Search for single production of a vectorlike T quark decaying into a Higgs boson and top quark with fully hadronic final states using the ATLAS detector

Aad, G.; ATLAS Collaboration

DOI
10.1103/PhysRevD.105.092012

Publication date
2022

Document Version
Final published version

Published in
Physical Review D

License
CC BY

Citation for published version (APA):
Search for single production of a vectorlike $T$ quark decaying into a Higgs boson and top quark with fully hadronic final states using the ATLAS detector

G. Aad et al.*

(ATLAS Collaboration)

(Received 19 January 2022; accepted 29 March 2022; published 25 May 2022)

A search is made for a vectorlike $T$ quark decaying into a Higgs boson and a top quark in 13 TeV proton-proton collisions using the ATLAS detector at the Large Hadron Collider with a data sample corresponding to an integrated luminosity of 139 fb$^{-1}$. The Higgs-boson and top-quark candidates are identified in the all-hadronic decay mode, where $H \to bb$ and $t \to bW \to bqq'$ are reconstructed as large-radius jets. The candidate Higgs boson, top quark, and associated $B$ hadrons are identified using tagging algorithms. No significant excess is observed above the background, so limits are set on the production cross section of a singlet $T$ quark at 95% confidence level, depending on the mass $m_T$ and coupling $\kappa_T$ of the vectorlike $T$ quark to Standard Model particles. In the considered mass range between 1.0 and 2.3 TeV, the upper limit on the allowed coupling values increases with $m_T$ from a minimum value of 0.35 for $1.07 < m_T < 1.4$ TeV to 1.6 for $m_T = 2.3$ TeV.

DOI: 10.1103/PhysRevD.105.092012

I. INTRODUCTION

The discovery of the Higgs boson [1,2] and measurements of the Higgs-boson couplings [3–6] by the ATLAS and CMS Collaborations confirm that the Standard Model of particle physics (SM) is an accurate description of nature at currently accessible energy scales. However, the SM still leaves many questions unanswered and is therefore not a complete theory. For example, radiative corrections to the Higgs-boson propagator from top-quark loops lead to a quadratic divergence in the mass of the Higgs boson [7]. The mechanism to cancel out the contribution from the top quark requires an unreasonable degree of fine-tuning to produce the observed 125 GeV Higgs boson. This so-called hierarchy problem is often considered to indicate that new physics naturally cancels out the divergent contributions to the Higgs-boson mass.

Vectorlike quarks are hypothetical spin-1/2 particles that arise in various models that address problems in the SM such as the hierarchy problem. Vectorlike quarks are color triplets whose left- and right-handed chiralities transform in the same way under weak isospin [8,9]. In little Higgs [10,11] and composite Higgs [12,13] models, the Higgs boson is naturally light because it is a pseudo Nambu-Goldstone boson arising from a spontaneously broken global symmetry [14]. Vectorlike quarks arise naturally in such models. Unlike the chiral current of SM quarks, vectorlike quarks have a pure vector current in the Lagrangian. In addition, vectorlike quarks do not acquire mass by interacting with the Higgs field, so they are not excluded by measurements of Higgs-boson properties.

In these models, vectorlike quarks are expected to couple preferentially to third-generation quarks [8,15] and can have both neutral- and charged-current decays. An up-type vectorlike $T$ quark with charge $+2/3$ can decay into $Wb$, $Zt$, or $Ht$, while a down-type $B$ quark with charge $-1/3$ can decay into $Wt$, $Zb$, or $Hb$ (and the charge conjugate states). To be consistent with results from precision electroweak measurements, the mass splitting between vectorlike quarks belonging to the same SU(2) multiplet should be small [16], preventing cascade decays such as $T \to WB$. Couplings between the vectorlike quarks and the first- and second-generation quarks are not excluded [17,18], but they are expected to be small.

Vectorlike quarks can be produced singly or in pairs in proton-proton ($pp$) collisions. There have been numerous searches for the pair production of vectorlike quarks [19–37] that have excluded $T$-quark masses below 1.37 TeV at 95% confidence level (C.L.) for a variety of decay modes. For $T$-quark masses above $\sim 1$ TeV, vectorlike quarks would mainly be produced singly if the couplings to SM particles were sufficiently large. Searches for single production of $T$ quarks have placed limits on $T$-quark production cross sections for $T$-quark masses between 1 and 2 TeV at 95% C.L. for various SM

*Full author list given at the end of the article.

Published by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. Further distribution of this work must maintain attribution to the author(s) and the published article’s title, journal citation, and DOI. Funded by SCOAP3.
The unknown parameters $\xi$, $\alpha_s$, and $\kappa_T$, that determine the $T$-quark branching ratios. In this analysis, the asymptotic limit of these parameters—such as $m_T$ goes to infinity—is assumed, leading to branching ratios of $1/2$, $1/4$, and $1/4$ for $T \rightarrow Wb$, $T \rightarrow Ht$, and $T \rightarrow Zt$, respectively. In this model, the unknown parameters $m_T$ and $\kappa_T$ define the expected $T$-quark cross section and resonance line shape. The search assumes this signal model in the interpretation of the data.

The $T$ quark is assumed to be a weak singlet state in this analysis; if additional multiplets of vectorlike quarks are assumed, the possible final states and branching ratios require an approach involving simultaneous consideration of several final states [16], which is beyond the scope of this paper.

The results reported here significantly extend the sensitivity to events in which a singly produced $T$ quark decays to $Ht$ followed by the hadronic decays $H \rightarrow bb$ and $t \rightarrow bW$. The use of fully hadronic decays allows the direct reconstruction of the $T$-quark final state, increasing the expected signal-to-background ratio in the signal region defined for the search. Compared with the most sensitive prior search [46], this search uses $\sim 4$ times more integrated luminosity, its sensitivity is improved by using tagging techniques resulting in a signal-to-background improvement of $\sim 3$, and it uses a data-driven multijet background estimate that reduces the uncertainty in the background estimate by an order of magnitude.

This fully hadronic final state is of particular interest for vectorlike quark masses above 1 TeV. The resulting high-$p_T$ jets from the top quark and Higgs boson are “boosted,” so that the decay products of the top quark and Higgs boson are collimated and captured in two large-radius (large-$R$) jets. This final state has the largest branching fraction of all the potential $Ht$ decay modes and the large-$R$ jets can be identified as either Higgs-boson or top-quark candidates through tagging algorithms that use the substructure within the jet [48,49]. In addition, bottom-quark jet identification ($b$-tagging) provides high background rejection with high efficiency given the three bottom-quark jets coming from the $H \rightarrow bb$ and $t \rightarrow Wb$ decays. Assuming the existence of single $T$-quark production, the signal would appear as an excess of events with $Ht$ invariant masses around the $T$-quark mass for values of $\kappa_T \lesssim 0.5$. Above this $\kappa_T$, the $Ht$ invariant mass distribution broadens to masses below the $T$-quark mass as $\kappa_T$ increases due to the convolution of increasing width and partonic densities.

The largest backgrounds come from boosted top-quark pair production and multijet events arising from the production of lighter high-$p_T$ quarks ($u$, $d$, $s$, $c$, and $b$) and/or gluons. The ATLAS [50–58] and CMS [59–63,63–69] Collaborations have published measurements of the $t\bar{t}$ differential cross sections at center-of-mass energies of $\sqrt{s} = 7$, 8, and 13 TeV in $pp$ collisions. The measured cross section for the production of top quarks with $p_T > 300$ GeV is $\sim 20\%$ lower than predicted by perturbative quantum chromodynamics (QCD) calculations performed at next-to-leading order (NLO) in the strong coupling constant $\alpha_s$. A control sample of fully reconstructed high-$p_T$ top-quark pairs is used with Monte Carlo (MC) models to normalize the expected background from top-quark pairs in the $Ht$ candidate sample. The multijet background is estimated using data-driven techniques developed for studies of events containing boosted top quarks [57].

This paper is organized as follows. Section II describes the ATLAS detector and Sec. III describes the datasets and MC samples that are used in this analysis. Section IV describes the object definition and event selection, while Sec. V summarizes the estimation of SM backgrounds to the $T$-quark signal. The systematic uncertainties are presented in Sec. VI and Sec. VII presents the results of the search. Conclusions are drawn in Sec. VIII.
II. ATLAS DETECTOR

The ATLAS detector [70] at the LHC is centered on the $pp$ collision point and covers nearly the whole 4π solid angle. It consists of an inner tracking detector surrounded by a 2 T superconducting solenoid, electromagnetic and hadronic calorimeters, and a muon spectrometer incorporating three large superconducting toroid magnets.

The inner detector, including the insertable B-layer added as a new innermost layer in 2014 [71,72], provides charged-particle tracking information from a pixel detector and silicon microstrip detector in the pseudorapidity range $|\eta| < 2.5$ and a transition radiation tracker covering $|\eta| < 2.0$.

