ATLAS NOTE

ATLAS-CONF-2020-016

9th June 2020

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

The ATLAS Collaboration

A search is presented for new phenomena in final state events characterized by high jet multiplicity, no isolated leptons (electrons and muons), and large multiplicity of 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 source of background, namely multijet production, is estimated using a data-driven technique based on the extrapolation from events with low $b$-jet multiplicity to the events with high $b$-jet multiplicity used in the search. No significant excess over the Standard Model expectation is observed and 95% confidence-level limits are extracted constraining simplified models of R-parity-violating supersymmetry that feature top-squark pair production. The exclusion limits reach up to 950 GeV in top-squark mass in the models considered.

ATLAS-CONF-2020-016
09 June 2020

# 1 Introduction

Events with a large number of high-$p_{\mathrm{T}}$ 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. An excess of events with such topology can be a signal of phenomena beyond the SM (BSM) [1–3]. Event topologies with five or more $b$-jets, small missing transverse momentum (the magnitude of which is denoted by $E_{\mathrm{T}}^{\mathrm{miss}}$) and no leptons (electrons and muons) have not been covered by present searches at the LHC.

Final states with high leptonic or hadronic multiplicity are commonly predicted by RPV supersymmetry (SUSY), a variant of SUSY [4, 5] in which R-parity is violated (RPV) [6]. RPV SUSY does not provide stable superpartners, and it gives rise to a wide variety of experimental signatures whose nature depends on which of the many RPV couplings are non-zero. The existence of light SUSY partners of third generation quarks, bottom- ($\tilde{b}$) and top- ($\tilde{t}$) squarks, is favoured by naturalness considerations [7, 8].

In the analysis presented here, a particular benchmark model is considered in order to interpret the measurement performed in the different jet and $b$-jet multiplicity regions. The process under consideration is the lightest top-squark pair production. The scenario assumes the lightest SUSY partners (LSP) to be a triplet of higgsino-like states ($\tilde{\chi}_2^0$, $\tilde{\chi}_1^{\pm}$, $\tilde{\chi}_1^0$) that are mass-degenerate. The top squark decays either to a chargino ($\chi_1^{\pm}$) and a bottom quark or into a neutralino ($\chi_{1,2}^0$) and a top quark, the chargino and neutralino decay through their higgsino components and the non-zero baryon number violating RPV coupling $\lambda''_{323} \approx O(10^{-2} - 10^{-1})$, to $bbs$ and $tbs$ quark triplets as shown in Figure 1. When $m_{\tilde{t}} - m_{\tilde{\chi}_{1,2}^0, \tilde{\chi}_1^{\pm}} \leq m_{\mathrm{top}}$, the $\tilde{t} \to t\tilde{\chi}_{1,2}^0$ decay is kinematically forbidden. In this case the top-squark branching ratio (BR) to $b\tilde{\chi}_1^{\pm}$ is equal to unity (Figure 1(a)), otherwise it is assumed to be 0.5 when $m_{\tilde{t}} - m_{\tilde{\chi}_{1,2}^0, \tilde{\chi}_1^{\pm}} \geq m_{\mathrm{top}}$. In the latter case, the rest of the BR is evenly divided between $\tilde{t} \to t\tilde{\chi}_{1,2}^0(\tilde{\chi}_{1,2}^0 \to tbs)$ (Figure 1(b)). For the supersymmetric particle masses under consideration, we consider only values of $\lambda''_{323}$ which ensure prompt neutralino decays and we omit more complex RPV decay patterns such as $\tilde{\chi}_1^{\pm} \to W^{\pm *}\tilde{\chi}_1^0(\tilde{\chi}_1^0 \to tbs)$ or $\tilde{\chi}_2^0 \to Z^*\tilde{\chi}_1^0(\tilde{\chi}_1^0 \to tbs)$ that could be substantial for very small values of $\lambda''_{323}$ [3].

Previous searches targeting RPV SUSY models of top-squark pair production decaying through the coupling $\lambda''_{323}$ have been carried out by the ATLAS and CMS collaborations. They 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 [9]), 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 (CMS [10]) in scenarios where the top squark is the LSP and decay directly via $\tilde{t} \to bs$. 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 $m_{\tilde{t}} - m_{\tilde{\chi}_{1,2}^0, \tilde{\chi}_1^{\pm}} \geq m_{\mathrm{top}}$ [11]. CMS has excluded top-squark masses between 100 and 720 GeV for the top squark decays into four quarks and for masses of the higgsino set to 75% of the squark mass [12].

