Search for phenomena beyond the Standard Model in events with large b-jet multiplicity using the ATLAS detector at the LHC

ATLAS Collaboration

DOI
10.1140/epjc/s10052-020-08730-0

Publication date
2021

Document Version
Final published version

Published in
European Physical Journal C

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).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

UvA-DARE is a service provided by the library of the University of Amsterdam (https://dare.uva.nl)
Search for phenomena beyond the Standard Model in events with large $b$-jet multiplicity using the ATLAS detector at the LHC

ATLAS Collaboration*
CERN, 1211 Geneva 23, Switzerland

Received: 2 October 2020 / Accepted: 4 December 2020 / Published online: 11 January 2021
© CERN for the benefit of the ATLAS collaboration 2021, corrected publication 2021

Abstract A search is presented for new phenomena in events characterised by high jet multiplicity, no leptons (electrons or muons), and four or more jets originating from the fragmentation of $b$-quarks ($b$-jets). The search uses 139 fb$^{-1}$ of $\sqrt{s} = 13$ TeV proton–proton collision data collected by the ATLAS experiment at the Large Hadron Collider during Run 2. The dominant Standard Model background originates from multijet production and is estimated using a data-driven technique based on an extrapolation from events with low $b$-jet multiplicity to the high $b$-jet multiplicities used in the search. No significant excess over the Standard Model expectation is observed and 95% confidence-level limits that constrain simplified models of R-parity-violating supersymmetry are determined. The exclusion limits reach 950 GeV in top-squark mass in the models considered.

1 Introduction

Events with a large number of high-transverse momentum ($p_T$) jets originating from the fragmentation of $b$-quarks ($b$-jets) are rarely produced by Standard Model (SM) processes in proton–proton ($pp$) collisions at the LHC. As a result, this signature can provide sensitivity to certain phenomena whose nature depends on which of the many RPV couplings are non-zero.

In the analysis presented here, a particular benchmark model is considered in order to interpret the measurements in the different jet and $b$-jet multiplicity regions. The process under consideration is the pair production of the top squark as the lightest of the coloured SUSY partners. The existence of light SUSY partners of third-generation quarks, bottom squarks ($\tilde{b}$) and top squarks ($\tilde{t}$), is favoured by naturalness considerations [7,8]. The scenario assumes the LSP to be a triplet of two neutralino ($\tilde{\chi}_1^0, \tilde{\chi}_2^0$) and one chargino ($\tilde{\chi}_1^\pm$) states that are mass-degenerate and carry dominantly higgsino components (in the following collectively referred to as “higgsinos”). The top squark decays either into a higgsino, $\tilde{\chi}_1^\pm$, and a bottom quark or into a neutralino, $\tilde{\chi}_{1,2}^0$, and a top quark. The chargino and neutralino decay, respectively, to $b$s and $t$s quark triplets, as shown in Fig. 1; this decay is mediated through their higgsino components via the non-zero baryon-number-violating RPV coupling $\lambda''$ [9,10].

When $m_{\tilde{t}} - m_{\tilde{\chi}_1^0} < m_{\text{top}}$ (Fig. 1a), the $t \rightarrow t\tilde{\chi}_1^0\tilde{\chi}_1^0\tilde{\chi}_1^0\tilde{\chi}_1^0$ decay is kinematically forbidden and the top-squark branching ratio ($B$) to $b\tilde{\chi}_1^\pm$ is equal to unity; when $m_{\tilde{t}} - m_{\tilde{\chi}_1^0} \geq m_{\text{top}}$ the value of $B$ is taken to be 0.5. In the latter case, the rest of the decay rate is evenly divided between the two neutralino states: $\tilde{t} \rightarrow t\tilde{\chi}_1^0\tilde{\chi}_1^0\tilde{\chi}_1^0\tilde{\chi}_1^0$ (Fig. 1b). For the supersymmetric particle masses under consideration, the analysis considers only values of $\lambda'' \sim O(10^{-2} - 10^{-1})$ [11] which ensure prompt neutralino and chargino decays and omit more complex RPV decay patterns such as $\tilde{\chi}_1^\pm \rightarrow W^{\pm*}\tilde{\chi}_1^0\tilde{\chi}_1^0\tilde{\chi}_1^0\tilde{\chi}_1^0$ or $\tilde{\chi}_2^0 \rightarrow Z^*\tilde{\chi}_1^0\tilde{\chi}_1^0\tilde{\chi}_1^0\tilde{\chi}_1^0$ that could be substantial for very small values of $\lambda''$ [3].

Previous searches targeting RPV SUSY models of pair-produced top squarks decaying through the coupling $\lambda''$ have been carried out by the ATLAS and CMS collaborations. Those searches already exclude top-squark masses in the ranges 100 GeV $\leq m_{\tilde{t}} \leq 470$ GeV and 480 GeV $\leq m_{\tilde{t}} \leq 610$ GeV (ATLAS [12]), and 80 GeV $\leq m_{\tilde{t}} \leq 270$ GeV, 285 GeV $\leq m_{\tilde{t}} \leq 340$ GeV and 400 GeV $\leq m_{\tilde{t}} \leq 505$ GeV.

*e-mail: atlas.publications@cern.ch
(CMS [13]) in scenarios where the top squark is the LSP and decays directly via $t \rightarrow b\tilde{t}$. For the direct top-squark production and $\lambda_{323}$-mediated decays of higgsino LSP scenarios, ATLAS has excluded top-squark masses up to 1.10 TeV, depending on the higgsino mass considered, in the region where $m_{\tilde{t}} - m_{\tilde{\chi}^0_{1\pm2}} \geq m_{\text{top}}$, by analysing lepton plus jets events [11]. CMS has excluded top-squark masses between 100 and 720 GeV for top-squark decays into four quarks in boosted topologies and with the mass of the higgsinos set to 75% of the squark mass [14].

This analysis considers events with six or more jets, of which at least four are identified as $b$-jets ($b$-tagged). There must be no identified electron or muon, and no requirement is made on the missing transverse momentum. In this channel, the dominant background is the non-resonant production of multijet events, referred to as ‘multijet’ in the following, and a data-driven method is applied to estimate its yield. Other backgrounds arise from top-quark pair production accompanied by extra $b$-jets or by a $Z$ or Higgs boson decaying into a $b$-quark pair. Results are reported as 95% confidence level (CL) exclusion limits on the top-squark production cross sections depending on the higgsino mass considered, in the region $|\eta| < 1.7$. The endcap and forward regions are instrumented with LAr calorimeters for both the electromagnetic and hadronic energy measurements up to $|\eta| = 4.9$. The muon spectrometer surrounds the calorimeters and is based on three large air-core toroidal superconducting magnets with eight coils each. The field integral of the toroids ranges between 2.0 and 6.0 T m across most of the detector. The muon spectrometer includes a system of precision tracking chambers and fast detectors for triggering. A two-level trigger system is used to select events to be recorded. The first-level trigger is implemented in hardware and uses a subset of the detector information to accept events at a rate of at most 100 kHz. This is followed by a software-based high-level trigger (HLT) that reduces the accepted event rate to $\sim 1.2$ kHz, on average.

### 2 ATLAS detector

The ATLAS experiment [15] at the LHC is a multipurpose particle detector with a forward–backward symmetric cylindrical geometry and a near $4\pi$ coverage in solid angle.¹ It consists of an inner tracking detector (ID) surrounded by a thin superconducting solenoid providing a 2 T axial magnetic field, electromagnetic and hadron calorimeters, and a muon spectrometer (MS). The inner tracking detector covers the pseudorapidity range $|\eta| < 2.5$. It consists of silicon pixel, silicon microstrip, and transition radiation tracking detectors. An additional innermost layer of the silicon pixel tracker, the insertable B-layer [16, 17], was installed in 2014 at an average radial distance of 3.3 cm from the beam-line to improve track reconstruction and flavour identification of quark-initiated jets. Lead/liquid-argon (LAr) sampling calorimeters provide electromagnetic energy measurements with high granularity. A steel/scintillator-tile calorimeter provides hadronic energy measurements and covers the central pseudorapidity range ($|\eta| < 1.7$). The endcap and forward regions are instrumented with LAr calorimeters for both the electromagnetic and hadronic energy measurements.