The calorimeter system covers the pseudorapidity range $|\eta| < 4.9$ and measures the positions and energies of electrons, photons, and charged and neutral hadrons. Within the region $|\eta| < 3.2$, electromagnetic calorimetry is provided by barrel and endcap high-granularity lead and liquid-argon sampling calorimeters. The hadronic sampling calorimeter uses either scintillator tiles or liquid argon as active material and steel, copper, or tungsten as absorber.

The muon spectrometer comprises separate trigger and high-precision tracking chambers measuring the tracks of muons in a magnetic field generated by superconducting air-core toroid magnets. The precision chamber system covers the region $|\eta| < 2.7$, while the muon trigger system covers the range $|\eta| < 2.4$.

A two-level trigger system is used to select which events to save for offline analysis [73]. The first level is implemented in hardware/firmware and uses a subset of the detector information to reduce the event rate from the 40 MHz proton bunch crossings to less than 100 kHz. This is followed by a software-based high-level trigger that reduces the event rate to approximately 1 kHz. An extensive software suite [74] is used in the reconstruction and analysis of real and simulated data, in detector operations, and in the trigger and data acquisition systems of the experiment.

III. DATA AND SIMULATED SAMPLES

This analysis studies $pp$ collisions with a center-of-mass energy of $\sqrt{s} = 13$ TeV recorded by the ATLAS detector between 2015 and 2018. Only data-taking periods in which all the subdetectors were operational are considered. The dataset corresponds to an integrated luminosity of 139 fb$^{-1}$ [75], measured using the LUCID-2 detector [76]. The events used in this analysis were collected by a set of triggers requiring at least one anti-$k_t$ jet [77,78] with a jet radius parameter of $R = 1.0$ [73]. The maximum $p_T$ threshold value of these triggers was 480 GeV, which was found to be fully efficient when requiring the offline reconstruction of at least one large-$R$ jet with $p_T > 500$ GeV and $|\eta| < 2.0$, as described in Sec. IV.

The main backgrounds for this search are from $\bar{t}t$ and multijet events. There are also small contributions from single-top-quark and $t\bar{t} + X (X = W, Z, H)$ events. The multijet background is estimated using a data-driven method described in Sec. V, while the other backgrounds, as well as the $T$-quark signal events, are estimated with MC simulations as described below. The multijet background estimate also includes backgrounds arising from electro-weak and QCD processes such as $W/Z +$ jets production.

The $T$-quark signal samples were produced at leading order, using the MadGraph5_aMC@NLO2.6.0 MC generator [79] to generate the hard interaction and the PYTHIA8 generator for parton shuffling and hadronization. The parton distribution function (PDF) set used is NNPDF3.0NLO [80]. Both $W$- and $Z$-mediated production contribute to single $T$-quark production and were included in the MC event generation, with the $Z$-mediated process having a cross section times smaller than the $W$-mediated process and comprising less than 1% of the total yield after the event selection described in Sec. IV. The matrix elements were calculated using the phenomenological model given in Ref. [47]. These include all tree-level processes, ensuring the inclusion of both resonant and nonresonant single $T$-quark production modes. The decay channel considered is $T \rightarrow Ht$, with $m_T$ and $\kappa_T$ as unknown parameters. The three additional parameters, $\xi_W$, $\xi_Z$, and $\xi_H$, that determine the $T$-quark branching ratios are set to the asymptotic limit in $m_T$, leading to branching ratios of 1/2, 1/4, and 1/4 for $T \rightarrow Wb$, $T \rightarrow Ht$, and $T \rightarrow Zt$, respectively. In order to accurately model the change in cross section and line shape as $m_T$ and $\kappa_T$ are varied, MC samples were created for a variety of mass and coupling values, with $m_T$ ranging from 1.0 to 2.3 TeV in steps of 0.1 TeV and $\kappa_T$ ranging from 0.1 to 1.6 in steps of 0.05 for $\kappa_T < 0.5$ and 0.1 for larger $\kappa_T$. All signal samples are normalized to cross sections that have been calculated at NLO in QCD [81]. These cross sections are computed in a $T$-quark narrow-width approximation and a correction factor is applied [82] to account for finite-width effects.

For all MC samples, the masses of the top quark ($m_{\text{top}}$) and Higgs boson were set to 172.5 and 125.0 GeV, respectively. The production of $t\bar{t}$ events was modeled using the Powheg Box v2 [83–86] generator. This provides matrix elements at NLO with the NNPDF3.0 NLO PDF. In addition, the $\eta_{\text{damp}}$ parameter, which controls the matching of the matrix element to the parton shower in Powheg and effectively regulates the high-$p_T$ radiation against which
the $t\bar{t}$ system recoils, was set to $1.5 \times m_{\text{top}}$ [87]. The functional form of the renormalization and factorization scales was set to the default scale $\sqrt{m_{\text{top}}^2 + p_T^2}$. The PYTHIA8.230 [88] parton-shower and hadronization models were employed, using a set of tuned parameter values called the A14 tune [89], and the NNPDF2.3LO set of PDFs [90]. The decays of bottom and charm hadrons were simulated using the EvtGen 1.6.0 program [91].

The $t\bar{t}$ sample is normalized to the cross section prediction at next-to-next-to-leading order (NNLO) in QCD including the resummation of next-to-next-to-leading logarithmic (NNLL) soft-gluon terms calculated using TOP++2.0 [92–98]. For $pp$ collisions at a center-of-mass energy of $\sqrt{s} = 13$ TeV, this cross section is $\sigma(t\bar{t})_{\text{NNLO+NNLL}} = 832 \pm 51$ fb. The cross section uncertainties due to the PDF and $\alpha_s$ are calculated using the PDF4LHC prescription [99] with the MSTW2008NNLO 68% C.L. [100,101], CT10NNLO [102,103], and NNPDF2.3LO 5f FFN [90] PDF sets, and are added in quadrature to the effect of the scale uncertainty.

The uncertainty due to initial-state radiation (ISR) was estimated by varying the Var3c A14 tune, renormalization scale $\mu_R$, factorization scale $\mu_F$, and the $h_{\text{damp}}$ parameter independently. The Var3c A14 tune variation corresponds to the variation of $\alpha_s$ for ISR in the A14 tune. The renormalization scale and factorization scales were varied by factors of 0.5 and 2.0 corresponding to an increase and decrease in ISR, respectively. The $h_{\text{damp}}$ uncertainty is measured by comparing the nominal $t\bar{t}$ sample with a sample using $h_{\text{damp}} = 3m_{\text{top}}$. The impact of final-state radiation (FSR) uncertainties was evaluated by increasing and decreasing the renormalization scale for emissions from the parton shower by a factor of 2.

The impact of using a different parton-shower and hadronization model was evaluated by comparing the nominal $t\bar{t}$ sample with a sample that was also generated by Powheg Box v2 but used Herwig 7.1.3 [104,105] instead of PYTHIA8.230 for parton showering and hadronization. The Herwig 7.1 default set of tuned parameters [105,106] and the MMHT2014LO PDF set [107] were employed.

To assess the uncertainty in the matching of NLO matrix elements to the parton shower, the Powheg $t\bar{t}$ sample was compared with a sample of events generated with MadGraph5_aMC@NLO2.6.0 but retaining the PYTHIA8.230 parton-shower and hadronization models. The MadGraph5_aMC@NLO2.6.0 calculation used the NNPDF3.0NLO set of PDFs, as in the Powheg sample, and PYTHIA8 again used the A14 tune and the NNPDF2.3LO set of PDFs.

The production of a single top quark in association with a $W$ boson ($tW$) was modeled using the Powheg Box v2 generator [84–86,108] at NLO in QCD with the five-flavor scheme and the NNPDF3.0NLO set of PDFs. The diagram removal scheme [109] was used to remove interference and overlap with $t\bar{t}$ production.

The PYTHIA[8.230] parton-shower and hadronization models were employed, using the A14 tune and the NNPDF2.3LO set of PDFs. The inclusive cross section for $tW$ production was corrected to the theory prediction calculated at NLO in QCD with NNLL soft-gluon corrections [110,111].

Single-top-quark $t$-channel production was modeled using the Powheg Box v2 [84–86,112] generator at NLO in QCD using the four-flavor scheme and the corresponding NNPDF3.0NLO set of PDFs. Parton showering and hadronization were performed with PYTHIA[8.230], using the A14 tune and the NNPDF2.3LO set of PDFs. The inclusive cross section was corrected to the theory prediction calculated at NLO in QCD with the Hather2.1 generator [113,114]. Single-top-quark $s$-channel MC events were not generated because the cross section for this process is much smaller than that for $tW$ production and the $t$-channel processes. However, the $s$-channel process makes a small contribution to the data-driven multijet background estimate and is therefore partially accounted for. The production of SM $tH$ is treated in a similar manner, as the background yield is negligible due to a combination of small cross section and low yield in the high-$p_T$ region.