A search for a BSM signal using this type of topology is presented in this note, which 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, and a data-driven method is applied to estimate its yields. Other sources of background arise from top-quark pair production accompanied by extra $b$-jets or by $Z$ or Higgs bosons decaying into a $b$-quark pair. Results are translated in 95% confidence level (CL) exclusion limits on the top-squark mass in the benchmark models described above. Model-independent limits on the possible contribution of BSM physics are also evaluated at large jet and $b$-tagged jet multiplicities.

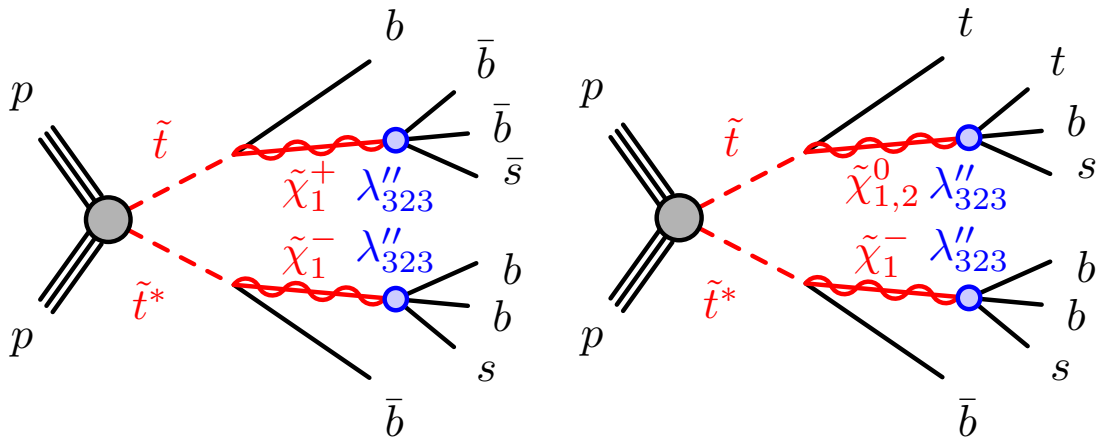



Figure 1: Diagrams of the considered signal processes involving pair production of top squarks $\tilde{t}$: (a) with the decay into a $b$-quark and the lightest chargino $\tilde{\chi}_1^\pm$ ($\tilde{t} \to b\tilde{\chi}_1^\pm$) with the subsequent decay of the chargino $\tilde{\chi}_1^\pm \to bbs$ and c.c. via a virtual top squark, and (b), for $m_{\tilde{t}} - m_{\tilde{\chi}_{1,2}^0, \tilde{\chi}_1^\pm} \geq m_{\text{top}}$ including the decay into a top-quark and the two lightest neutralinos $\tilde{\chi}_{1,2}^0$ with the subsequent decay $\tilde{\chi}_{1,2}^0 \to tbs$. Where not explicitly shown, the anti-squarks ($\tilde{t}^*$) decay into the charge-conjugate final states of those indicated.

## 2 ATLAS detector

The ATLAS experiment [13] at the LHC is a multi-purpose 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 $|\eta| < 2.5$. An additional innermost layer of the silicon pixel tracker, the insertable B-layer [14], was installed before 2015 at an average radial distance of 3.3 cm from the beam line to improve track reconstruction and flavour identification of quark-initiated jets. It consists of silicon pixel, silicon microstrip, and transition radiation tracking detectors. Lead/liquid-argon (LAr) sampling calorimeters provide electromagnetic energy measurements with high granularity. A hadron (steel/scintillator-tile) calorimeter covers the central pseudorapidity range ($|\eta| < 1.7$). The end-cap and forward regions are instrumented with LAr calorimeters for both EM 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. The first-level trigger is implemented in hardware and uses a subset of the detector information to reduce the accepted rate to at most nearly 100 kHz. This is followed by a software-based High Level Trigger (HLT) that reduces the accepted event rate to ~1.2 kHz, on average.

[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 $\Delta R \equiv \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$.