### 3 Data collection and simulated event samples

This search is based on 139 fb$^{-1}$ of centre-of-mass energy $\sqrt{s} = 13$ TeV $pp$ collision data, collected between 2015 and 2018, that satisfy beam, detector and data-quality criteria. The uncertainty in the combined 2015–2018 integrated luminosity is 1.7% [18], obtained using the LUCID-2 detector [19] for the primary luminosity measurements. The average number of interactions ($\langle \mu \rangle$ in the same and nearby bunch crossings (pile-up) varies from $\langle \mu \rangle = 13.4$ (2015 dataset) to $\langle \mu \rangle = 36.1$ (2018 dataset), with a highest $\langle \mu \rangle = 37.8$ (2017 dataset) and an average $\langle \mu \rangle = 33.7$. Data were collected using a four-jet trigger which, in the HLT, requires four jets each having $|\eta| < 2.5$, with $p_T > 100$ GeV for the 2015–2016 data period and $p_T > 120$ GeV for the 2017–2018 data period. Data events used for the validation of the data-driven multijet background were collected using the lowest unprescaled single-lepton triggers; the lowest trigger $p_T$ threshold used for muons is 20 (26) GeV in 2015 (2016–2018), while for electrons the trigger $p_T$ threshold is 24 (26) GeV in 2015–2017 (2018).

Monte Carlo (MC) simulations are used to model the SUSY signals, as well as to aid in the description of the background processes. In the remainder of this section, the simulation of the signal and of the main background processes contributing to the selected events in data is described. For all the simulated physics processes, the top-quark mass is assumed to be $m_{\text{top}} = 172.5$ GeV and the Higgs boson mass is taken to be $m_H = 125$ GeV. The generation of the simulated event samples includes the effect of multiple $pp$ interactions in the same and neighbouring bunch crossings, as well as the effect of pile-up on the detector response. These interactions were produced using PYTHIA 8.230 [20] with a set of tuned parameters called the A3 tune [21] and the NNPDF2.3

¹ 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 upwards. Cylindrical coordinates $(r, \phi)$ are used in the transverse plane, $\phi$ being the azimuthal angle around the $z$-axis. The pseudorapidity is defined in terms of the polar angle $\theta$ as $\eta = -\ln \tan(\theta/2)$. Angular distance is measured in units of $\Delta R \equiv \sqrt{(\Delta \eta)^2 + (\Delta \phi)^2}$.  

© Springer
Fig. 1 Diagrams of the signal processes involving pair production of top squarks $\tilde{t}$: (a) with the decay into a $b$-quark and the lightest chargino $\tilde{\chi}_1^+ (\tilde{t} \rightarrow b\tilde{\chi}_1^+)$ with the subsequent decay of the chargino $\tilde{\chi}_1^+ \rightarrow bb\tilde{s}$ and charge conjugate (c.c.), and (b) the decay into a top quark and the two lightest neutralinos $\tilde{\chi}_{1,2}^0$ with the subsequent decay $\tilde{\chi}_{1,2}^0 \rightarrow tbs$

leading-order (LO) [22] parton distribution function (PDF) set.

All generated MC samples were processed through a simulation [23] of the detector geometry and response using either GEANT4 [24] or a fast simulation [25] of the calorimeter response and were then processed by the same reconstruction software used on data. To model the parton shower, hadronisation, and underlying event, the PYTHIA 8 generator was used with the NNPDF2.3 LO PDF set and the A14 [26] set of tunable parameters. The decays of bottom and charm hadrons were modelled using EvtGen [27]. Simulated MC events are weighted such that the object identification efficiencies, energy scales and energy resolutions match those determined from data control samples [28,29].

MC samples for multijet production were generated using PYTHIA 8.230 with leading-order matrix elements for dijet production and a $p_T$-ordered parton shower. EvtGen v1.6.0 was used for bottom and charm hadron decays. The renormalisation and factorisation scales were set to the geometric mean of the squared transverse masses of the two outgoing partons, $\sqrt{(p_{T,1}^2 + m_1^2)(p_{T,2}^2 + m_2^2)}$.

The production of $t\bar{t}$ events (referred to as $t\bar{t}$ + jets) was modelled using the POWHEG-BOX v2 [30–33] generator at next-to-leading order (NLO) with the NNPDF3.0 NLO [34] PDF set and with the $h\text{damp}$ parameter set to 1.5 $m_{\text{top}}$ [35]. PYTHIA 8.230 was used for the parton shower and EvtGen v1.6.0 for bottom and charm hadron decays. The $t\bar{t}$ + jets sample was generated inclusively in the number of jets using fast simulation. The MC sample cross-section is corrected to the theory prediction at next-to-next-to-leading order (NNLO) in QCD including resummation of next-to-next-to-leading logarithmic (NNLL) soft gluon terms by means of the \textsc{Top++} (v2.0) program [36–42]. The generated events may have jets which do not originate from the decay of the $t\bar{t}$ system. These additional jets are used to categorise the events depending on the flavour of the matching parton. Particle jets are reconstructed from all stable particles generated in the event (excluding muons and neutrinos) using the anti-$k_T$ algorithm [43] with a radius parameter $R = 0.4$ and are required to have $p_T > 15$ GeV and $|\eta| < 2.5$. Events having at least one such particle jet, matched within $\Delta R < 0.3$ to a generated $b$-hadron having $p_T > 5$ GeV and not originating from a top-quark decay, are labelled as $t\bar{t} + \geq 1 b$ events. Similarly, events which are not already categorised as $t\bar{t} + \geq 1 b$, and where at least one particle jet is matched to a $c$-hadron not originating from a $W$ boson decay, are labelled as $t\bar{t} + \geq 1 c$ events. Events labelled as either $t\bar{t} + \geq 1 b$ or $t\bar{t} + \geq 1 c$ are referred to as $t\bar{t}$ + HF events (HF for ‘heavy flavour’). The remaining events, including those with no additional jets, are labelled as $t\bar{t}$ + light events (light for ‘light flavour’).

The $Wt$ single-top-quark background was generated at NLO in QCD by POWHEG-BOX v2 with the NNPDF3.0 NLO PDF set. Overlaps between the $t\bar{t}$ and $Wt$ final states were removed using the ‘diagram removal’ scheme [44]. PYTHIA 8.230 was used for the parton shower and EvtGen v1.6.0 for bottom and charm hadron decays. Samples of single-top events are normalised to the cross-section calculated at NLO in QCD with NNLL soft gluon corrections [45,46].
The production of $t\bar{t}V$ events was modelled using the MadGraph5_aMC@NLO v2.3.3 [47] generator at NLO with the NNPDF3.0 NLO PDF set. Pythia 8.210 was used for the parton shower and EvtGen v1.2.0 for bottom and charm hadron decays.

The production of $t\bar{t}H$ events was modelled using the POWHEG-BOX v2 generator to NLO with the NNPDF3.0 NLO PDF set. Pythia 8.230 was used for the parton shower and EvtGen v1.6.0 for bottom and charm hadron decays. The cross-sections are calculated at NLO QCD and NLO electroweak accuracy using the generator MadGraph5_aMC@NLO [48].

Signal events were produced using the MadGraph5_aMC@NLO v2.3.3 generator at NLO with the NNPDF2.3 LO PDF, and the fast simulation of the detector response. Pythia 8.230 was used for the parton shower and EvtGen v1.6.0 for bottom and charm hadron decays. Signal cross-section calculations include approximate next-to-next-to-leading-order (NNLO\textsubscript{Approx}) supersymmetric QCD corrections and the resummation of soft gluon emission at NNLL accuracy [49]. The nominal cross-section and its uncertainty are taken from an envelope of predictions using different PDF sets as well as different factorisation and renormalisation scales. Top-squark masses between 600 GeV and 1 TeV PDF sets as well as different factorisation and renormalisation scales. Top-squark masses between 600 GeV and 1 TeV are considered.

4 Event reconstruction

Events are required to have a primary vertex reconstructed from at least two tracks with transverse momentum $p_T > 500$ MeV. When several vertices are found in a given bunch crossing, the vertex with the largest summed $p_T^2$ of the associated tracks is selected as the primary vertex.

Electrons are reconstructed from energy deposits (clusters) in the electromagnetic calorimeter matched to tracks reconstructed in the ID [50,51] and are required to have $p_T > 10$ GeV and $|\eta| < 2.47$. Candidates in the calorimeter barrel–endcap transition region ($1.37 < |\eta| < 1.52$) are excluded. Electron tracks must match the primary vertex of the event: the longitudinal impact parameter is required to satisfy $|z_0| < 0.5$ mm, while the transverse impact parameter is required to satisfy $|d_0|/\sigma_{d_0} < 5$, where $\sigma_{d_0}$ represents the uncertainty in the measured $|d_0|$ values. Loose electrons are identified using the ‘Medium’ identification criterion provided by a likelihood-based discriminant [52]. Tight electrons are required to pass the ‘TightLH’ selection [52] and the ‘Gradient’ isolation criteria [52] and $p_T > 27$ GeV.