The production of $t\bar{t}$ in association with a Higgs boson ($t\bar{t}H$) was modeled by the Powheg Box v2 [83–86,115] generator at NLO. The production of $t\bar{t}$ in association with a $W$ or $Z$ boson was modeled using the MadGraph5_aMC@NLO2.6.0 generator at NLO. Parton showering and hadronization for these processes was performed by PYTHIA[8.210] and the decays of bottom and charm hadrons were simulated using EvtGen 1.2.0. The cross sections for the $t\bar{t} + W/Z/H$ processes were calculated using MadGraph5_aMC@NLO2.6.0 at NLO QCD and NLO electroweak accuracies using Ref. [116]. The $t\bar{t} + Z$ cross section was corrected to take into account contributions from off-shell $Z$ bosons with masses down to 5 GeV. The predicted values of the cross sections at 13 TeV are $0.88^{+0.09}_{-0.11}, 0.60^{+0.08}_{-0.07}$, and $0.51^{+0.04}_{-0.05}$ pb for $t\bar{t} + Z$, $t\bar{t} + W$, and $t\bar{t} + H$, respectively, where the uncertainties reflect QCD scale variations.

The effect of multiple interactions in the same and neighboring bunch crossings (pile-up) was modeled by overlaying the simulated hard-scattering event with inelastic $pp$ events generated with the PYTHIA[8.186] MC generator [117] using the NNPDF2.3LO set of PDFs and the A3 set of tuned parameters [118].

The detector response was simulated using the GEANT4 framework [119,120], and the data and MC events are reconstructed with the same software algorithms.

**IV. OBJECT RECONSTRUCTION AND EVENT SELECTION**

**A. Object definition**

This analysis makes use of jets, electrons, muons, and event-based quantities formed from their combinations.
Electron candidates are identified from high-quality inner-detector tracks matched to calorimeter energy deposits consistent with an electromagnetic shower [121]. The calorimeter deposits must form a cluster with $E_T > 25$ GeV, $|\eta| < 2.47$, and be outside the transition region $1.37 \leq |\eta| \leq 1.52$ between the barrel and end cap calorimeters. A likelihood-based requirement is used to suppress misidentified jets, and calorimeter- and track-based isolation requirements are imposed using the gradient working point [121], which provides uniform rejection in $\eta$ and improved rejection as $p_T$ increases.

Muon candidates are reconstructed using high-quality inner-detector tracks combined with tracks reconstructed in the muon spectrometer [122]. Only muon candidates with $p_T > 25$ GeV and $|\eta| < 2.5$ are considered. Calorimeter- and track-based isolation criteria similar to those used for electrons are used [123]. To reduce the impact of nonprompt leptons, muons within $\Delta R = 0.4$ of a jet are removed.

The anti-$k_t$ algorithm implemented in the FastJet package [77,78] is used to define three types of jets for this analysis: (1) VRTrack jets with a variable-radius parameter with values between $R = 0.02$ and $R = 0.4$ [124], (2) small-$R$ jets with $R = 0.4$, and (3) large-$R$ jets with $R = 1.0$. These are reconstructed independently of each other. The VRTrack jets make use of tracking information from the inner detector, the large-$R$ jets use information from topological clusters [125] in the calorimeter, and the small-$R$ jets use both tracking information and topological clusters [126].

Only jets that have $p_T > 25$ GeV and $|\eta| < 2.5$ are considered. To reduce pileup effects, the jet-vertex tagger (JVT) algorithm [127] is used to reject small-$R$ jets that do not originate from the primary interaction vertex. The primary vertex is selected as the one with the largest $\Sigma p_T^2$, where the sum is over all tracks with transverse momentum $p_T > 0.5$ GeV that are associated with the vertex. This JVT algorithm is applied only to small-$R$ jets with $p_T < 60$ GeV and $|\eta| < 2.4$.

The topological clusters used as input to the small-$R$ and large-$R$ jet reconstruction are calibrated using the local calibration method [128]. The jet energy scale is energy and $\eta$ dependent with calibration factors derived from simulation and in situ measurements [125,129,130]. The large-$R$ jet candidates are required to have $|\eta| < 2.0$ and $p_T > 350$ GeV. The $\eta$ requirement is imposed to optimize the $T$-quark signal-to-background ratio and to select jets in a kinematic regime where the object tagging is efficient and well understood. The $p_T$ requirement ensures that the large-$R$ jets are sufficiently collimated to contain most of the decay products of the top quark or Higgs boson. A trimming algorithm [131] with parameters $R_{\text{sub}} = 0.2$ and $f_{\text{cut}} = 0.05$ is applied to suppress gluon radiation and further mitigate pileup effects. The small-$R$ jets are used to validate the modeling of large-$R$ jets arising from the $t\bar{t}$ and multijet backgrounds and are not used directly in the event selection. Only small-$R$ jets with $p_T > 25$ GeV and $|\eta| < 2.5$ are considered, so as to match the VRTrack jet candidates.

### B. Higgs boson, top quark, and $b$-jet tagging

This analysis searches for Higgs bosons, top quarks, and $b$-hadron jets ($b$-jets) to identify $T$-quark candidates that undergo a $T \to Ht$ decay, followed by $H \to b\bar{b}$, $t \to Wb$, and $W \to q\bar{q}'$ decays. Distinct tagging algorithms are employed to identify these three objects.

Higgs-boson candidates are identified by requiring the large-$R$ jet mass [128] to be between 100 and 140 GeV, along with an upper bound on the jet-substructure variable $r_{21}$ [132,133], which is a relative measure of whether the jet has a two-body or one-body structure. The upper bound on $r_{21}$ is chosen as a function of the jet $p_T$ in order to achieve a tagging efficiency of 70% for Higgs bosons, independent of their $p_T$. The tagger provides a rejection factor between five and ten for light-quark and gluon jets.

The top-quark-tagging algorithm uses a deep-neural-network (DNN) scheme [48]. It makes use of jet-substructure variables to discriminate between top-quark jets and jets arising from $W$, $Z$, Higgs bosons, gluons, and lighter quarks. An 80% efficiency working point is used, which is defined for all top-quark jets whose decay products are clustered together into the large-$R$ jet. In addition, only jets with a reconstructed mass between 140 and 225 GeV are considered. The orthogonal mass window requirements for tagging Higgs bosons and top quarks ensure that a jet can only be identified as either a top-quark or Higgs-boson candidate.

The $b$-tagging algorithm used is known as DL1, a DNN-based tagging scheme that uses the secondary vertex information and the impact parameters of the charged tracks in a VRTrack jet [49]. The working point chosen for this algorithm results in 70% tagging efficiency for $b$-jets in $t\bar{t}$ events, with a rejection of $\sim 10$ and $\sim 400$ for charm and light quarks, respectively. This algorithm is applied to all VRTrack jets that have been geometrically matched to the large-$R$ jets by requiring that the jet axes have an $\eta - \phi$ distance $\Delta R < 1.0$.

### C. Event preselection

A preselection is performed to obtain a sample of candidate signal and background events. Each event is required to have a primary vertex with five or more associated tracks with $p_T > 0.5$ GeV [134].

To identify the fully hadronic $Ht$ decay topology, events must have at least two large-$R$ jets with $p_T > 350$ GeV and $|\eta| < 2.0$. The highest-$p_T$ jet is required to have $p_T > 500$ GeV to ensure that the inclusive jet trigger used to record the events has 100% efficiency. The two highest-$p_T$ large-$R$ jets are referred to as the leading and
second-leading jets. All other large-\( R \) jets are ignored. The large-\( R \) jets must have a mass between 100 and 225 GeV.

To remove candidates where a \( t\bar{t} \) event has resulted in a lepton + jet or dilepton final state, events are rejected if they have an identified electron or muon candidate, as described in Sec. IVA.

This preselection defines the data sample used in the \( T \)-quark search, which comprises about four million events.

### D. Event classification by tagging states

The leading and second-leading large-\( R \) jet candidates are examined to determine if either jet satisfies the Higgs-boson-tagging or top-quark-tagging criteria. In addition, each VRTrack jet contained within a large-\( R \) jet is examined to determine if it is \( b \)-tagged. In what follows, a \( b \)-tagged VRTrack jet associated with a large-\( R \) jet is referred to as a \( b \)-tag.

With these tagging definitions, the events are classified according to the tagging states of each large-\( R \) jet: the jet could be neither Higgs-boson tagged nor top-quark tagged, be Higgs-boson tagged, or be top-quark tagged. The jet also could have no \( b \)-tags, 1\( b \)-tag, or \( \geq 2 \)\( b \)-tags. Altogether, a large-\( R \) jet could be in one of nine different tagging states, so a \( 9 \times 9 \) matrix is defined as shown in Fig. 2 to categorize all possible tagging states of the two jets in an event.

Three sets of regions are defined: a signal region \( SR \), a \( t\bar{t} \) normalization region \( NR \), and eight validation regions, as illustrated in Fig. 2. The signal region consists of those events where one jet is Higgs-boson tagged with \( \geq 2 \)\( b \)-tags and the other jet is top-quark tagged with \( \geq 1 \)\( b \)-tag, and comprises four event-tagging states as illustrated in Fig. 2.