## 3 Data collection and simulated event samples

This search is based on 139 fb$^{-1}$ of $\sqrt{s}$ = 13 TeV $pp$ collision data, after applying beam, detector and data-quality criteria, collected between 2015 and 2018. In 2015–2016, the average number of interactions per bunch crossing (pile-up) was $\langle\mu\rangle$ = 20, increasing to $\langle\mu\rangle$ = 38 in 2017 and to $\langle\mu\rangle$ = 37 in 2018. The uncertainty in the combined 2015–2018 integrated luminosity is 1.7% [15], obtained using the LUCID-2 detector [16] for the primary luminosity measurements. Data are collected using a four-jet trigger which, in the HLT, requires four jets each having $|\eta| < 2.5$, $p_{\mathrm{T}} > 100$ GeV for the 2015-2016 data period and $p_{\mathrm{T}} > 120$ GeV for the 2017-2018 data period. For the validation of the data-driven multijet background, events are selected using the lowest unprescaled single-lepton triggers; the lowest trigger $p_{\mathrm{T}}$ threshold used for muons is 20 (26) GeV in 2015 (2016-2018), while for electrons the trigger $p_{\mathrm{T}}$ threshold is 24 (26) GeV.

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 main background processes contributing to the selected events in data and of the signal is described. For all the simulated physics processes, the top mass is assumed to be $m_{\mathrm{top}}$ = 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 per bunch crossing, as well as the effect on the detector response due to pile-up. These interactions were produced using Pythia 8.2 [17] with the A3 tuned parameters [18] and NNPDF2.3 [19] LO parton distribution functions (PDF) set. All simulated MC samples are processed through a simulation [20] of the detector geometry and response using either GEANT4 [21] or a fast simulation of the calorimeter response [22] and then reconstructed by the same reconstruction software as the data. Simulated MC events are weighted so that the object identification efficiencies, energy scales and energy resolutions match those determined from data control samples.

MC samples for multijet production are generated using Pythia 8.2 with leading-order matrix elements for dijet production and interfaced to a $p_{\mathrm{T}}$-ordered parton shower. The renormalisation and factorisation scales are set to the geometric mean of the squared transverse masses of the two outgoing partons, $\sqrt{(p_{\mathrm{T},1}^2 + m_1^2)(p_{\mathrm{T},2}^2 + m_2^2)}$. The `NNPDF2.3lo` PDF set [23] and the A14 [24] tune are used in the partonic final state generation, the shower, and in the simulation of the multi-parton interactions.

The production of $t\bar{t}$ events (referred to as $t\bar{t}$+jets) is modelled using the PowhegBox [25–28] v2 generator at NLO with the NNPDF3.0NLO [19] PDF set and with the $h_{\mathrm{damp}}$ parameter[2] set to 1.5 $m_{\mathrm{top}}$ [29]. The events are interfaced to Pythia 8.230, configured with the A14 tune and `NNPDF2.3lo` PDF set, to model the parton shower, hadronisation, and the underlying event. The decays of bottom and charm hadrons are performed by EvtGen v1.6.0 [30]. The $t\bar{t}$+jets sample is generated inclusively using fast simulation. Events are further categorised depending on the flavour of partons that are matched to particle jets which do not originate from the decay of the $t\bar{t}$ system. Particle jets are reconstructed from all stable particles generated in the event (not counting muons and neutrinos) using the anti-$k_t$ algorithm [31] with a radius parameter $R$ = 0.4 and are required to have $p_{\mathrm{T}} > 15$ 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 $p_{\mathrm{T}} > 5$ 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

[2] The $h_{\mathrm{damp}}$ parameter is a resummation damping factor and one of the parameters that controls the matching of Powheg matrix elements to the parton shower and thus effectively regulates the high-$p_{\mathrm{T}}$ radiation against which the $t\bar{t}$ system recoils.

$t\bar{t}$+≥1$b$, and where at least one particle jet is matched to a charm quark not originating from a $W$ boson decay, are labelled as $t\bar{t}$+≥1$c$ events. Events labelled as either $t\bar{t}$+≥1$b$ or $t\bar{t}$+≥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-jet events.

The $Wt$ single top quark background is generated with Powheg-Box 2.0 using the `NNPDF3.0nnlo` PDF set. Overlaps between the $t\bar{t}$ and $Wt$ final states are removed using the "diagram removal" scheme [32]. Single top quark samples are interfaced to Pythia 8.230 with the Perugia 2012 underlying-event tuned parameters. The EvtGen v1.2.0 program is used to model properties of the bottom and charm hadron decays.