Muons are reconstructed by matching either track segments or full tracks in the MS to tracks in the ID [53]. Combined tracks are then re-fitted using information from both detector systems. Muon tracks must match the primary vertex of the event: the longitudinal impact parameter is required to satisfy $|z_0| < 0.5$ mm, while the transverse impact parameter is required to satisfy $|d_0|/\sigma_{d_0} < 3$. Loose muons are those that pass the ‘Loose’ muon selection [53] and have $p_T > 10$ GeV and $|\eta| < 2.5$, and Tight muons are those that pass the ‘Medium’ muon selection [53], satisfy the ‘FixedCutTightTrackOnly’ isolation criterion [53], and have $p_T > 27$ GeV.

Jets are reconstructed from three-dimensional topological energy clusters [54] in the calorimeter using the anti-$k_T$ jet algorithm [43] with a radius parameter of 0.4. Reconstructed jets are then corrected to the particle level by the application of a jet energy scale calibration that is derived from simulation and by in situ corrections obtained from 13 TeV data [55]. Jets used in this analysis are required to have $p_T > 25$ GeV and $|\eta| < 2.5$ after calibration.

To avoid selecting jets from pile-up, low-$p_T$ ($p_T < 120$ GeV) jets in the central ($|\eta| < 2.5$) region of the detector are required to satisfy the jet-vertex tagger [56] configured such that it has an efficiency of approximately 92% to identify jets from a primary vertex. This requirement is applied to both data and simulation. Quality criteria are imposed to identify jets arising from non-collision sources or detector noise (using the BadLoose operating point [57]), and any event containing at least one such jet is removed. This removal produces a negligible loss of efficiency for signal events.

The $b$-jets are identified via a $b$-tagging algorithm that uses multivariate techniques to combine information from the impact parameters of displaced tracks as well as topological properties of secondary and tertiary decay vertices reconstructed within the jet. This analysis uses the MV2c10 tagger [58], trained on a hybrid sample of simulated $t\bar{t}$ and $Z'$ events statistically enriched at high-$p_T$ in order to discriminate $b$-jets from a background consisting of light- (93%) and $c$-labelled (7%) jets [29]. A weight is calculated corresponding to the probable presence of a $b$-quark or a $c$-quark, and jets are confirmed $b$-tagged if they satisfy a minimum requirement on the MV2c10 $b$-tagging weight corresponding to an average efficiency in $t\bar{t}$ events of 60% for $b$-jets, 4% for $c$-jets and a rejection factor of approximately 1200 for light-flavour jets across the jet $p_T$ range.

An overlap removal procedure is carried out to resolve ambiguities between jets and lepton candidates. To prevent treating electron energy deposits as jets, the closest jet within $\Delta R = \sqrt{(\Delta \eta)^2 + (\Delta \phi)^2} = 0.2$ of a selected electron is
removed.\(^4\) If the nearest jet surviving that selection is within \(\Delta R_y = 0.4\) of the electron, the electron is discarded. To reduce the background from heavy-flavour decays inside jets, muons are removed if they are separated from the nearest jet by \(\Delta R_y < 0.4\). However, if that jet has fewer than three associated tracks, the muon is kept and the jet is removed instead.

5 Analysis strategy

Events selected for further analysis are required to have at least five jets, of which at least two must be \(b\)-tagged. The four highest-\(p_T\) jets are required to be on the trigger efficiency plateau, namely to have \(p_T > 120\) GeV or \(p_T > 140\) GeV, depending on the jet-\(p_T\) trigger requirement in 2015–2016 or 2017–2018, and have \(|\eta| < 2.5\). All other jets present in the event are required to have \(p_T > 25\) GeV and \(|\eta| < 2.5\). A lepton veto is applied: events that contain loose muons or electrons with \(p_T > 10\) GeV, whether isolated or non-isolated, are discarded.

After the selections described above, the largest background contribution to the measurement is from non-resonant multijet production from light-quark and gluonic final states. The next largest is from \(t\bar{t}+\)jets production. Other small background contributions originate from the production of a single top quark and from the production of a \(t\bar{t}\) pair in association with either a vector boson or a Higgs boson. The estimation of the multijet background using a data-driven method and the validation of this estimate without significant bias from potential signal contamination are the main challenges for this analysis.

To probe top-squark pair production and estimate the contribution of signal top squarks in data, a model-dependent fit of the yield of events with jet multiplicity \(N_j = 6, 7, 8\) and \(\geq 9\) and \(b\)-tagged jet multiplicity \(N_b = 4\) and \(\geq 5\) is performed. These \((N_j, N_b)\) regions are indicated as SR\(_j\) in Table 1. The signal contribution predicted for different values of \(m_{\tilde{t}}\) and \(m_{\tilde{\chi}_{1}^{0,\pm}}\) is considered in all bins and is scaled by one common signal-strength parameter \((\mu_{\tilde{t}\tilde{\chi}})\). For the model considered here, the product of acceptance and reconstruction efficiency \((A \times \epsilon)\) is of order \(5 \times 10^{-2}\) for \(N_j \geq 9\) and \(N_b \geq 5\). Figure 2 shows the number of signal events obtained from the model as a function of \(N_j\) and \(N_b\) compared to the estimated backgrounds. Their evaluation is described in Sect. 6. The signal yields are concentrated at high jet and \(b\)-tagged jet multiplicity, while the backgrounds are concentrated at low \(b\)-tagged jet multiplicity. To validate the background estimates, intervals with \(N_j = 6, 7, 8\) and \(\geq 9\), and \(N_b = 3\) and 4, subsequently referred to as VR-MJ, are used. In these, a region-dependent selection is applied, based on a maximum accepted value of the centrality mass \(C_\text{mass}\), defined as:

\[
C_\text{mass} = \frac{H_T}{\sqrt{(\sum_{i=1}^{N_j} E_i)^2 - (\sum_{i=1}^{N_j} p_T)^2}},
\]

i.e. the ratio of the scalar sum of all jet \(p_T\) in the event \((H_T)\) to the invariant mass of the set of observed jets. The signal-to-background ratio decreases monotonically with decreasing \(C_\text{mass}\) for all \(N_j\) and \(N_b\) values. The value of the maximum value of \(C_\text{mass}\) \((C_\text{mass}^\text{max})\) is chosen such that the signal-to-background ratio is less than 5%. Values of the \(C_\text{mass}^\text{max}\) limits used are listed in Table 1.

A separate, model-independent test is used to search for, and to set generic exclusion limits on, potential contributions from a hypothetical BSM signal by comparing the observed number of events with background predictions in two dedicated signal regions, one with \(N_j \geq 9\) and \(N_b \geq 5\) and the other with \(N_j \geq 8\) and \(N_b \geq 5\) (labelled SR\(_{\text{discovery}}\) in Table 1), that were not explored in previous searches at the LHC.

6 Multijet background estimation

The predominant multijet background is estimated via a data-driven method, subsequently referred to as the tag-rate function method for multijet events (TRF\(_{\text{MJ}}\)) [59,60]. The aim is to extrapolate the \(b\)-tag multiplicity distributions from \(N_j = 5\), where the signal contamination for models not already excluded by other LHC searches is negligible, to larger \(N_j\) values. The TRF\(_{\text{MJ}}\) method uses a tag-rate function to quantify the experimental probability of \(b\)-tagging an additional jet in samples of events with at least two, or at least three, \(b\)-tagged jets. This per-jet probability is then used to estimate the shape of the multijet \(b\)-tag multiplicity distribution for each \(N_j\) value.

Events that satisfy the selection criteria described in Sect. 5 and that have exactly five jets, of which at least two are \(b\)-tagged, are used to determine the \(b\)-tagging probability. The data are first corrected by subtracting the expected non-multijet background found in simulation, approximately 5% of the total. After excluding the two jets in each event with the highest \(b\)-tagging weight, the probability that each remaining jet is \(b\)-tagged, denoted \(\epsilon_3\), is calculated for this jet. A similar procedure is used to calculate the probability \(\epsilon_2\) of additional \(b\)-tagged jets in events with at least three \(b\)-tagged jets. These \(\epsilon\) probabilities are parameterised as a function of both the \(p_T\) of the remaining jet divided by \(H_T\), and the minimum \(\Delta R\) between that jet and the two (for \(\epsilon_2\)) or three (for \(\epsilon_3\)) jets with the largest \(b\)-tagging weight in the event \((\Delta R_{\text{min}})\). This choice of variables for the parameterisation is made to minimise the residual differences (non-closure) between the
Table 1 The strategy of the analysis. For the model-dependent fit, the signal regions (SR\textsubscript{t}) consist of events with \(N_j = 6, 7, 8\) and \(\geq 9\) jets and \(N_b = 4\) and \(\geq 5\). These are used independently in the final fit. For the model-independent fit, two dedicated signal regions (SR\textsubscript{discovery}), with \((N_j \geq 9, N_b \geq 5)\) and \((N_j \geq 8, N_b \geq 5)\), are used. The validation regions (VR-MJ), which are based on a maximum value of the centrality mass, \(C_{\text{max}}^\text{VR-MJ}\), introduced for the description of the VRs in Sect. 5, are also indicated.