The \( t\bar{t} \) normalization region is designed to contain the highest-purity sample of \( t\bar{t} \) pair events. The four event-tagging states that define this region are those with both large-\( R \) jets being top-quark tagged and each having at least 1\( b \)-tag. Top-quark-tagged jets with \( \geq 2 \)\( b \)-tags are included, as these typically result from mistagging a charm-quark jet arising from a \( W \rightarrow c\bar{s} \) decay. This region is used to study top-quark-tagging performance and to validate the top-quark acceptance and background estimates.

The validation regions are used to validate the background estimation techniques used in the \( SR \) and \( NR \). The regions with a leading large-\( R \) jet top-quark tagged with 1\( b \)-tag and the second-leading large-\( R \) jet not being top-quark tagged with 1\( b \)-tag (VR1 and VR2) or the second-leading large-\( R \) jet being either Higgs-boson tagged or top-quark tagged with no \( b \)-jets (VR3 and VR4) validate the multijet and non-all-hadronic \( t\bar{t} \) background estimates. The validation regions defined by the event-tagging states with a Higgs-boson-tagged large-\( R \) jet with \( \geq 2 \)\( b \)-tags and with the other jet top-quark tagged with no associated \( b \)-tags (VR5 and VR6) are expected to be dominated by mistagged events and are used to cross-check the mistagging estimates for the Higgs-boson, top-quark, and \( b \)-jet tagging schemes. The validation regions defined by one jet that is neither Higgs-boson tagged nor top-quark tagged with 1\( b \)-tag and the other jet being top-quark tagged with \( \geq 2 \)\( b \)-tags (VR7 and VR8) are considered to validate the background modeling involving \( \geq 3 \)\( b \)-tags.

![FIG. 2. The \( 9 \times 9 \) matrix that represents the 81 exclusive event-tagging states defined by the nine possible tagging states of each large-\( R \) jet. Each event can be in only one of the 81 event-tagging states. The red event-tagging states comprise the events in the signal region (\( SR \)), the blue event-tagging states comprise the \( t\bar{t} \) normalization region (\( NR \)), and the yellow event-tagging states are validation regions labeled \( V1 \) through \( V1 \). The gray event-tagging states are regions used to estimate the multijet background, as described in Sec. VA.](image-url)
V. BACKGROUND ESTIMATION AND VALIDATION

The two largest contributions to the \( SR \) and \( NR \) are multijet events and \( t \bar{t} \) production, with smaller contributions arising from events with only one hadronically decaying top quark or from \( t \bar{t} \) production in association with a \( W \), \( Z \), or Higgs boson. The multijet background in all the regions is estimated using a data-driven technique employing sidebands and control regions dominated by multijet events and originally developed to study boosted \( t \bar{t} \) production [53,58,135]. This background is found to be the largest source of candidate events in the signal region and is determined iteratively, as described in Sec. VA.

The second-largest background in the \( SR \) consists of events in which a pair of boosted top quarks decay hadronically to produce two top-quark jets. This background is estimated using MC calculations normalized by the event yield in the \( NR \) after subtracting other backgrounds. As this subtraction requires an estimate of the multijet background, the background-subtracted event yield is determined iteratively, as described in Sec. VB.

The third-largest contribution in the \( SR \) is from the non-all-hadronic \( t \bar{t} \) process where one top quark decayed semileptonically and the other hadronically. In this case, the final-state leptons are not reconstructed or are misidentified and not rejected by the electron and muon veto. The rate of this process, estimated using MC samples, is normalized using the observed \( t \bar{t} \) event yield in the \( NR \).

Other contributions from SM processes with at least one top-quark jet are estimated using MC samples as described in Sec. VC.

A. The data-driven multijet background estimate

The multijet background is estimated by a data-driven method using events from specifically chosen event-tagging states to estimate the multijet background event yields in the signal, normalization, and validation regions. The event-tagging states used are dominated by multijet backgrounds and have small contributions from events with one or more top quarks. These event-tagging states also have potential contributions from \( T \)-quark production of less than 1% for all choices of \( T \)-quark masses and couplings considered in this search. Contributions from \( W/\gamma + \text{jets} \) are negligible due to a combination of a relatively low cross section for high-\( p_T \) hadronically decaying bosons [136] and the tagging requirements. For a given event-tagging state, the number of events from all MC backgrounds is subtracted from the observed number of events with that event-tagging state. This provides an estimate of the number of multijet events in each event-tagging state. As noted above, the \( t \bar{t} \) background subtracted from each region is normalized to the event yield in the \( NR \) that depends on the multijet estimate in that region. Hence, the multijet background and \( t \bar{t} \) yield are calculated iteratively. This procedure is similar to the algorithm used in Ref. [58].

For example, consider the \( SR \) event-tagging state defined by requiring that the leading jet is top-quark tagged with 1\( b \)-tag and the second-leading jet is Higgs-boson tagged with \( \geq 2 \)-\( b \)-tags. The method uses the numbers of multijet events \( N_A, N_B, N_C, \) and \( N_D, \) after MC background subtraction, in four regions A, B, C, and D. In region A the leading jet is neither top-quark tagged nor Higgs-boson tagged with no \( b \)-tags and the second-leading jet is neither top-quark tagged nor Higgs-boson tagged with \( \geq 2 \)-\( b \)-tags. In region B the leading jet is neither top-quark tagged nor Higgs-boson tagged with \( \geq 2 \)-\( b \)-tags and the second-leading jet is Higgs-boson tagged with \( \geq 2 \)-\( b \)-tags. In region C the leading jet is top-quark tagged with 1\( b \)-tag and the second-leading jet is neither top-quark tagged nor Higgs-boson tagged with no \( b \)-tags. In region D, which is one of the \( SR \) event-tagging states, the leading jet is top-quark tagged with 1\( b \)-tag and the second-leading jet is Higgs-boson tagged with \( \geq 2 \)-\( b \)-tags. If the tagging efficiencies of the two large-\( R \) jets are uncorrelated, then the ratio of the numbers of multijet events in two distinct event-tagging states that differ only by the tagging state for one of the large-\( R \) jets will be independent of the tagging state of the other large-\( R \) jet. In this example, the ratio of \( N_D \) to \( N_C \) is equal to the ratio of \( N_B \) to \( N_A \) since the ratios only differ by the tagging state of the leading large-\( R \) jet. Hence, the number of multijet events in region D is

\[
N_D = \frac{N_B \times N_C}{N_A}.
\]

A corresponding method is performed for each of the event-tagging states of the \( SR \) and \( NR \), and for all the validation regions using different event-tagging states to define regions A, B, and C. Since the \( t \bar{t} \) background subtraction in regions A, B, and C is normalized to the \( t \bar{t} \) event yield in the \( NR \), which requires an estimate of the multijet background, the calculation of the multijet background and the \( t \bar{t} \) event yield in each region is iterated as described in Sec. VB.

The assumption of uncorrelated jet-tagging states is only approximately true. The level of correlation is determined by examining ratios of the numbers of events with specific event-tagging states that do not overlap with the \( SR, NR, \) or validation regions, shown as the gray event-tagging states in Fig. 2. The observed corrections between the jet-tagging states defined by the top-quark, Higgs-boson, and \( b \)-tagging criteria are applied to the multijet background estimates for each of the event-tagging states that define the \( SR, NR, \) and the eight validation regions, with the total corrections varying from 1.01 to 1.10 with uncertainties ranging from 0.03 to 0.06. In the calculation of the multijet background for the event-tagging state belonging to the \( SR \) illustrated above, there are four corrections applied as a
The multijet estimates are calculated independently for each of the four event-tagging states that make up the $SR$ and $NR$, after which they are summed. This provides a fully data-driven multijet background estimate in each region.

For example, to calculate the correlation between the mistagging probabilities when the leading jet is top-quark tagged and the second-leading jet is Higgs-boson tagged, the event yields in four regions, $E$, $F$, $G$, and $H$ are considered. In region $E$ the leading jet is neither top-quark tagged nor Higgs-boson tagged with no $b$-tags and the second-leading jet is neither top-quark tagged nor Higgs-boson tagged with no $b$-tags. In region $F$ the leading jet is top-quark tagged with no $b$-tags and the second-leading jet is neither top-quark tagged nor Higgs-boson tagged with no $b$-tags. In region $G$ the leading jet is neither top-quark tagged nor Higgs-boson tagged with no $b$-tags and the second-leading jet is Higgs-boson tagged with no $b$-tags. In region $H$ the leading jet is top-quark tagged with no $b$-tags and the second-leading jet is Higgs-boson tagged with no $b$-tags. The ratio between the number of events in regions $E$ and $F$ is related to the ratio of events in regions $G$ and $H$ by

$$\frac{N_E}{N_F} = K \frac{N_G}{N_H},$$

where $K$ is the measure of the correlation in mistagging probabilities between the leading jet being top-quark tagged and the second-leading jet being Higgs-boson tagged, with both large-$R$ jets having no associated $b$-tags.