The production of $t\bar{t}V$ events is modelled using the MadGraph5_aMC@NLO v2.3.3 [33] generator at NLO with the `NNPDF3.0nlo` [19] PDF set. The events are interfaced to Pythia 8.230 using the A14 tuned parameters and the NNPDF2.3 LO PDF set. The decays of bottom and charm hadrons are simulated using the EvtGen v1.2.0 program [30].

The production of $t\bar{t}H$ events is modelled using the PowhegBox [25–28] generator to NLO with the NNPDF3.0NLO PDF set. The events are interfaced with Pythia 8.230 using the A14 tuned parameters and the NNPDF2.3 LO PDF set. The cross-sections are calculated at NLO QCD and NLO EW accuracy using MadGraph5_aMC@NLO [34]. The total cross-section is $507^{+35}_{-50}$ fb, where the uncertainties are from variations of renormalization and factorization scales as well as $\alpha_S$ variations.

Signal events are produced using MadGraph5_aMC@NLO v2.3.3 generator at NLO with the NNPDF2.3 LO PDF then interfaced to Pythia 8.210 and fast simulation of the detector response. Signal cross-section calculations include approximate next-to-next-to-leading order ($\text{NNLO}_{\text{Approx}}$) supersymmetric QCD corrections and the resummation of soft gluon emission at next-to-next-to-leading-logarithmic (NNLL) accuracy [35]. The nominal cross-section and its uncertainty are taken from an envelope of predictions using different PDF sets as well as different factorization and renormalization scales. Top-squark masses between 600 GeV and 1 TeV and higgsino masses between 100 GeV and 950 GeV are considered.

## 4 Event reconstruction

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

Electrons are reconstructed from energy deposits (clusters) in the electromagnetic calorimeter matched to tracks reconstructed in the ID [36, 37] and are required to have $p_{\text{T}} > 10$ GeV and $|\eta| < 2.47$. Candidates in the calorimeter barrel–endcap transition region ($1.37 < |\eta| < 1.52$) are excluded. *Loose* Electrons must satisfy the 'Medium' identification criterion described in Ref. [37] that is based on a likelihood discriminant combining observables related to the shower shape in the calorimeter and to the track matching the electromagnetic cluster. Electron tracks must match the primary vertex of the event: the longitudinal impact parameter[3] 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 on the measured $|d_0|$ values. *Tight* electrons

[3] The transverse impact parameter is defined as the distance of closest approach in the transverse plane between a track and the beam-line. The longitudinal impact parameter corresponds to the $z$-coordinate distance between the point along the track atwhich the transverse impact parameter is defined and the primary vertex.

are required to pass the 'TightLH' requirement, satisfy the 'Gradient' isolation criteria [37], and have $p_{\mathrm{T}} > 27$ GeV.

Muons are reconstructed by matching either track segments or full tracks in the MS to tracks in the ID [38]. 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 $|d_0|/\sigma_{d_0} < 3$. For the purposes of this note, *Loose* muons are those that pass the 'Loose' muon selection and have $p_{\mathrm{T}} > 10$ GeV and $|\eta| < 2.5$, and *Tight* muons are those that pass the 'Medium' muon selection, satisfy the 'FixedCutTightTrackOnly' isolation criterion [38], and have $p_{\mathrm{T}} > 27$ GeV.

Candidate jets are reconstructed from three-dimensional topological energy clusters [39] in the calorimeter using the anti-k$_T$ jet algorithm with a radius parameter of 0.4. Each topological cluster is calibrated to electromagnetic scale prior to jet reconstruction. Reconstructed jets are then corrected to the particle level by the application of a jet energy scale (JES) calibration that is derived from simulation and by *in situ* corrections obtained from 13 TeV data [40, 41]. Jets used in this analysis are required to have $p_{\mathrm{T}} > 25$ GeV and $|\eta| < 2.5$ after calibration.

To avoid selecting jets from pile-up, low $p_{\mathrm{T}}$ ($p_{\mathrm{T}} < 120$ GeV) jets in the central ($|\eta| < 2.5$) region of the detector are required to satisfy the Jet Vertex Tagger (JVT) [42] configured such that it has an efficiency of approximately 92% to identify jets from a primary vertex. This requirement is applied both to data and simulation. Quality criteria are imposed to identify jets arising from non-collision sources or detector noise (using the `LooseBad` operating points) and any event containing at least one such jet is removed [43]. This removal produces a negligible loss of efficiency for signal events.