<table>
<thead>
<tr>
<th>(N_j)</th>
<th>(N_b)</th>
<th>VR-MJ (C_{\text{max}}^\text{VR-MJ})</th>
<th>(N_b)</th>
<th>(N_b)</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>3</td>
<td>(C_{\text{max}}^\text{VR-MJ} = 1.2)</td>
<td>(N_b)</td>
<td>(N_b)</td>
</tr>
<tr>
<td>7</td>
<td>(\geq 5)</td>
<td>(C_{\text{max}}^\text{VR-MJ} = 0.9)</td>
<td>(N_b)</td>
<td>(N_b)</td>
</tr>
<tr>
<td>8</td>
<td>(\geq 5)</td>
<td>(C_{\text{max}}^\text{VR-MJ} = 0.7)</td>
<td>(N_b)</td>
<td>(N_b)</td>
</tr>
<tr>
<td>(\geq 9)</td>
<td>(\geq 5)</td>
<td>(C_{\text{max}}^\text{VR-MJ} = 0.5)</td>
<td>(N_b)</td>
<td>(N_b)</td>
</tr>
</tbody>
</table>

TRF\textsubscript{MJ} prediction and the number of events obtained when selecting \(b\)-jets directly in the most sensitive signal regions in the multijet events simulated by MC. The dependence of \(\varepsilon_2\) and \(\varepsilon_3\) on both \(p_T/H\text{\scriptsize T}\) and \(\Delta R_{\text{min}}\) is shown in Fig. 3. The rapid variation with \(\Delta R_{\text{min}}\) is consistent with the dependence expected from multi-\(b\)-jet production due to gluon-splitting. The \(p_T/H\text{\scriptsize T}\) dependence, more visible at small \(\Delta R_{\text{min}}\), reflects the variation of the \(b\)-tagging efficiency with jet \(p_T\).

Following the methods of Ref. [61], in the second step of the TRF\textsubscript{MJ} method the expected number of events with each different number of \(b\)-tagged jets is estimated for each \(N_j\) value by weighting all events with \(N_b \geq 2\) by the \textit{event probability} of having \(N_b = 2, 3, 4\) and \(\geq 5\), respectively. Upon subtracting the non-multijet background contribution [59], the event probabilities are estimated using both \(\varepsilon_2\) and \(\varepsilon_3\), after first excluding the two jets with the highest \(b\)-tagging weight. For \(N_b = 2\) the event probabilities are estimated directly from \(\varepsilon_2\), treating the tagging probability for each jet as independent. For \(N_b = 3, 4\) and \(\geq 5\), a two-step procedure is employed. First, a ‘pseudodata sample’ with \(N_b \geq 3\) is emulated, using \(\varepsilon_2\) in events with \(N_b \geq 2\). The additional emulated \(b\)-tagged jet is chosen randomly from the remaining \(N_j - 2\) jets by using their probability-dependent \(b\)-tagging weights [60]. This emulated sample is then used to estimate the event probabilities, this time relying on \(\varepsilon_3\). The probability of finding \(N_b = 4\) and \(N_b \geq 5\) is estimated using the emulated \(N_b \geq 3\) sample via \(\varepsilon_3\). Due to too few events in the control sample from which the \(\varepsilon_2\) and \(\varepsilon_3\) values are extracted, it is not possible to estimate the probability of \(b\)-tagging an additional jet in a sample of events with at least four \(b\)-tagged jets.

![Fig. 2 Predicted numbers of events as a function of jet multiplicity, \(N_j\), and \(b\)-tagged jet multiplicity, \(N_b\), for a SM background (multijet and top-quark production) and b top-squark pair production in the \(t \to b\tilde{\chi}_1^\pm (\tilde{\chi}_1^\pm \to b\bar{b})\) (and c.c.) channel, for \(m_{\tilde{t}} = 1000\) GeV and \(m_{	ilde{\chi}_1^\pm} = 950\) GeV](image)

6.1 Validation of TRF\textsubscript{MJ} method

The TRF\textsubscript{MJ} method is validated using two different comparisons with data: in the VR-MJ regions defined in Sect. 5, and in a separate set of \(Z + \) jets-enriched events. Figure 4 shows...
Fig. 3 Two-dimensional distributions of the probability $a \varepsilon_2$ or $b \varepsilon_3$ of $b$-tagging an additional jet in a sample of events with $a$ at least two or $b$ at least three $b$-tagged jets as a function of the ratio of jet transverse momentum to $H_T$, $p_T/H_T$, and the minimum $\Delta R$ between the jet and the $a$ two or $b$ three $b$-tagged jets with the highest $b$-tagging weight in the event, $\Delta R_{\text{min}}$. The choice of binning is made so as to avoid empty bins.

A comparison between measured and estimated event rates in VR-MJ. The data and predictions are in agreement within systematic uncertainties (described in Sect. 7).

An independent test of the method is performed in $Z + \text{jets}$-enriched events, referred as ‘VR-ZJ’, where additional jets are produced by radiation and where $b \bar{b}$ pairs arise from gluon splitting. In order to select events where a $Z$ boson decays into pairs of electrons or muons, events are required to pass a single-lepton trigger. Two opposite-sign, same-flavour, tight electrons or muons are required to each have $p_T > 27 \text{ GeV}$ and a pair mass larger than $60 \text{ GeV}$. Events are required to have at least five jets with $p_T > 25 \text{ GeV}$ and $|\eta| < 2.5$, of which at least two must be $b$-tagged. The tagging probabilities $\varepsilon_2$ and $\varepsilon_3$ are derived from five-jet VR-ZJ events and used to predict the number of events with $N_j = 6, 7, 8, \geq 9$ and $N_b = 4, \geq 5$. As shown in Fig. 5, this statistically limited test further validates the TRFMJ method.

7 Systematic uncertainties

Several sources of systematic uncertainty are considered that can affect the overall normalisation of signal and background samples and their relative contribution for different values of $N_j$ and $N_b$. In estimating the dominant multijet background from the data, systematic uncertainties arise from the assumptions made in obtaining the TRFMJ background estimates. Uncertainties related to the theoretical modelling and due to the description of the detector response in simulated events are relevant only for the signal and background MC samples.

The main assumption of the TRFMJ method is that it is possible to define per-jet $b$-tagging probabilities ($\varepsilon_2$ and $\varepsilon_3$) in events with at least two or at least three $b$-tagged jets and, in particular, that the variables used for the parameterisation are sensitive to the heavy-flavour composition of the jet sample. A second assumption is that the per-jet probabilities are independent of the jet multiplicity and, therefore, may be derived in a specific region, namely that with exactly five jets, and applied to regions with $N_j = 6, 7, 8$ and $\geq 9$ jets.

The validity of these assumptions is verified using MC simulations. The TRFMJ method is applied to Pythia 8 MC dijet events, and the larger of (a) the residual non-closure and (b) the statistical uncertainty in the number of events with a given $b$-tagged jet multiplicity, is symmetrised and taken to be the systematic uncertainty associated with the method. Table 2 shows the final TRFMJ systematic uncertainty in the multi-jet background estimation in each $(N_j, N_b)$ region. For $N_b = 4$ the TRFMJ uncertainties are dominated by the non-closure component, while for $N_b \geq 5$, the statistical component dominates. The TRFMJ uncertainties are the source of the largest systematic uncertainty for the analysis.

The second largest contribution to the total systematic uncertainty arises from the modelling of the $t \bar{t}$+$\text{jets}$ background. The diagrams that contribute to $t \bar{t} + \geq 1b, t \bar{t} + \geq 1c$, and $t \bar{t}$+light production are different, and the associated uncertainties may affect these processes differently in different regions. As a result, all uncertainties in $t \bar{t}$+jets background modelling, except the uncertainty in the inclusive cross-section, are considered to be uncorrelated among $t \bar{t} + \geq 1b, t \bar{t} + \geq 1c$, and $t \bar{t}$+light.
Fig. 4 Comparison between data and the predicted number of events with $N_j = 6, 7, 8$ and $\geq 9$ and $N_b = 3$ and $4$ in the VR-MJ validation regions, which are based on a maximum value of the centrality mass, $c_{\text{mass}}$. The bottom panel displays the ratios of data to the total prediction, uncertainty bars are statistical only. The systematic uncertainties listed in Sect. 7 are represented by the blue hatched area.