The value of $K$ in this example is $0.976 \pm 0.004$, where the uncertainty is statistical only, and is applied as a correction to the multijet background estimate.

Each correlation is measured in an analogous way using the numbers of events in pairs of event-tagging states. The pairs chosen have MC background contributions less than 8% of the observed event yield, thus reducing the systematic uncertainties arising from the subtraction of the MC background contributions. The multijet background estimate taking into account the tagging correlations is calculated bin-by-bin for each distribution so that the shape of the multijet background distribution is measured as well as the total background event yield.

Since the multijet background depends on the $t\bar{t}$ background subtraction, the two are determined iteratively as described in the next section.

**B. Top-quark pair yields and multijet backgrounds in the $NR$ and $SR$**

Previous measurements of the $t\bar{t}$ differential cross sections for highly boosted top quarks [58] show that the observed cross section is lower than MC predictions by $\sim 20\%$. To avoid the uncertainty this would create in the $t\bar{t}$ background contribution and the multijet estimates in each region, the ratio of the observed rate to the predicted rate of events in the $NR$, $\alpha_{\text{norm}}$, is used to normalize the predicted $t\bar{t}$ background contributions in the $SR$, validation regions, and the event-tagging states used for the multijet estimate. The value of $\alpha_{\text{norm}}$ is determined after the initial multijet estimate that uses the nominal $t\bar{t}$ prediction by requiring the

![FIG. 3. Invariant mass distributions for (a) the leading small-$R$ jet associated with the leading large-$R$ jet in the $t\bar{t}$ normalization region where both the leading and second-leading large-$R$ jets are top-quark tagged with at least 1$b$-tag and (b) the leading large-$R$ jet, in the same region. The predicted distributions include the estimated backgrounds and a hypothetical $T$-quark signal with $m_T = 1.6$ TeV and $\kappa_T = 0.5$. The blue hashed lines correspond to the sum in quadrature of the statistical and systematic uncertainties of the prediction in a given bin. The lower panels show the ratio of the data to the prediction, along with the uncertainty in the ratio. A ratio outside the bounds of the axis is represented by a blue arrow. The last bin includes the event overflows. Contributions to the predicted yield are stacked in the same order as they appear in the legend.](https://example.com/fig3)
predicted event yield in the NR to match the observed yield. However, the multijet estimate itself is a function of $\alpha_{\text{norm}}$, as the estimation technique described in the previous section requires the subtraction of the $t\bar{t}$ background contribution in the multijet-dominated event-tagging states during its calculation. Thus, both the multijet estimate and $\alpha_{\text{norm}}$ are calculated iteratively using

$$\alpha_{\text{norm}}^{n+1} = \frac{N_{\text{Data}} - N_{\text{Multijet}}(\alpha_{\text{norm}}^n) - N_{\text{top-related}}} {N_{\text{MC}}},$$

where $\alpha_{\text{norm}}^n$ is the value of $\alpha_{\text{norm}}$ resulting from the $n$th iteration, $N_{\text{Data}}$ is the observed event yield in the NR, $N_{\text{Multijet}}(\alpha_{\text{norm}}^n)$ is the data-driven multijet background event yield from the $n$th iteration in the NR, $N_{\text{top-related}}$ is the sum of the backgrounds from single-top-quark, $t\bar{t} + W$, $t\bar{t} + Z$, and $t\bar{t} + H$ production that are estimated by MC calculations in the NR, and $N_{\text{MC}}$ is the sum of the $t\bar{t}$ events with all-hadronic and non-all-hadronic decays in the NR.

In each iteration of the multijet estimate $N_{\text{MC}}$ is scaled by $\alpha_{\text{norm}}^{n+1}$ before subtraction. This calculation converges to

![Invariant mass distributions for the leading small-$R$ jet associated with the leading large-$R$ jet for (a) $VR_1$ defined by requiring the leading large-$R$ jet be top-quark tagged with $1b$-tag and the second-leading jet is Higgs-boson tagged with $1b$-tag, (b) $V_1$ defined by requiring the leading large-$R$ jet be top-quark tagged with $1b$-tag and the second-leading jet is neither Higgs-boson tagged nor top-quark tagged with $1b$-tag, (c) $V_1$ defined by requiring the leading large-$R$ jet be top-quark tagged with $1b$-tag and the second-leading jet is top-quark tagged with no $b$-tag, and (d) $V_1$ defined by requiring the leading large-$R$ jet be top-quark tagged with $1b$-tag and the second-leading jet is Higgs-boson tagged with no $b$-tag. The predicted distribution includes the estimated backgrounds and a hypothetical $T$-quark signal with $m_T = 1.6$ TeV and $\kappa_T = 0.5$. The blue hashed lines correspond to the sum in quadrature of the statistical and systematic uncertainties of the prediction in a given bin. The lower panels show the ratio of the data to the prediction, along with the uncertainty in the ratio. A ratio outside the bounds of the axis is represented by a blue arrow. The last bin includes the event overflows. Contributions to the predicted yield are stacked in the same order as they appear in the legend.

FIG. 4.
subpercent level in four iterations to a value of $\alpha_{\text{norm}} = 0.82 \pm 0.01$, where only statistical uncertainties are considered. This is consistent with cross section measurements of boosted $t\bar{t}$ production [57]. The $t\bar{t}$ contribution predicted by the MC calculations in the SR is scaled by $\alpha_{\text{norm}}$. The resulting $t\bar{t}$ yield estimates are $8587 \pm 1369$ events and $174 \pm 35$ events in the NR and SR, respectively, where the uncertainties include the systematic uncertainties described in Sec. VI. This estimate of the $t\bar{t}$ yield in the NR is used only for the iterative multijet background estimate.

The multijet yields in the NR and SR after this iterative calculation are estimated to be $1452 \pm 57$ and $316 \pm 9$ events, respectively. The uncertainties in the multijet estimates, including the uncertainties in the tagging correlations, consist of the statistical uncertainties in the event-tagging states used for the calculation and the systematic

---

**FIG. 5.** Dijet invariant mass distributions for the two large-$R$ jets in four validation regions: (a) VR5 defined by requiring a leading jet Higgs-boson tagged with $\geq 2b$-tags and second-leading jet top-quark tagged with no associated $b$-tag, (b) VR6 defined by requiring a leading jet top-quark tagged with no associated $b$-tag and second-leading jet Higgs-boson tagged with $\geq 2b$-tags, (c) VR7 defined by requiring a leading jet top-quark tagged with $\geq 2b$-tags and second-leading jet neither top-quark tagged nor Higgs-boson tagged with $1b$-tag, and (d) VR8 defined by requiring a leading jet neither top-quark tagged nor Higgs-boson tagged with $1b$-tag and second-leading jet top-quark tagged with $\geq 2b$-tags. The predicted distributions include the estimated backgrounds and a hypothetical $T$-quark signal with $m_T = 1.6$ TeV and $\kappa_T = 0.5$. The blue hashed lines correspond to the sum in quadrature of the statistical and systematic uncertainties of the prediction in a given bin. The lower parts of each panel show the ratio of the data to the prediction, along with the uncertainty in the ratio. A ratio outside the bounds of the axis is represented by a blue arrow. The last bin includes the event overflows. Contributions to the predicted yield are stacked in the same order as they appear in the legend.
uncertainties arising from the MC background subtraction, as described in Sec. VI.

C. Other top-quark backgrounds

Single-top-quark production in the $Wt$ final state and the $t$-channel represent a small contribution to the total background prediction, which is estimated using the Powheg-+Pythia8 MC calculation described in Sec. III. The $s$-channel single-top-quark process is not included as an explicit contribution because of its small cross section and because a part of it is already taken into account in the data-driven multijet estimate. The uncertainty in the single-top-quark background is increased by 50% to account for the uncertainty in this contribution.

The estimated single-top-quark yields in the $\bar{t}t$ NR and SR are $93 \pm 52$ and $8 \pm 6$ events, respectively.

The backgrounds from production of a top-quark pair in association with a $W$, $Z$, or Higgs boson are also estimated using the MC samples described in Sec. III. The estimated yields in the NR and SR are $115 \pm 25$ events and $9 \pm 2$ events, respectively.

D. Validation of background calculations

Kinematic variables with the ability to distinguish between $\bar{t}t$ and multijet contributions in the NR and the validation regions are examined to further validate the background modeling. The potential contribution of $T$-quark production to these regions is $<1\%$. The distributions of the mass of the leading small-$R$ jet associated with the leading large-$R$ jet events in the NR where both large-$R$ jets are top-quark tagged and have $\geq 1b$-tags are shown in Fig. 3(a). A $W$-mass peak is observed, which arises when the $W$-boson decay products are collimated into a small-$R$ jet, along with a low-mass peak arising from light quarks and bottom quarks. A shoulder is seen around the top-quark mass, which arises from a small number of highly boosted top-quark jets where all the decay products of the top quark are clustered into the small-$R$ jet. The observed distribution is well modeled with a large $\bar{t}t$ contribution and a smaller multijet distribution. The invariant mass distribution of the leading large-$R$ jet in the same sample, shown in Fig. 3(b), confirms the interpretation that this region is dominated by $\bar{t}t$ production.