Jets originating from the hadronisation of a $b$-quark, referred to as $b$-tagged 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 [44], trained on simulated $t\bar{t}$ events to discriminate $b$-jets from a background consisting of light- (90%) and $c$-labeled (10%) jets [45]. Jets are considered $b$-tagged if they satisfy a lower requirement on the MV2c10 $b$-tagging weight corresponding to an average efficiency in $t\bar{t}$ events of 60% for $b$-jets, 5% for $c$-jets and a rejection factor of approximately 1200 for light-jets across the jet $p_{\mathrm{T}}$ 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 $\Delta R_y = \sqrt{(\Delta y)^2 + (\Delta\phi)^2} = 0.2$ 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

For the analysis selection, events are required to have at least five jets, of which at least two must be $b$-tagged. The four highest $p_{\mathrm{T}}$ jets are required to be on the trigger efficiency plateau, namely to have

[4] The rapidity is defined as $y = \frac{1}{2} \ln \frac{E+p_z}{E-p_z}$ where $E$ is the energy and $p_z$ is the longitudinal component of the momentum along the beam pipe.

| Analysis Regions | | $N_b$ 3 | 4 | $\geq 5$ |
|---|---|---|---|---|
| $N_j$ | 6 | $\leftarrow$ VR-MJ $\rightarrow C_{mass}^{max}$ | $\longleftarrow$ $SR_{stop}$ $\longrightarrow$<br>$\leftarrow$ VR-MJ $\rightarrow C_{mass}^{max}$ | $\longleftarrow$ $SR_{stop}$ $\longrightarrow$ |
| | 7 | $\leftarrow$ VR-MJ $\rightarrow C_{mass}^{max}$ | $\longleftarrow$ $SR_{stop}$ $\longrightarrow$<br>$\leftarrow$ VR-MJ $\rightarrow C_{mass}^{max}$ | $\longleftarrow$ $SR_{stop}$ $\longrightarrow$ |
| | 8 | $\leftarrow$ VR-MJ $\rightarrow C_{mass}^{max}$ | $\longleftarrow$ $SR_{stop}$ $\longrightarrow$<br>$\leftarrow$ VR-MJ $\rightarrow C_{mass}^{max}$ | $\leftarrow$ $SR_{stop}$, $SR_{discovery}$ $\rightarrow$ |
| | $\geq 9$ | $\leftarrow$ VR-MJ $\rightarrow C_{mass}^{max}$ | $\longleftarrow$ $SR_{stop}$ $\longrightarrow$<br>$\leftarrow$ VR-MJ $\rightarrow C_{mass}^{max}$ | $\leftarrow$ $SR_{stop}$, $SR_{discovery}$ $\rightarrow$ |

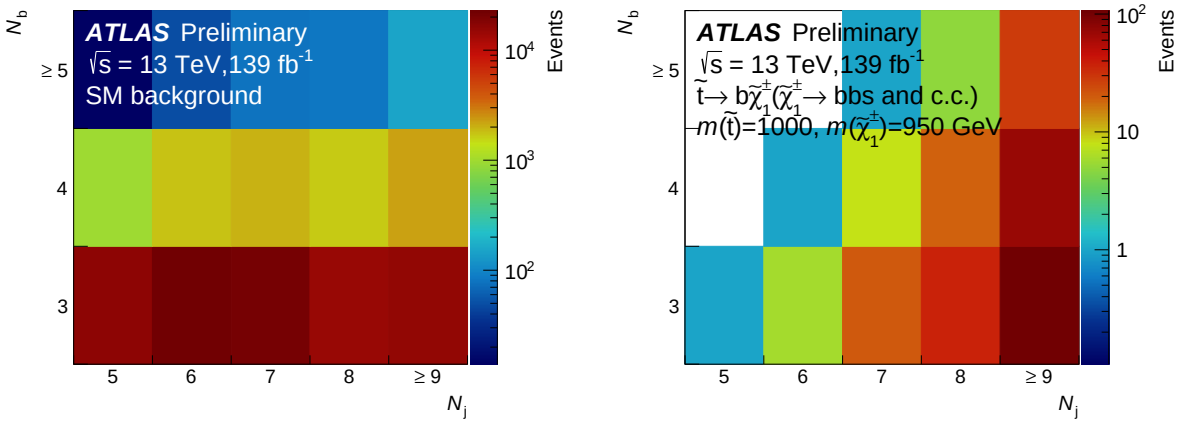
Table 1: Table illustrating the strategy of the analysis. For the model-dependent fit, the signal regions ($SR_{stop}$) consists of events with $N_j$ = 6, 7, 8 and $\geq 9$ jets and $N_b$ = 4 and $\geq 5$. These are used independently in the final fit. For the model-independent fit, two dedicated signal regions ($SR_{discovery}$), with $N_j \geq 9$ and $N_b \geq 5$ and $N_j \geq 8$ and $N_b \geq 5$ are used. The validation regions (VR-MJ) based on an upper value of the centrality mass, $C_{mass}^{max}$, are also indicated.