Fig. 5 Comparison between data and the number of events with $N_j = 6, 7, 8$ and $\geq 9$ and $N_b = 4$ and $\geq 5$ predicted by the TRF$_{\text{MJ}}$ method (grey histogram) in the VR-ZJ region, defined by the requirement of two isolated leptons with invariant mass larger than 60 GeV. The bottom panel displays the ratios of data to the TRF$_{\text{MJ}}$ prediction, uncertainty bars are statistical only. Systematic uncertainties in the TRF$_{\text{MJ}}$ prediction are represented by the blue hatched area.

The uncertainty in the inclusive $t\bar{t}$ NNLO+NNLL production cross-section is taken to be ±6% [42]. This uncertainty includes effects from varying the factorisation and renormalisation scales, the PDF, $\alpha_S$, and the top-quark mass. The normalisations of the $t\bar{t}+\geq 1c$ and $t\bar{t}+\geq 1b$ yields are taken from their fractional contribution to the nominal $t\bar{t}$+jets sample as generated using the POWHEG-BOX program. In addition to the uncertainty in the inclusive $t\bar{t}$ cross-section, an additional uncertainty of 50%, based on the measurement of the $t\bar{t}+\geq 1b$ and $t\bar{t}+\geq 1c$ normalisation factors reported in Ref. [62], is assigned to the $t\bar{t}+\geq 1c$ and $t\bar{t}+\geq 1b$ production cross-sections.

The impact of the parton shower and hadronisation model uncertainties on the $t\bar{t}$+jets, $t\bar{t}H$ and $Wt$ single-top-quark
Table 2 Systematic uncertainties in the data-driven estimation of the multijet background using the TRF MJ method. The uncertainties are assessed using PYTHIA 8 MC dijet events for each value of jet multiplicity \(N_j\) and \(b\)-tagged jet multiplicity \(N_b\) used in the final fit.

<table>
<thead>
<tr>
<th>TRF MJ uncertainty</th>
<th>(N_b)</th>
<th>(N_j)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>4</td>
<td>(\geq 5)</td>
</tr>
<tr>
<td>(N_j)</td>
<td>6</td>
<td>9%</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>9%</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>13%</td>
</tr>
<tr>
<td>(N_j \geq 9)</td>
<td></td>
<td>16%</td>
</tr>
</tbody>
</table>

yields is evaluated by comparing the sample from the nominal generator set-up with a sample produced with the NLO POWHEG-BOX v2 generator using the NNPDF3.0 NLO PDF set. The latter events are interfaced with HERWIG 7.04 [63,64], using the H7UE set of tuned parameters [64] and the MMHT2014LO PDF set [65], and processed using fast simulation of the detector response. The difference between the two predictions of the \(t\bar{t} + \geq 1b\) event yield ranges from 20\% (33\%) for \(N_j = 6\) and \(N_b = 4 (\geq 5)\) to 46\% (60\%) in the region with \(N_j \geq 9\) and \(N_b = 4 (\geq 5)\).

To assess the uncertainty due to the choice of matching scheme, the POWHEG-BOX sample is compared with a sample produced by MADGRAPH5_aMC@NLO and PYTHIA 8. For the calculation of the hard scattering, MADGRAPH5_aMC@NLO v2.6.0 with the NNPDF3.0 NLO PDF set is used. The events are processed with PYTHIA 8.230, using the A14 set of tuned parameters and the NNPDF2.3 LO PDF set, and the fast simulation of the detector response. The uncertainty, which is obtained from the difference in yield between the two models and is symmetrised, affects both the normalisation and the \(N_j\) and \(N_b\)-dependence of background rates. It is largest for large values of the jet and \(b\)-tagged jet multiplicities. For \(t\bar{t}+\geq 1b\), it reaches 25\% for \(N_j = 8, \geq 9\) and \(N_b = 4\), and 41\% (32\%) for \(N_j = 8 (\geq 9)\) and \(N_b \geq 5\).

The effect of renormalisation and factorisation scale uncertainties and PDF uncertainties is evaluated for \(t\bar{t}H\) and \(t\bar{t}V\) events. For the former, the scales are varied simultaneously by common factors of 2.0 and 0.5. For the latter, the envelope of the 100 variations for NNPDF3.0 NLO [34] are taken into account. An uncertainty of \(\pm 5\%\) is assigned to the total cross-section for single-top production [45,66,67]. For both the \(t\bar{t}H\) and single-top events, additional uncertainties due to initial- and final-state radiation and the choice of generator are evaluated in a manner similar to that used for \(t\bar{t}+\) jets. The uncertainty in the amount of interference between \(Wt\) and \(t\bar{t}\) production at NLO is assessed by comparing samples using the default ‘diagram removal’ scheme with those using an alternative ‘diagram subtraction’ scheme [44]. All modelling uncertainties from non-\(t\bar{t}\)+jets simulated backgrounds are, after investigation, found to be negligible.

The uncertainties assigned to the expected signal yield for the SUSY benchmark processes considered include the experimental uncertainties related to the luminosity and to the detector modelling, which are dominated by the modelling of the jet energy scale and the \(b\)-tagging efficiencies. For example, for the \(\tilde{t} \rightarrow b\tilde{\chi}_1^+ (\tilde{\chi}_1^+ \rightarrow \tilde{b}\tilde{b}\tilde{\nu}\text{c.c.)}\) signal model, the \(b\)-tagging uncertainties in the region \(N_j \geq 9\) and \(N_b = 4\) are approximately 10\%, and the jet-related uncertainties of the signal yields are in the range of 3–5\%. The uncertainties in the signal yields related to the modelling of additional jet radiation are studied by varying the factorisation, renormalisation, and jet-matching scales as well as the parton-shower tune in the simulation. The corresponding uncertainties are small for most of the signal parameter space and are largest for small top-squark masses, where they reach 7\%. The uncertainty in the signal cross-section ranges between 8\% and 11\% for a top-squark mass in the range 600–1000 GeV.

8 Results

The events are allocated to \((N_j, N_b)\) regions with different signal-to-background ratios in order to constrain systematic uncertainties and to improve the separation of signal and background. Then, in each region, the total signal and background yields, shown in Tables 3 and 4, are used in combination as the input for the statistical analysis to extract the final results.

Hypothesis testing is performed using a modified frequentist method as implemented in ROOSTATS [68] and is based on a profile likelihood which takes into account the systematic uncertainties as nuisance parameters. This procedure minimises the impact of systematic uncertainties on the search sensitivity by taking advantage of the highly populated, background-dominated \((N_j, N_b)\) regions included in the likelihood fit. The signal-strength parameter, \(\mu_{ij}\), defined for positive values and corresponding to the signal normalisation, is unconstrained in the profile-likelihood fit. The normalisation of each component of the background and \(\mu_{ij}\), are determined simultaneously from the fit to the data.

Individual sources of systematic uncertainty are taken as uncorrelated. Contributions from \(t\bar{t} + \geq 1b, t\bar{t} + \geq 1c, t\bar{t} + \text{light, } t\bar{t} + V, t\bar{t}H\) and single-top-quark backgrounds are constrained by the uncertainties of the respective theoretical calculations, the uncertainty in the luminosity (described in Sect. 3), and experimental data. The TRF MJ uncertainty is taken as uncorrelated across regions because of its large statistical component. In all cases, the profile-likelihood-ratio test is used to establish 95\% confidence intervals using the CLs [69] prescription. The likelihood is configured differently for the model-independent and model-dependent hypothesis tests.
Table 3 Event yields from background predictions and data in the regions with $N_j = 6, 7, 8$ or $\geq 9$ and $N_b = 4$. The quoted uncertainties are the sum in quadrature of the statistical and systematic uncertainties in the yields for all samples. The individual background uncertainties can be larger than the total uncertainty due to correlations between parameters.