The distributions of the jet mass for the leading small-$R$ jet associated with the leading large-$R$ jet are shown in Fig. 4 for validation regions VR1–VR4. The relative sizes of the $\bar{t}t$ and multijet contributions vary between these validation regions, further testing the robustness of their modeling and normalization. There is agreement between the observed and predicted distributions in both normalization and shape, except for a small excess in the prediction of events for VR1. This is further discussed in Sec. VI B.

Further validation of the multijet background estimates is illustrated in Fig. 5, where the distributions of the invariant mass of the two leading jets, or dijet system, are shown for the four validation regions dominated by multijet backgrounds. The distributions for events with a top-quark-tagged jet with no $b$-tags and a Higgs-boson-tagged jet with $\geq 2b$-tags (VR5 and VR6) are shown in Figs. 5(a) and 5(b), respectively. Distributions for events with a jet with $1b$-tag, but no top-quark or Higgs-boson tag, and another jet with a top-quark tag and $\geq 2b$-tags (VR7 and VR8) are shown in Figs. 5(c) and 5(d), respectively. There is also agreement between the observed and predicted distributions.

VI. SYSTEMATIC UNCERTAINTIES

Systematic uncertainties that affect the interpretation of the data are estimated using data and MC samples. Table I. Size of the postfit uncertainties in the $T$-quark signal cross section for a $T$-quark mass of 1.6 TeV and coupling $\kappa_T = 0.5$. The fitted cross section is $-10$ fb and is consistent with zero. The background uncertainty is the sum in quadrature of the systematic uncertainty on the multijet background and the statistical uncertainties on the MC-derived backgrounds. The total uncertainty of 25 fb is the sum in quadrature of the total systematic uncertainty and statistical uncertainty. The uncertainty arising from simultaneously fitting the $\bar{t}t$ normalization factor is included in the total systematic uncertainty. The individual uncertainties do not add up in quadrature to the total uncertainty because of their correlations in the fit.

<table>
<thead>
<tr>
<th>Category</th>
<th>Uncertainty in $\sigma$ $(pp \rightarrow T + X \rightarrow Ht + X)$ (fb)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Detector uncertainties</td>
<td></td>
</tr>
<tr>
<td>$b$-jet tagging</td>
<td>6.1</td>
</tr>
<tr>
<td>Top-quark jet tagging</td>
<td>5.9</td>
</tr>
<tr>
<td>Jet mass resolution</td>
<td>3.0</td>
</tr>
<tr>
<td>Jet mass scale</td>
<td>2.3</td>
</tr>
<tr>
<td>Jet energy scale</td>
<td>1.8</td>
</tr>
<tr>
<td>Jet energy resolution</td>
<td>1.7</td>
</tr>
<tr>
<td>Higgs-boson tagging</td>
<td>1.6</td>
</tr>
<tr>
<td>Other detector uncertainties</td>
<td>0.3</td>
</tr>
<tr>
<td>Modeling uncertainties</td>
<td></td>
</tr>
<tr>
<td>Other $\bar{t}t$ modeling uncertainties</td>
<td>4.9</td>
</tr>
<tr>
<td>$\bar{t}t$ parton shower and hadronization</td>
<td>1.9</td>
</tr>
<tr>
<td>$\bar{t}t$ matrix element</td>
<td>2.4</td>
</tr>
<tr>
<td>Background uncertainty</td>
<td>7.3</td>
</tr>
<tr>
<td>Signal MC statistical uncertainty</td>
<td>4.9</td>
</tr>
<tr>
<td>$\bar{t}t$ normalization ($\alpha_\beta$)</td>
<td>1.5</td>
</tr>
<tr>
<td>Other top-quark-background theory uncertainties</td>
<td>1.8</td>
</tr>
<tr>
<td>Total uncertainties</td>
<td></td>
</tr>
<tr>
<td>Total statistical uncertainty</td>
<td>19</td>
</tr>
<tr>
<td>Total systematic uncertainty</td>
<td>15</td>
</tr>
<tr>
<td>Total uncertainty</td>
<td>25</td>
</tr>
</tbody>
</table>
Variations corresponding to a $+1\sigma$ and $-1\sigma$ confidence interval are derived for each uncertainty.

These systematic uncertainties are broken down into detector-related and modeling uncertainties. They do not have a significant dependence on the choice of $T$-quark mass and coupling, so an example of the size of the systematic uncertainties arising from the likelihood fit described in Sec. VII (the "postfit" results) is provided in Table I for $m_T = 1.6$ TeV and $\kappa_T = 0.5$.

### A. Detector-related uncertainties

The most significant detector-related systematic uncertainties arise from the measurements of jet properties and tagging efficiencies.

Uncertainties associated with the large-$R$ jets arise from the jet energy scale (JES), jet mass scale (JMS), jet mass response (JMR), jet energy resolution (JER), and the JVT requirement. The uncertainties in the JES, JMS, and JMR are evaluated by using in situ measurements [125]. The JES is measured in events where a large-$R$ jet recoils against well-defined reference objects (photons, $Z$ bosons, or calibrated small-$R$ jets). The JMS and JMR uncertainties are measured using both a double-ratio method that compares the calorimeter-to-tracker response ratio between data and simulation [125] and a fit to the $W$-boson mass peak in high-$p_T$ lepton + jets $t\bar{t}$ events. The JER uncertainty is measured by studying dijet mass resolution and the effect of energy flow near the jet radius [128]. The JVT uncertainty arises from the correction factors used to match the efficiencies in the MC samples to data.

The efficiency for tagging $b$-jets is measured in data using dilepton $t\bar{t}$ events [49]. Correction factors are applied to the jets in the MC sample so that the $b$-jet tagging efficiency as a function of jet $p_T$ in MC events matches that in data events. Uncertainties arising in the evaluation of the efficiencies are propagated to the correction factors. The largest source of $b$-jet tagging uncertainty is the extrapolation of tagging efficiencies to $b$-jets with $p_T > 300$ GeV, as $b$-jet tagging calibrations use data with $p_T < 300$ GeV.

A $T$-quark hypothesis with $m_T = 1.6$ TeV and $\kappa_T = 0.5$ is used in these plots. The blue hashed lines correspond to the sum in quadrature of the statistical and systematic uncertainties of the prediction in a given bin. The lower panels show the ratio of the data to the prediction, along with the uncertainty in the ratio. A ratio outside the bounds of the axis is represented by a blue arrow. The last bin includes the event overflows. Contributions to the predicted distributions are stacked in the same order as they appear in the legend.
The $t\bar{t}$ background estimate has systematic uncertainties from initial/final-state radiation (ISR/FSR), the renormalization scale, factorization scale, PDF, parton-shower algorithm, matrix-element calculation, and $h_{\text{damp}}$ parameter value. The effects of ISR/FSR, renormalization scale, and factorization scale uncertainties are evaluated using the method described in Sec. III. The PDF uncertainties are evaluated by use of the PDF4LHC15 Hessian uncertainties, where the 30 variations are combined into one nuisance parameter. Uncertainties arising from the choice of parton-shower and hadronization algorithms are evaluated by comparing the nominal Powheg+Pythia8 sample with the Powheg+Herwig sample. The uncertainty arising from the matrix-element calculation is assessed by comparing the nominal MC sample with the MadGraph5_aMC@NLO2.6.0+Pythia8 sample.