$p_T > 120$ GeV or $p_T > 140$ GeV, depending on the jet $p_T$ trigger requirement, and have $|\eta| < 2.5$. All other jets present in the event are required to have $p_T > 25$ GeV and $|\eta| < 2.5$. Events having isolated *loose* muons or electrons with $p_T > 10$ GeV are discarded.

After the selections described above, the largest background contribution to the measurement is multijet production from light quark and gluonic final states. The next largest is $t\bar{t}$+jets production which is well below 10% of the total background in any signal region defined for the analysis. Other small background contributions originate from the production of a single top quark and from the associated production of either a vector boson or the Higgs boson and a $t\bar{t}$ pair. The multijet background is estimated using a data-driven method validated using MC.

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 jets and $b$-tagged jet multiplicity, $N_b$ = 4 and $\geq 5$, is performed. These ($N_j$, $N_b$) regions are indicated as $SR_{stop}$ in Table 1. The signal contribution predicted for different values of $m_{\tilde{t}}$ and $m_{\tilde{\chi}_1^\pm, \tilde{\chi}_{1,2}^0}$ is considered in all bins and is scaled by one common signal-strength parameter. For the model considered here, the product of acceptance and reconstruction efficiency ($A \times \epsilon$) is of order $\sim 5 \times 10^{-2}$ 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 whose evaluation is described in Section 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 whose evaluation is described in Section 6, 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 based on an upper value of the centrality mass ($C_{mass}$), defined as the ratio of the scalar sum of all jet $p_T$ in the event ($H_T$) and the invariant mass of the vectorial sum of all jets four-momentum, is applied. The $C_{mass}^{max}$ selections limit the modelled SUSY signal contributions to less than 5%.

A separate, model-independent test is used to search for and to set generic exclusion limits on potential contributions from a hypothetical BSM signal presenting an experimental signature with large jet and




(a) (b)

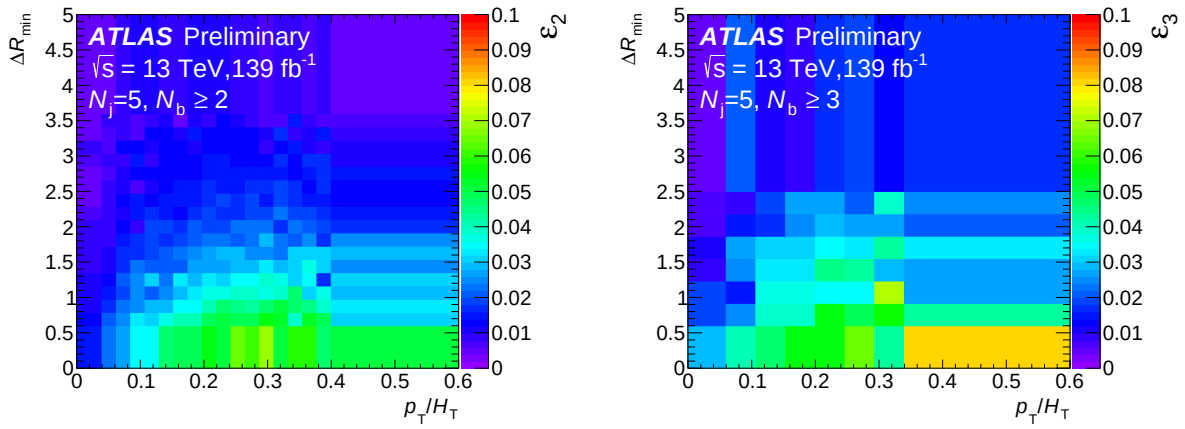
Figure 2: $N_j$ versus $N_b$ distribution of the predicted number of events for (a) SM background (multijet and top quark production) and (b) stop pair production in the $\tilde{t} \to \bar{b}\tilde{\chi}_1^{\pm}$ ($\tilde{\chi}_1^{\pm} \to bbs$ and c.c.) channel for $m_{\tilde{t}}$=1000 GeV and $m_{\tilde{\chi}_1^{\pm}}$=950 GeV (b).