<table>
<thead>
<tr>
<th>Process</th>
<th>($N_j$, $N_b$)</th>
<th>(6, 4)</th>
<th>(7, 4)</th>
<th>(8, 4)</th>
<th>($\geq 9$, 4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Multijet</td>
<td>1760 ± 170</td>
<td>1920 ± 180</td>
<td>1510 ± 210</td>
<td>1870 ± 350</td>
<td></td>
</tr>
<tr>
<td>$t\bar{t}$ + light</td>
<td>6 ± 4</td>
<td>8.0 ± 3.4</td>
<td>6 ± 4</td>
<td>8 ± 7</td>
<td></td>
</tr>
<tr>
<td>$t\bar{t}$ + $\geq 1c$</td>
<td>4.1 ± 2.9</td>
<td>8 ± 5</td>
<td>11 ± 6</td>
<td>22 ± 17</td>
<td></td>
</tr>
<tr>
<td>$t\bar{t}$ + $\geq 1b$</td>
<td>45 ± 26</td>
<td>110 ± 70</td>
<td>160 ± 100</td>
<td>350 ± 260</td>
<td></td>
</tr>
<tr>
<td>$t\bar{t}$ + $W$</td>
<td>0.055 ± 0.032</td>
<td>0.26 ± 0.07</td>
<td>0.30 ± 0.10</td>
<td>1.34 ± 0.28</td>
<td></td>
</tr>
<tr>
<td>$t\bar{t}$ + $Z$</td>
<td>1.8 ± 0.4</td>
<td>4.3 ± 1.0</td>
<td>6.0 ± 1.5</td>
<td>10.9 ± 2.3</td>
<td></td>
</tr>
<tr>
<td>$Wt$</td>
<td>1.7 ± 2.0</td>
<td>5 ± 5</td>
<td>5.1 ± 3.1</td>
<td>10 ± 11</td>
<td></td>
</tr>
<tr>
<td>$t\bar{t}H$</td>
<td>4.9 ± 0.9</td>
<td>10.5 ± 1.7</td>
<td>14.2 ± 2.4</td>
<td>29 ± 8</td>
<td></td>
</tr>
<tr>
<td>Total background</td>
<td>1820 ± 170</td>
<td>2060 ± 190</td>
<td>1710 ± 220</td>
<td>2300 ± 400</td>
<td></td>
</tr>
<tr>
<td>Data</td>
<td>1660</td>
<td>1901</td>
<td>1624</td>
<td>2237</td>
<td></td>
</tr>
</tbody>
</table>

Table 4 Event yields from background predictions and data in the regions with $N_j = 6, 7, 8$ or $\geq 9$ and $N_b \geq 5$. The quoted uncertainties are the sum in quadrature of the statistical and systematic uncertainties in the yields for all samples. The individual background uncertainties can be larger than the total uncertainty due to correlations between parameters.

<table>
<thead>
<tr>
<th>Process</th>
<th>($N_j$, $N_b$)</th>
<th>(6, $\geq 5$)</th>
<th>(7, $\geq 5$)</th>
<th>(8, $\geq 5$)</th>
<th>($\geq 9$, $\geq 5$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Multijet</td>
<td>49 ± 13</td>
<td>75 ± 23</td>
<td>74 ± 14</td>
<td>123 ± 20</td>
<td></td>
</tr>
<tr>
<td>$t\bar{t}$ + light</td>
<td>&lt;0.01</td>
<td>0.3 ± 0.6</td>
<td>&lt;0.01</td>
<td>0.00 ± 0.04</td>
<td></td>
</tr>
<tr>
<td>$t\bar{t}$ + $\geq 1c$</td>
<td>&lt;0.01</td>
<td>0.016 ± 0.029</td>
<td>0.3 ± 0.4</td>
<td>0.26 ± 0.31</td>
<td></td>
</tr>
<tr>
<td>$t\bar{t}$ + $\geq 1b$</td>
<td>1.2 ± 0.9</td>
<td>3.9 ± 2.7</td>
<td>7 ± 6</td>
<td>28 ± 25</td>
<td></td>
</tr>
<tr>
<td>$t\bar{t}$ + $W$</td>
<td>&lt;0.01</td>
<td>0.005 ± 0.007</td>
<td>0.021 ± 0.025</td>
<td>0.090 ± 0.035</td>
<td></td>
</tr>
<tr>
<td>$t\bar{t}$ + $Z$</td>
<td>0.05 ± 0.05</td>
<td>0.22 ± 0.12</td>
<td>0.7 ± 0.4</td>
<td>0.7 ± 0.7</td>
<td></td>
</tr>
<tr>
<td>$Wt$</td>
<td>&lt;0.01</td>
<td>&lt;0.01</td>
<td>0.00 ± 0.13</td>
<td>0.9 ± 1.2</td>
<td></td>
</tr>
<tr>
<td>$t\bar{t}H$</td>
<td>0.12 ± 0.05</td>
<td>0.49 ± 0.13</td>
<td>0.82 ± 0.21</td>
<td>2.9 ± 1.5</td>
<td></td>
</tr>
<tr>
<td>Total background</td>
<td>50 ± 13</td>
<td>80 ± 23</td>
<td>84 ± 15</td>
<td>156 ± 27</td>
<td></td>
</tr>
<tr>
<td>Data</td>
<td>35</td>
<td>75</td>
<td>80</td>
<td>179</td>
<td></td>
</tr>
</tbody>
</table>

For the model-independent test, a profile-likelihood fit is performed independently in the two SR$_{\text{discovery}}$ regions with ($N_j \geq 8$, $N_b = 5$) and ($N_j \geq 9$, $N_b \geq 5$). This test is used to search for, and to compute generic exclusion limits on, the potential contribution from a hypothetical BSM signal in the given SR$_{\text{discovery}}$ regions.

For the model-dependent test, assuming a specific top-squark model with variable mass values, tests of the signal-pluss-background hypothesis, i.e. $\mu_{t\tilde{t}}$, are formed for a series of values of $m_{\tilde{t}}$ and $m_{\tilde{t}^\pm}$. These are used to derive exclusion limits for the specific top-squark model. The full set of regions, $N_j = 6, 7, 8$ and $\geq 9$ and $N_b = 4$ and $\geq 5$, is employed in the likelihood. The expected signal contribution, as predicted by the given model, is considered in all regions and is scaled by $\mu_{t\tilde{t}}$.

Figure 6 shows the observed numbers of data events compared with the fitted background model. The likelihood fit is configured using the model-dependent set-up where all bins are input to the fit, and $\mu_{t\tilde{t}}$ is set to zero. This configuration is also referred to as the background-only fit and includes no free-floating parameters, only nuisance parameters with Gaussian constraints. An example signal model is also shown in the figure to illustrate the separation between the signal and the background.

8.1 Model-independent interpretation

The model-independent results are calculated from the observed number of events and the background predictions in the two SR$_{\text{discovery}}$ regions. The observed number of events and the backgrounds obtained from the fits are shown for both SR$_{\text{discovery}}$ regions in Table 5.
Fig. 6 Expected background and observed number of events in different jet and b-tag multiplicity bins. The background is estimated by including all bins in a background-only fit and is plotted separately for each contribution. An example signal yield for $t \rightarrow b \tilde{\chi}_{\pm}^1$ (and c.c.) production with $m_{\tilde{t}} = 600 \text{ GeV}$ and $m_{\tilde{\chi}_{\pm}^1} = 550 \text{ GeV}$ is overlaid. The bottom panel displays the ratios of data to the total prediction, uncertainty bars are statistical only. All uncertainties, which can be correlated across bins, are included in the error bands.

Table 5 Fitted background yields in ($N_j \geq 8$, $N_b \geq 5$) and ($N_j \geq 9$, $N_b \geq 5$) signal regions. The individual background uncertainties can be larger than the total uncertainty due to correlations between parameters.

<table>
<thead>
<tr>
<th>Process</th>
<th>$N_j \geq 8$, $N_b \geq 5$</th>
<th>$N_j \geq 9$, $N_b \geq 5$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Multijet</td>
<td>200 $\pm$ 40</td>
<td>123 $\pm$ 20</td>
</tr>
<tr>
<td>$t\bar{t} + \geq 1c$</td>
<td>0.6 $\pm$ 0.6</td>
<td>0.29 $\pm$ 0.33</td>
</tr>
<tr>
<td>$t\bar{t} + \geq 1b$</td>
<td>26 $\pm$ 20</td>
<td>20 $\pm$ 15</td>
</tr>
<tr>
<td>$t\bar{t} + W$</td>
<td>0.11 $\pm$ 0.05</td>
<td>0.09 $\pm$ 0.04</td>
</tr>
<tr>
<td>$t\bar{t} + Z$</td>
<td>1.4 $\pm$ 0.7</td>
<td>0.8 $\pm$ 0.7</td>
</tr>
<tr>
<td>$Wt$</td>
<td>0.9 $\pm$ 0.8</td>
<td>0.9 $\pm$ 1.2</td>
</tr>
<tr>
<td>$t\bar{t}H$</td>
<td>3.7 $\pm$ 1.6</td>
<td>2.9 $\pm$ 1.4</td>
</tr>
<tr>
<td>Total background</td>
<td>230 $\pm$ 40</td>
<td>147 $\pm$ 20</td>
</tr>
<tr>
<td>Data</td>
<td>259</td>
<td>179</td>
</tr>
</tbody>
</table>

Model-independent 95% CL upper limits on the expected and observed number of BSM events, $N_{\text{exp}}^{95}$ and $N_{\text{obs}}^{95}$, that may contribute to the signal regions are computed from the observed number of events and the fitted background. Normalising these results by the integrated luminosity, $L$, of the data sample, allows them to be interpreted as upper limits on the visible BSM cross-section $\sigma_{\text{bsm}}^{95}$, defined as:

$$\sigma_{\text{bsm}}^{95} = \sigma_{\text{prod}} \times A \times \epsilon = \frac{N_{\text{obs}}^{95}}{L},$$

where $\sigma_{\text{prod}}$ is the production cross-section. The resulting limits are presented in Table 6. In addition, the $p_0$ values, which quantify the probability that a background-only hypothesis results in a fluctuation giving an event yield equal to or larger than the one observed in the data, are calculated, as are the corresponding Gaussian significance values $Z$.