Although the non-all-hadronic $t\bar{t}$ background is relatively small in the SR and NR, a 5% excess of predicted events relative to the data is observed in VR1 defined by the event-tagging state with the leading large-$R$ jet top-quark tagged with $1b$-tag and the second-leading jet Higgs-boson tagged with $1b$-tag. This validation region is estimated to have a non-all-hadronic $t\bar{t}$ background fraction of approximately 15% and it is possible that the observed excess is due to mismodeling of this background. A conservative uncertainty of 62%, which covers the excess if it is attributed entirely to the non-all-hadronic

| TABLE II. The event yields are shown for the $t\bar{t}$ normalization region and for the signal region after the likelihood fit (postfit). The requirement of the $Ht$ invariant mass being greater than 1 TeV has been imposed on the signal region. The postfit yields incorporate the statistical, detector, and modeling uncertainties as well as their correlations. |
|---------------------------------|---------------------------------|---------------------------------|
| Region                          | $t\bar{t}$ normalization         | $Ht$ signal region              |
|                                 | Postfit                          | Postfit                          |
| $t\bar{t}$ all-hadronic         | 8366 ± 216                       | 147 ± 17                         |
| $t\bar{t}$ non-all-hadronic     | 189 ± 133                        | 14 ± 10                          |
| Single top quark                | 92 ± 49                          | 8 ± 6                            |
| $t\bar{t} + W/Z/H               | 117 ± 25                         | 9 ± 2                            |
| Multijet events                 | 1452 ± 57                        | 316 ± 9                          |
| Signal events ($m_T = 1.6$ TeV, $\kappa_T = 0.5$) | $-9 ± 21$                        |                                  |
| Predicted background            | 10216 ± 150                      | 494 ± 22                         |
| Data (139 fb$^{-1}$)            | 10231                            | 471                              |
The dijet invariant mass formed from the tagged large-\(R\) jets is interpreted as a combination of the expected SM backgrounds and a \(T\)-quark signal. The dijet mass in the \(SR\) is the invariant mass of the Higgs-boson and top-quark candidates while in the \(NR\) it is the invariant mass of the two top-quark candidates. The dijet invariant mass distributions for the \(SR\) and \(NR\) are shown in Fig. 6, assuming a

**FIG. 8.** Observed and expected 95% C.L. upper limits on the single \(T\)-quark production cross section as a function of \(m_T\) for values of \(\kappa_T\) ranging from 0.1 to 1.1 in steps of 0.2 for figures (a) through (f), respectively. The green (yellow) band is the 68\% (95\%) confidence interval around the median expected limit, as determined using pseudo-experiments. The predicted cross sections of single \(T\)-quark production are shown in red.
T-quark signal contribution with $m_T = 1.6$ TeV and $\kappa_T = 0.5$ scaled to the theory cross section of 41 fb. The overall acceptance times efficiency of T-quark detection in the all-hadronic final state is 1.6% for this choice of mass and couplings, taking into account the kinematic requirements and tagging efficiencies. The predicted background rates and shapes are in good agreement with the observed distributions.

The dijet mass is used as a discriminant in the SR and NR to test for the presence of a T-quark signal. Two parameters of interest are defined: $\sigma_{\text{obs}}$, the observed cross section for single production of a T-quark, and $\alpha$, the SR and NR $t\bar{t}$ background normalization.

A binned-likelihood fit is performed in which a T-quark signal and the background model is fitted to the SR dijet mass distribution and simultaneously the NR background model is fitted to the NR dijet mass distribution. The fit is performed for events with a dijet mass greater than 1 TeV. The fit model in the SR is the sum of the background distributions and a T-quark signal distribution with a given mass, coupling, and signal cross section $\sigma_{\text{obs}}$. In the NR the very small contribution from the T-quark signal is

**FIG. 9.** Observed and expected 95% C.L. upper limits on the single T-quark coupling $\kappa_T$ as a function of $m_T$ are shown as solid and dashed lines, respectively. The green (yellow) band is the 68% (95%) confidence interval around the median expected limit, as determined using pseudo-experiments. All values of $\kappa_T$ above the solid line are excluded. The dashed curves represent contours of fixed $\Gamma/m_T$.

**FIG. 10.** Observed (a) and expected (b) 95% C.L. lower limits on the T-quark mass as a function of the T-quark width-to-mass ratio and the branching fraction of the $T \to Ht$ decay ($\Gamma_T$ is the T-quark width). The branching fractions ($B$) for $T \to Ht$ and $T \to Zt$ are kept equal. The branching fraction for $T \to Wb$ is $1 - B(T \to Ht) - B(T \to Zt)$. The color scale on the right side of each plot defines the 95% C.L. limit on the T-quark mass. Masses below the observed limit are excluded. The dashed white contour lines denote isolines of equal exclusion on the mass in units of TeV.
neglected. The signal cross section is allowed to take negative values in the fit whereas $\sigma_{\text{fit}}$ is constrained to be positive. The fit of the $t\bar{t}$ background in the NR and SR measures $\sigma_{\text{fit}}$ using both regions and thus provides a scaled $t\bar{t}$ background contribution in the SR.

The fit incorporates the systematic uncertainties as Gaussian nuisance parameters. Additional bin-by-bin uncertainties are included to account for the statistical uncertainties in the predicted multijet and MC backgrounds. The $t\bar{t}$ contributions to the NR and SR are fully correlated in the fit. The likelihood is then profiled [138] as a function of each nuisance parameter and used as the test statistic to determine the statistical significance of the fit results.

Figure 7 shows the dijet mass distributions for the SR and NR after the fit (postfit) assuming a signal hypothesis with $m_T = 1.6$ TeV and $\kappa_T = 0.5$. The observed and predicted event yields in the NR and SR are given in Table II. The fitted value of $\sigma_{\text{fit}} = 0.79 \pm 0.12$ is consistent with the $t\bar{t}$ normalization factor $\sigma_{\text{norm}} = 0.82 \pm 0.01$ determined from the background-subtracted event yield in the NR (the uncertainty on $\sigma_{\text{norm}}$ is statistical only). There is good agreement between the predicted postfit signal region background yield of $494 \pm 22$ events and the observed yield of 471 events, consistent with no significant excess in data above SM backgrounds over the entire $Ht$ invariant mass distribution as seen in Fig. 7(a). The fit of the $m_T = 1.6$ TeV and $\kappa_T = 0.5$ signal hypothesis results in $\sigma(pp \rightarrow T + X \rightarrow Ht + X) = -10 \pm 25$ fb, further confirming no excess of events at $Ht$ masses around 1.6 TeV.

Similarly, fit results with $T$-quark cross sections consistent with zero are obtained for $T$-quark masses between 1.0 and 2.3 TeV and for $\kappa_T$ values from 0.1 to 1.6. Based on these fit results, for $1.0 < m_T < 2.3$ TeV there is no significant evidence of a $T$ quark decaying to the $Ht$ final state.

The fit results are used to set 95% C.L. upper limits on the single-$T$-quark production cross section for $1.0 < m_T < 2.3$ TeV and $0.1 < \kappa_T < 1.6$ using the C.L.s method [139]. The predicted cross sections assume a singlet $T$ quark with a $T \rightarrow Ht$ branching fraction of 1/4. Figure 8 shows the 95% C.L. upper limits as a function of $m_T$ for different values of $\kappa_T$. The cross section limits range from $\sim 10$ to $\sim 200$ fb, depending on $\kappa_T$. The decrease in sensitivity for masses from 1.0 to 1.2 TeV arises from the change in signal shape due to the $p_T$ requirements on the Higgs-boson and top-quark candidate jets. The $p_T$ requirements shape the distribution to peak at roughly 1.2 TeV, which can be seen in Figs. 6 and 7. Figure 9 shows the exclusion limits as a function of $m_T$ and $\kappa_T$. Figure 10 shows the observed and expected 95% C.L. limits on the $T$-quark mass as a function of the $T$-quark width-to-mass ratio $\Gamma/m_T$ and the branching fraction for $T$-quark decay into a Higgs boson and a top quark.

For the considered mass range of 1.0 to 2.3 TeV the upper limit on allowed values of $\kappa_T$ rises from a minimum value of 0.3 starting at $m_T = 1.1$ TeV, up to 1.6 for $m_T = 2.3$ TeV. At 95% C.L., this analysis excludes $T$ quarks with $\Gamma/m_T \geq 0.05$ for $1.05 < m_T < 1.2$ TeV, with the mass limits rising with $\Gamma/m_T$ to exclude $m_T < 1.7$ TeV for $\Gamma/m_T \geq 0.5$.

**VIII. CONCLUSION**

A search is reported for the single production of a vectorlike singlet $T$ quark decaying into a Higgs boson and a top quark both of which decay hadronically. The search uses 139 fb$^{-1}$ of 13 TeV proton-proton collision data collected with the ATLAS detector at the LHC. The final states are fully reconstructed by clustering the decay products into two large-$R$ jets. The use of fully hadronic decays allows the direct reconstruction of the $T$-quark final state, increasing the signal-to-background ratio for the search. The results significantly extend the sensitivity for the production of $T$ quarks decaying fully hadronically. The search sensitivity is further improved by a larger dataset than used previously, tagging techniques with greater background rejection, and a data-driven multijet background estimate that reduces the uncertainty in the background modeling. The cross section upper limits are typically a factor of 2 lower than previous searches.

The analysis is performed by searching for an excess above SM backgrounds in the $Ht$ invariant mass distribution. This distribution shows no evidence of significant contributions from single $T$-quark production and is consistent with the expected SM background sources. Therefore, limits are set at 95% C.L. on the production cross section of a $T$ quark decaying to the $Ht$ final state. These depend on the $T$-quark mass and coupling to SM particles and range from $\sim 10$ to $\sim 200$ fb, depending on the assumed $\kappa_T$ value for the couplings. In the resonance mass range between 1.0 and 2.3 TeV, the upper limit on the allowed coupling values rises with $m_T$ from a minimum value of 0.3 for $m_T = 1.1$ TeV to 1.6 for $m_T = 2.3$ TeV. This analysis excludes $T$ quarks with $\Gamma/m_T \geq 0.05$ for $1.05 < m_T < 1.2$ TeV, with the mass limits rising with $\Gamma/m_T$ to exclude $m_T < 1.7$ TeV for $\Gamma/m_T \geq 0.5$.

