant parameters in PYTHIA in a range given by current experimental data [57]), and (3) the choice of parton-shower and fragmentation models (based on the comparison of HERWIG and PYTHIA). These result in $t\bar{t}$ normalization uncertainties of 55%, 1% and 26%, respectively.

The uncertainty on the jet energy scale [27] affects the normalization of the $t\bar{t}$ signal, $t\bar{t}$ background and non-$t\bar{t}$ backgrounds by $\pm 6\%$, $(+22/-25)\%$, and $(+19/-10)\%$, respectively. The uncertainties due to the jet energy resolution are 2%, 3% and 3%, respectively. Uncertainties associated with the jet mass scale and resolution, affecting the selection of $W^{\text{type1}}_{\text{had}}$ candidates, are smaller in magnitude but are also taken into account. Uncertainties on the modeling of the $b$-tagging algorithms affect the identification of $b$, $c$ and light jets [28,58,59], and collectively result in uncertainties for the $t\bar{t}$ signal, as well as the $t\bar{t}$ and non-$t\bar{t}$ backgrounds, of $(5–6)\%$. Other systematic uncertainties such as those on jet reconstruction efficiency or the effect of multiple $pp$ interactions on the modeling of $E_{\text{miss}}$ have been verified to be negligible.

In summary, taking into account all systematic uncertainties discussed above, the total uncertainty on the normalization affecting the tight selection for a $t\bar{t}$ signal with $m_t = 500$ GeV, $t\bar{t}$ and non-$t\bar{t}$ backgrounds is $11\%$, $67\%$ and $50\%$, respectively.

8. Statistical analysis

In the absence of any significant data excess, the $m_{\text{reco}}$ spectrum shown in Fig. 2(b) is used to derive 95% CL upper limits on the $t\bar{t}$ production cross section using the CL_s method [60,61].

---

**Fig. 2.** Distribution of $m_{\text{reco}}$ for the combined $e+\text{jets}$ and $\mu+\text{jets}$ channels after the (a) loose and (b) tight selection. The data (solid black points) are compared to the SM prediction. The total uncertainty on the background estimation (see Section 7 for details) is shown as a black hashed band. Also shown, stacked on top of the SM background, are the expected contributions from a signal with mass $m_t = 500$ GeV for the case of $BR(t' \rightarrow Wb) = 1$ (red shaded histogram), corresponding to a fourth-generation $t'$ quark, as well as the case of $BR(t' \rightarrow Wb) = BR(t' \rightarrow Ht) = 0.5$ (dashed black histogram). The overflow has been added to the last bin. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this Letter.)

**Fig. 3.** Observed (solid line) and expected (dashed line) 95% CL upper limits on the $t\bar{t}$ cross section as a function of the $t'$ quark mass. The surrounding shaded bands correspond to the $\pm 1$ and $\pm 2$ standard deviations around the expected limit. The thin red line and band show the theoretical prediction and its $\pm 1$ standard deviation uncertainty. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this Letter.)
This method employs a log-likelihood ratio \( LLR = -2 \log(L_{s+b}/L_b) \) as test-statistic, where \( L_{s+b} (L_b) \) 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 fraction of pseudo-experiments for the signal-plus-background (background-only) hypothesis with \( LLR \) larger than a given threshold defines \( CL_{s+b} (CL_b) \). Such threshold is set to the observed (median) \( LLR \) for the observed (expected) limit. Signal cross sections for which \( CL_s = CL_{s+b}/CL_b < 0.05 \) are deemed to be excluded at 95% CL. Dividing by \( CL_s \) minimizes the possibility of mistakenly excluding a small signal due to a downward fluctuation of the background.

9. Results

The resulting observed and expected upper limits on the \( t\bar{t}^{'} \) production cross section are shown in Fig. 3 as a function of \( m_{t^{'}} \), and compared to the theoretical prediction, assuming \( BR(t^{'} \to Wb) = 1 \). The total uncertainty on the theoretical cross section includes the contributions from scale variations and PDF uncertainties. An observed (expected) 95% CL limit \( m_{t^{'}} > 656 (638) \) 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 fourth-generation \( t^{'} \) quark decaying exclusively into a \( W \) boson and a \( b \) quark. This limit is also applicable to a down-type vector-like quark with electric charge of \(-4/3\) and decaying into a \( W \) boson and a \( b \) quark.