8.2 Model-dependent interpretation

For each signal model probed, the fit is configured using the model-dependent set-up, as detailed in the first part of Sect. 8. Figure 7 shows exclusion limits at the 95% confidence level in the top-squark production model when $B(t \rightarrow b \tilde{\chi}_{\pm}^1)$ is assumed to be unity. For this model, top-squark masses are excluded up to 950 GeV for chargino masses close to the kinematic threshold for producing this final state. For lower values of the chargino mass, the limit weakens such that for chargino masses of around 200 GeV, the top-squark mass is constrained to be more than 800 GeV. In this phase space region, the signal is concentrated at lower $N_j$ and $N_b$ values where the background is larger.

The limits for higgsino LSPs are shown in Fig. 8. In the region $m_{\tilde{t}} - m_{\tilde{\chi}_{\pm}^0} \geq m_{\text{top}}$, the sensitivity of the analysis is lower than in the pure $t \rightarrow b \tilde{\chi}_{\pm}^1$ case because contributions to the signal that have one leptonically decaying top quark fail the lepton-veto requirement. The large contribution of the multijet background reduces the present sensitivity relative to a previous ATLAS search that analysed events characterised by the presence of a lepton plus jets [11].

9 Conclusion

A search for physics beyond the Standard Model in events with high jet multiplicity and a large number of b-tagged jets is described in this paper. The search uses 139 fb$^{-1}$ of $\sqrt{s} = 13$ TeV proton–proton collision data collected by the ATLAS experiment at the LHC. In contrast to many previous
Fig. 7 Observed and expected exclusion contours for the $\tilde{t}$ and $\chi_1^{\pm}$ masses in a top-squark production model with RPV decays of the $\chi_1^{\pm}$. Limits are shown for $B(\tilde{t} \rightarrow b\chi_1^{\mp})$ equal to unity. The contours of the band around the expected limit are the $\pm 1\sigma$ variations, including all uncertainties except theoretical uncertainties in the signal cross-section. The dotted lines around the observed limit illustrate the change in the observed limit as the nominal signal cross-section is scaled up and down by the theoretical uncertainty. All limits are computed at 95% CL. The results are constrained by the kinematic limits of the top-squark decay into a chargino and a bottom quark (diagonal line).

Fig. 8 Observed and expected exclusion contours for the $\tilde{t}$ and $\chi_1^{\pm}$ masses in a top-squark production model with RPV decays of the $\chi_1^{\pm}$ and of the $\tilde{\chi}_1^0$. Limits are shown in the case of a higgsino LSP. The contours of the band around the expected limit are the $\pm 1\sigma$ variations, including all uncertainties except theoretical uncertainties in the signal cross-section. The dotted lines around the observed limit illustrate the change in the observed limit as the nominal signal cross-section is scaled up and down by the theoretical uncertainty. All limits are computed at 95% CL. The results are constrained by the kinematic limits of the top-squark decay into a chargino and a bottom quark (upper diagonal line) and into a neutralino and a top quark (lower diagonal line), respectively. Also shown are the limits from Ref. [11].
Table 6 Observed 95% CL model-independent upper limits on the visible BSM cross-section \( \sigma_{\text{obs}}^{95} \), expressed in fb, along with the observed (expected) limits, \( N_{\text{obs}}^{95} \) (\( N_{\text{exp}}^{95} \)), on the number of excess events. The limits are determined for two signal regions, \( (N_j \geq 8, N_b \geq 5) \) and \( (N_j \geq 9, N_b \geq 5) \). The \( \rho_0 \) value quantifies the probability that the background-only hypothesis would result in a fluctuation that gives an event yield equal to or larger than the one observed in the data, and \( Z \) is the corresponding Gaussian significance.

<table>
<thead>
<tr>
<th>Signal region</th>
<th>( \sigma_{\text{obs}}^{95} ) [fb]</th>
<th>( N_{\text{obs}}^{95} )</th>
<th>( N_{\text{exp}}^{95} )</th>
<th>( \rho_0 ) (Z)</th>
</tr>
</thead>
<tbody>
<tr>
<td>( N_j \geq 8, N_b \geq 5 )</td>
<td>0.76</td>
<td>105</td>
<td>85(^{+30}_{-24} )</td>
<td>0.24 (0.7)</td>
</tr>
<tr>
<td>( N_j \geq 9, N_b \geq 5 )</td>
<td>0.54</td>
<td>75</td>
<td>52(^{+20}_{-15} )</td>
<td>0.11 (1.2)</td>
</tr>
</tbody>
</table>

Acknowledgements

We would like to thank Gilbert Moultaka for helpful discussions on the phenomenological aspects of the analysis. 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; CNPq and FAPESP, Brazil; NSERC, NRC and CFI, Canada; CERN; ANID, Chile; CAS, MOST and NSFC, China; COLCIENCIAS, Colombia; MSM CR, MPO CR and VSC CR, Czech Republic; DNRF and DNSRC, Denmark; IN2P3-CNRS and CEA-DRF/IRFU, France; SRNSFG, Georgia; BMBF, HGF and MPG, Germany; GSRT, Greece; RGC and Hong Kong SAR, China; ISF and Benoziyo Center, Israel; INFN, Italy; MEXT and JSPS, Japan; CNRST, Morocco; NWO, Netherlands; RCN, Norway; MNiSW and NCN, Poland; FCT, Portugal; MINEPA, Romania; MES of Russia and NRC KI, Russia Federation; JINR; MEST, Serbia; MSSR, Slovakia; ARRS and MIZS, Slovenia; DST/NRF, 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, United States of America. In addition, individual groups and members have received support from BCKDF, CANARIE, Compute Canada and CRC, Canada; ERC, ERDF, Horizon 2020, Marie Skłodowska-Curie Actions and COST, European Union; Investissements d’Avenir Labex, Investissements d’Avenir Idex and ANR, France; DFG and AvH Foundation, Germany; Herakleitos, Thales and Aristeia programmes co-financed by EU-ESF and the Greek NSRF, Greece; BSF-NSF and GIF, Israel; La Caixa Banking Foundation, CERCA Programme Generalitat de Catalunya and PROMETEO and GenT Programmes Generalitat Valenciana, Spain; Gönül Gustafsson’s 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 (USA), the Tier-2 facilities worldwide and large non-WLCG resource providers. Major contributors of computing resources are listed in Ref. [70].

Data Availability Statement

This manuscript has no associated data or the data will not be deposited. [Authors’ comment: All ATLAS scientific output is published in journals, and preliminary results are made available in Conference Notes. All are openly available, without restriction on use by external parties beyond copyright law and the standard conditions agreed by CERN. Data associated with journal publications are also made available: tables and data from plots (e.g. cross section values, likelihood profiles, selection efficiencies, cross section limits, ... are stored in appropriate repositories such as HEPDATA (http://hepdata.cedar.ac.uk/). ATLAS also strives to make additional material related to the paper available that allows a reinterpretation of the data in the context of new theoretical models. For example, an extended encapsulation of the analysis is often provided for measurements in the framework of RIVET (http://rivet.hepforge.org/)." This information is taken from the ATLAS Data Access Policy, which is a public document that can be downloaded from http://opendata.cern.ch/record/413 [opendata.cern.ch].]

Open Access

This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

Funded by SCOAP3.