These results provide significantly improved mass and coupling limits on vectorlike quark models involving a $T$ quark decaying into a Higgs boson and a top quark. The exclusion limits set by this analysis extend the limits set by previous searches.

**ACKNOWLEDGMENTS**

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; BMWFW and FWF, Austria; ANAS, Azerbaijan; SSTC, Belarus; CNPq and
FAPESP, Brazil; NSERC, NRC and CFI, Canada; CERN; ANID, Chile; CAS, MOST and NSFC, China; Minciencias, Colombia; MEYS CR, Czech Republic; DNRF and DNSRC, Denmark; IN2P3-CNRS and CEA-DRF/IRFU, France; SRNSFG, Georgia; BMBF, HGF and MPG, Germany; GSR, Greece; RGC and Hong Kong SAR, China; IF and Benoziyo Center, Israel; INFN, Italy; MEXT and JSPS, Japan; NWO, Netherlands; RCN, Norway; MEiN, Poland; FCT, Portugal; MNE/IFA, Romania; JINR; MES of Russia and NRC KI, Russian Federation; MESTD, Serbia; MSSR, Slovakia; ARRS and MIZŠ, Slovenia; DSI/INRF, South Africa; MICINN, Spain; SRC and Wallenberg Foundation, Sweden; SERI, SNSF and Cantons of Bern and Geneva, Switzerland; MOST, Taiwan; TAEK, Turkey; STFC, United Kingdom; DOE and NSF, U.S. In addition, individual groups and members have received support from BCKDF, CANARIE, Compute Canada and CRC, Canada; COST, ERC, ERDF, Horizon 2020 and Marie Skłodowska-Curie Actions, European Union; Investissements d’Avenir Labex, Investissements d’Avenir Idex and ANR, France; DFG and AvH Foundation, Germany; Herakleitos, Thales and Aristeia programs cofinanced by EU-ESF and the Greek NSRF, Greece; BSF-NSF and GIF, Israel; Norwegian Financial Mechanism 2014-2021, Norway; NCN and NAWA, Poland; La Caixa Banking Foundation, CERCA Programme Generalitat de Catalunya and PROMETEO and GenT Programmes Generalitat Valenciana, Spain; Göran Gustafssons Stiftelse, Sweden; The Royal Society and Leverhulme Trust, United Kingdom. The crucial computing support from all WLCG partners is acknowledged gratefully, in particular from CERN, 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 (U.S.), the Tier-2 facilities worldwide and large non-WLCG resource providers. Major contributors of computing resources are listed in Ref. [140].

[21] ATLAS Collaboration, Search for Pair Production of a New b’ Quark that Decays into a Z Boson and a Bottom


[26] CMS Collaboration, Search for pair production of a new heavy quark that decays into a \(W\) boson and a light quark in \(pp\) collisions at \(\sqrt{s} = 8\) TeV with the ATLAS detector, Phys. Rev. D 92, 112007 (2015).


[33] ATLAS Collaboration, Search for pair production of heavy vector-like quarks decaying to high-\(p_T\) \(W\) bosons and \(b\) quarks in the lepton-plus-jets final state in \(pp\) collisions at \(\sqrt{s} = 13\) TeV with the ATLAS detector, J. High Energy Phys. 10 (2017) 141.


their transverse momentum in $\sqrt{s} = 8$ TeV proton-proton collisions using the ATLAS detector, Phys. Rev. D 93, 032009 (2016).


[60] CMS Collaboration, Measurement of the semileptonic $t\bar{t}+\gamma$ production cross section in pp collisions at $\sqrt{s} = 8$ TeV, J. High Energy Phys. 10 (2017) 006.


1Department of Physics, University of Adelaide, Adelaide, Australia
2Department of Physics, University of Alberta, Edmonton, Alberta, Canada
3aDepartment of Physics, Ankara University, Ankara, Turkey
3bDivision of Physics, TOBB University of Economics and Technology, Ankara, Turkey
4LAPP, Univ. Savoie Mont Blanc, CNRS/IN2P3, Annecy, France
5APC, Université Paris Cité, CNRS/IN2P3, Paris, France
6High Energy Physics Division, Argonne National Laboratory, Argonne, Illinois, USA
7Department of Physics, University of Arizona, Tucson, Arizona, USA
8Physics Department, National and Kapodistrian University of Athens, Athens, Greece
9Physics Department, National Technical University of Athens, Zografou, Greece
10Department of Physics, University of Texas at Austin, Austin, Texas, USA
11Institute of Physics, Azerbaijan Academy of Sciences, Baku, Azerbaijan
12Institut de Física d’Altes Energies (IFAE), Barcelona Institute of Science and Technology, Barcelona, Spain
13Institut für High Energy Physics, Chinese Academy of Sciences, Beijing, China
14a Physics Department, Tsinghua University, Beijing, China
14bDepartment of Physics, Nanjing University, Nanjing, China
14cSchool of Physical Science and Technology, University of Science and Technology of China, Hefei, China
14dPhysics Department, University of Belgrade, Belgrade, Serbia
15Department for Physics and Technology, University of Bergen, Bergen, Norway
16Physics Division, Lawrence Berkeley National Laboratory and University of California, Berkeley, California, USA
17Institut für Physik, Humboldt Universität zu Berlin, Berlin, Germany
18Albert Einstein Center for Fundamental Physics and Laboratory for High Energy Physics, University of Bern, Bern, Switzerland
19School of Physics and Astronomy, University of Birmingham, Birmingham, United Kingdom
20Department of Physics, Bogazici University, Istanbul, Turkey
21aDepartment of Physics Engineering, Gaziantep University, Gaziantep, Turkey
21bDepartment of Physics, Istanbul University, Istanbul, Turkey
22aFaculty of Sciences and Center for Investigations, Universidad Antonio Nariño, Bogotá, Colombia

SEARCH FOR SINGLE PRODUCTION OF A VECTORLIKE T … PHYS. REV. D 105, 092012 (2022)
SEARCH FOR SINGLE PRODUCTION OF A VECTORLIKE $T$ …

PHYS. REV. D 105, 092012 (2022)
Deceased.

Also at Instituto de Fisica Teorica, IFT-UAM/CSIC, Madrid, Spain.
Also at Institute of Physics, Azerbaijan Academy of Sciences, Baku, Azerbaijan.
Also at TRIUMF, Vancouver BC, Canada.
Also at Physics Department, An-Najah National University, Nablus, Palestinian Authority.
Also at Department of Physics, University of Fribourg, Fribourg, Switzerland.
Also at Department of Physics and Astronomy, University of Louisville, Louisville, Kentucky, USA.
Also at Departament de Fisica de la Universitat Autonoma de Barcelona, Barcelona, Spain.
Also at Moscow Institute of Physics and Technology State University, Dolgoprudny, Russia.
Also at Department of Physics, Ben Gurion University of the Negev, Beer Sheva, Israel.
Also at Università di Napoli Parthenope, Napoli, Italy.
Also at Institute of Particle Physics (IPP), Canada.
Also at Bruno Kessler Foundation, Trento, Italy.
Also at Department of Physics, St. Petersburg State Polytechnical University, St. Petersburg, Russia.
Also at Borough of Manhattan Community College, City University of New York, New York, New York, USA.
Also at Department of Financial and Management Engineering, University of the Aegean, Chios, Greece.
Also at Centro Studi e Ricerche Enrico Fermi, Italy.
Also at Department of Physics, California State University, East Bay, Hayward, USA.
Also at Institucio Catalana de Recerca i Estudis Avancats, ICREA, Barcelona, Spain.
Also at Graduate School of Science, Osaka University, Osaka, Japan.
Also at Physikalisches Institut, Albert-Ludwigs-Universität Freiburg, Freiburg, Germany.
Also at University of Chinese Academy of Sciences (UCAS), Beijing, China.
Also at Joint Institute for Nuclear Research, Dubna, Russia.
Also at Yeditepe University, Physics Department, Istanbul, Turkey.
Also at Institute of Theoretical Physics, Ilia State University, Tbilisi, Georgia.
Also at CERN, Geneva, Switzerland.
Also at Hellenic Open University, Patras, Greece.
Also at Center for High Energy Physics, Peking University, China.
Also at The City College of New York, New York, New York, USA.
Also at Department of Physics, California State University, Sacramento, USA.
Also at D´epartement de Physique Nucléaire et Corpusculaire, Université de Genève, Genève, Switzerland.
Also at Faculty of Physics, M.V. Lomonosov Moscow State University, Moscow, Russia.
Also at Institut für Experimentalphysik, Universität Hamburg, Hamburg, Germany.
Also at National Research Nuclear University MEPhI, Moscow, Russia.
Also at Department of Physics and Astronomy, Michigan State University, East Lansing MI, USA.