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 by \( BR(t^{'} \to Zt) = 1 - BR(t^{'} \to Wb) - BR(t^{'} \to Ht) \). To probe this two-dimensional branching-ratio plane, the signal samples with the original branching ratios as generated by Pythia are weighted. The resulting 95% CL exclusion limits are shown in Fig. 4 for different values of \( m_{t^{'}} \). For instance, a \( t^{'} \) quark with a mass of 550 GeV and \( BR(t^{'} \to Wb) > 0.63 \) is excluded at \( \geq 95 \% \) CL, regardless of the value of its branching ratios to \( Ht \) and \( Zt \). All the decay modes contribute to the final sensitivity when setting limits. For example, assuming \( m_{t^{'}} = 550 \) GeV, the efficiency of the \( t^{'} \) selection with at least four jets is 2.67%, 0.64%, 0.81%, 0.27%, 0.24% and 0.25%, for decays to \( WbWb \), \( WbHt \), \( WbZt \), \( ZtHt \), \( ZtZt \) and \( HtHt \), respectively. The default predictions from Pythia for the weak-isospin singlet and doublet cases are shown. A weak-isospin singlet \( t^{'} \) quark with 400 \( \leq m_{t^{'}} \leq 500 \) GeV is excluded at \( \geq 95 \% \) CL. It should be noted that since this analysis is optimized for \( m_{t^{'}} > 400 \) GeV (recall the \( H_{t^{'}} > 750 \) GeV requirement), it is not sensitive for vector-like quark scenarios where \( m_{t^{'}} < 400 \) GeV. The doublet scenarios are shown in Fig. 4 to illustrate the fact that this analysis has no sensitivity in these cases.

10. Conclusion

The strategy followed in this search, directly exploiting the distinct boosted signature expected in the decay of a heavy \( t^{'} \) quark, has resulted in the most stringent limits to date on a fourth-generation \( t^{'} \) quark. This approach shows great promise for improved sensitivity in future LHC searches at higher centre-of-mass energy and integrated luminosity. This search is also interpreted more generically in the context of vector-like quark models, resulting in the first quasi-model-independent exclusions in the two-dimensional plane of \( BR(t^{'} \to Wb) \) versus \( BR(t^{'} \to Ht) \), for different values of the \( t^{'} \) quark mass.

Fig. 4. 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 Pythia event generator for the weak-isospin singlet and doublet cases are shown as plain circle and star symbols, respectively. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this Letter.)
Acknowledgements

We thank CERN for the very successful operation of the LHC, as well as the support staff from our institutions without whom ATLAS could not be operated efficiently.

We acknowledge the support of ANPCyT, Argentina; YerPhI, Armenia; ARC, Australia; BMWF and FWF, Austria; ANAS, Azerbaijan; SSTC, Belarus; CERN, Russia, and CFI, Canada; CERN; CONICYT, Chile; CAS, MOST and NSFC, China; COLCIENCIAS, Colombia; MSMT CR, MPO CR and VSC CR, Czech Republic; DNRF, DNSRC and Lundbeck Foundation, Denmark; EPLANET and ERC, European Union; IN2P3-CNR, CEA-DSM/IRFU, France; GNSF, Georgia; BMBF, DFG, MPG and AvH Foundation, Germany; GSRT, Greece; ISF, MINEVA, GIF, DIP and Benoziyo Center, Israel; INFN, Italy; MEXT and JSPS, Japan; CRTA, Korea; LNAE and CONICET, Argentina; King Abdullah University of Science and Technology, Saudi Arabia; NSC, Taiwan; TAEK, Turkey; STFC, the Royal Society and Leverhulme Trust, SER, SNSF and Cantons of Bern and Geneva, Switzerland; NSC, Taiwan; NFR, South Africa; MICINN, Spain; SRC and Wallenberg Foundation, Sweden; SER, NSFF and Cantons of Bern and Geneva, Switzerland; NSC, Taiwan; TAEK, Turkey; STFC, the Royal Society and Leverhulme Trust, United Kingdom; DOE and NSF, United States of America.

The crucial computing support from all WLCG partners is acknowledged gratefully, in particular from CERN and the ATLAS Tier-1 facilities at TRIUMF (Canada), NDGF (Denmark, Norway, Sweden), CC-IN2P3 (France), KIT/GridKA (Germany), INFN-CNAF (Italy), NL-T1 (Netherlands), PIC (Spain), ASGC (Taiwan), RAL (UK) and BNL (USA) and in the Tier-2 facilities worldwide.

Open access

This article is published Open Access at sciedirect.com. It is distributed under the terms of the Creative Commons Attribution License 3.0, which permits unrestricted use, distribution, and reproduction in any medium, provided the original authors and source are credited.

References