References


---

1 Department of Physics, University of Adelaide, Adelaide, Australia
2 Physics Department, SUNY Albany, Albany, NY, USA
3 Department of Physics, University of Alberta, Edmonton, AB, Canada
4 (a)Department of Physics, Ankara University, Ankara, Turkey; (b)Istanbul Aydin University, Application and Research Center for Advanced Studies, Istanbul, Turkey; (c)Division of Physics, TOBB University of Economics and Technology, Ankara, Turkey
5 LAPP, Université Grenoble Alpes, Université Savoie Mont Blanc, CNRS/IN2P3, Annecy, France
6 High Energy Physics Division, Argonne National Laboratory, Argonne, IL, USA
7 Department of Physics, University of Arizona, Tucson, AZ, USA
8 Department of Physics, University of Texas at Arlington, Arlington, TX, USA
9 Physics Department, National and Kapodistrian University of Athens, Athens, Greece
10 Physics Department, National Technical University of Athens, Zografou, Greece
11 Department of Physics, University of Texas at Austin, Austin, TX, USA

© Springer
12 (a) Faculty of Engineering and Natural Sciences, Bahcesehir University, Istanbul, Turkey; (b) Faculty of Engineering and Natural Sciences, Istanbul Bilgi University, Istanbul, Turkey; (c) Department of Physics, Bogazici University, Istanbul, Turkey; (d) Department of Physics Engineering, Gaziantep University, Gaziantep, Turkey
13 Institute of Physics, Azerbaijan Academy of Sciences, Baku, Azerbaijan
14 Institut de Física d’Altes Energies (IFAE), Barcelona Institute of Science and Technology, Barcelona, Spain
15 (a) Institute of High Energy Physics, Chinese Academy of Sciences, Beijing, China; (b) Physics Department, Tsinghua University, Beijing, China; (c) Department of Physics, Nanjing University, Nanjing, China; (d) University of Chinese Academy of Science (UCAS), Beijing, China
16 Institute of Physics, University of Belgrade, Belgrade, Serbia
17 Department for Physics and Technology, University of Bergen, Bergen, Norway
18 Physics Division, Lawrence Berkeley National Laboratory and University of California, Berkeley, CA, USA
19 Institut für Physik, Humboldt Universität zu Berlin, Berlin, Germany
20 Albert Einstein Center for Fundamental Physics and Laboratory for High Energy Physics, University of Bern, Bern, Switzerland
21 School of Physics and Astronomy, University of Birmingham, Birmingham, UK
22 (a) Facultad de Ciencias y Centro de Investigaciones, Universidad Antonio Nariño, Bogotá, Colombia; (b) Departamento de Física, Universidad Nacional de Colombia, Bogotá, Colombia
23 (a) Dipartimento di Fisica, INFN Bologna and Universita’ di Bologna, Bologna, Italy; (b) INFN Sezione di Bologna, Bologna, Italy
24 Physikalisches Institut, Universität Bonn, Bonn, Germany
25 Department of Physics, Boston University, Boston, MA, USA
26 Department of Physics, Brandeis University, Waltham, MA, USA
27 (a) Transilvania University of Brasov, Brasov, Romania; (b) Horia Hulubei National Institute of Physics and Nuclear Engineering, Bucharest, Romania; (c) Department of Physics, Alexandru Ioan Cuza University of Iasi, Iasi, Romania; (d) Physics Department, National Institute for Research and Development of Isotopic and Molecular Technologies, Cluj-Napoca, Romania; (e) University Politehnica Bucharest, Bucharest, Romania; (f) West University in Timisoara, Timisoara, Romania
28 (a) Faculty of Mathematics, Physics and Informatics, Comenius University, Bratislava, Slovak Republic; (b) Department of Subnuclear Physics, Institute of Experimental Physics of the Slovak Academy of Sciences, Kosice, Slovak Republic
29 Physics Department, Brookhaven National Laboratory, Upton, NY, USA
30 Departamento de Física, Universidad de Buenos Aires, Buenos Aires, Argentina
31 California State University, Long Beach, CA, USA
32 Cavendish Laboratory, University of Cambridge, Cambridge, UK
33 (a) Department of Physics, University of Cape Town, Cape Town, South Africa; (b) iThemba Labs, Western Cape, South Africa; (c) Department of Mechanical Engineering Science, University of Johannesburg, Johannesburg, South Africa; (d) Department of Physics, University of South Africa, Pretoria, South Africa; (e) School of Physics, University of the Witwatersrand, Johannesburg, South Africa
34 Department of Physics, Carleton University, Ottawa, ON, Canada
35 (a) Faculté des Sciences Ain Chock, Réseau Universitaire de Physique des Hautes Energies, Université Hassan II, Casablanca, Morocco; (b) Faculté des Sciences, Université Ibn-Tofail, Kénitra, Morocco; (c) Faculté des Sciences Semlalia, Université Cadi Ayyad, LPHEA, Marrakech, Morocco; (d) Faculté des Sciences, Université Mohamed Premier and LPTPM, Oujda, Morocco; (e) Faculté des sciences, Université Mohammed V, Rabat, Morocco
36 CERN, Geneva, Switzerland
37 Enrico Fermi Institute, University of Chicago, Chicago, IL, USA
38 LPC, Université Clermont Auvergne, CNRS/IN2P3, Clermont-Ferrand, France
39 Nevis Laboratory, Columbia University, Irvington, NY, USA
40 Niels Bohr Institute, University of Copenhagen, Copenhagen, Denmark
41 (a) Dipartimento di Fisica, Università della Calabria, Rende, Italy; (b) INFN Gruppo Collegato di Cosenza, Laboratori Nazionali di Frascati, Frascati, Italy
42 Physics Department, Southern Methodist University, Dallas, TX, USA
43 Physics Department, University of Texas at Dallas, Richardson, TX, USA
44 National Centre for Scientific Research “Demokritos”, Agia Paraskevi, Greece
45 (a) Department of Physics, Stockholm University, Stockholm, Sweden; (b) Oskar Klein Centre, Stockholm, Sweden
Department of Particle Physics and Astrophysics, Weizmann Institute of Science, Rehovot, Israel
Department of Physics, University of Wisconsin, Madison, WI, USA
Fakultät für Mathematik und Naturwissenschaften, Fachgruppe Physik, Bergische Universität Wuppertal, Wuppertal, Germany
Department of Physics, Yale University, New Haven, CT, USA

a Also at Borough of Manhattan Community College, City University of New York, New York, NY, USA
b Also at Centro Studi e Ricerche Enrico Fermi, Roma, Italy
c Also at CERN, Geneva, Switzerland
d Also at CPPM, Aix-Marseille Université, CNRS/IN2P3, Marseille, France
e Also at Département de Physique Nucléaire et Corpusculaire, Université de Genève, Geneva, Switzerland
f Also at Departament de Física de la Universitat Autònoma de Barcelona, Barcelona, Spain
g Also at Department of Financial and Management Engineering, University of the Aegean, Chios, Greece
h Also at Department of Physics and Astronomy, Michigan State University, East Lansing, MI, USA
i Also at Department of Physics and Astronomy, University of Louisville, Louisville, KY, USA
j Also at Department of Physics, Ben Gurion University of the Negev, Beer Sheva, Israel
k Also at Department of Physics, California State University, East Bay, USA
l Also at Department of Physics, California State University, Fresno, USA
m Also at Department of Physics, California State University, Sacramento, USA
n Also at Department of Physics, King’s College London, London, UK
o Also at Department of Physics, St. Petersburg State Polytechnical University, St. Petersburg, Russia
p Also at Department of Physics, University of Fribourg, Fribourg, Switzerland
q Also at Dipartimento di Matematica, Informatica e Fisica, Università di Udine, Udine, Italy
r Also at Faculty of Physics, M.V. Lomonosov Moscow State University, Moscow, Russia
s Also at Giresun University, Faculty of Engineering, Giresun, Turkey
t Also at Graduate School of Science, Osaka University, Osaka, Japan
u Also at Hellenic Open University, Patras, Greece
v Also at IJCLab, Université Paris-Saclay, CNRS/IN2P3, 91405, Orsay, France
w Also at Institut Catalana de Recerca i Estudis Avancats, ICREA, Barcelona, Spain
x Also at Institut für Experimentalphysik, Universität Hamburg, Hamburg, Germany
y Also at Institute for Mathematics, Astrophysics and Particle Physics, Radboud University Nijmegen/Nikhef, Nijmegen, The Netherlands
z Also at Institute for Nuclear Research and Nuclear Energy (INRNE) of the Bulgarian Academy of Sciences, Sofia, Bulgaria

aa Also at Institute for Particle and Nuclear Physics, Wigner Research Centre for Physics, Budapest, Hungary
ab Also at Institute of Particle Physics (IPP), Montreal, Canada
ac Also at Institute of Physics, Azerbaijan Academy of Sciences, Baku, Azerbaijan
ad Also at Instituto de Física Teórica, IFT-UAM/CSIC, Madrid, Spain
ae Also at Joint Institute for Nuclear Research, Dubna, Russia
af Also at Louisiana Tech University, Ruston, LA, USA
ag Also at Moscow Institute of Physics and Technology State University, Dolgoprudny, Russia
ah Also at National Research Nuclear University MEPhI, Moscow, Russia
ai Also at Physics Department, An-Najah National University, Nablus, Palestine
aj Also at Physikalisches Institut, Albert-Ludwigs-Universität Freiburg, Freiburg, Germany
ak Also at The City College of New York, New York, NY, USA
al Also at TRIUMF, Vancouver, BC, Canada
am Also at Università di Napoli Parthenope, Naples, Italy
an Also at University of Chinese Academy of Sciences (UCAS), Beijing, China
ao *Deceased