$b$-tagged jet multiplicity. In the search for a hypothetical BSM signal, two dedicated signal regions with $N_j \geq 9$ and $N_b \geq 5$ and $N_j \geq 8$ and $N_b \geq 5$ are defined ($SR_{discovery}$ in Table 1). The two regions are not mutually exclusive hence an event can enter more than one $SR_{discovery}$ region.

## 6 Multijet background estimation

The multijet background is estimated via a data-driven method, subsequently referred to as the tag-rate function method for multijet events ($TRF_{MJ}$) [46, 47], 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 $TRF_{MJ}$ method uses a tag-rate function, $\varepsilon$, to quantify the experimental probability of $b$-tagging an additional jet in a sample of events with at least two ($\varepsilon_2$) or at least three ($\varepsilon_3$) $b$-tagged jets, and to estimate the $b$-tag multiplicity shape in each $N_j$ value.

In the first step, the $b$-tagging probability is measured in events passing the selection criteria described in Section 5 with an additional requirement to have exactly five jets with $p_T > 25$ GeV and $|\eta| < 2.5$, and at least two $b$-tagged jets. The data are first corrected by subtracting the expected yield from non-multijet background based on simulation. Excluding the two jets with the highest $b$-tagging weight in the event, $\varepsilon_2$ is defined as the probability of $b$-tagging any other jet in the event. In the same way $\varepsilon_3$ is defined from events with at least three $b$-tagged jets, where the three jets with the highest $b$-tagging weight are excluded. The two $\varepsilon$'s are parameterised as a function of both the $p_T$ of each remaining jet used for the $\varepsilon$ computation divided by $H_T$, and the minimum $\Delta R$ between that jet and the two, for $\varepsilon_2$, or three, for $\varepsilon_3$, jets with the largest $b$-tagging weight in the event ($\Delta R_{min}$). This choice of variables for the parametrization is made to minimize the residual differences between the $TRF_{MJ}$ prediction and the number of events in the most sensitive signal regions in MC multijet events (non-closure). The dependence of the values of $\varepsilon_2$ and $\varepsilon_3$ on both $p_T/H_T$ and $\Delta R_{min}$ are shown in Figure 3. The rapid variation of both $\varepsilon_2$ and $\varepsilon_3$ on $\Delta R_{min}$ is consistent




(a) (b)

Figure 3: Two-dimensional distributions of (a) $\varepsilon_2$ and (b) $\varepsilon_3$ as a function of the ratio of jet transverse momentum to $H_{\mathrm{T}}$, $p_{\mathrm{T}}/H_{\mathrm{T}}$, and the minimum $\Delta R$ between the jet and the two (a) or three (b) $b$-tagged jets in the event with the highest $b$-tagging weight, $\Delta R_{\mathrm{min}}$. The choice of binning is made so as to avoid empty bins.

with the dependence expected from multi-$b$-jet production due to gluon-splitting. The $p_{\mathrm{T}}/H_{\mathrm{T}}$ dependence, more visible at small $\Delta R_{\mathrm{min}}$, reflects the variation of the $b$-tagging efficiency with jet $p_{\mathrm{T}}$.

Following the methods from Ref. [48], in the second step of the $\mathrm{TRF_{MJ}}$ method the number of events with different number of $b$-tagged jets is estimated for each $N_{\mathrm{j}}$ value by weighting all events with $N_{\mathrm{b}} \geq 2$ by the event probability of having $N_{\mathrm{b}} = 2$, 3, 4 and $\geq$5, respectively. Upon subtracting non-multijet background contribution [46], the event probabilities are estimated by using both $\varepsilon_2$ and $\varepsilon_3$, after first excluding the two jets with the highest $b$-tagging weight. For the estimate in $N_{\mathrm{b}} = 3$, 4 and $\geq$5, a two-step procedure is employed. First, a "pseudodata sample" with $N_{\mathrm{b}} \geq 3$ is emulated, using $\varepsilon_2$ in events with $N_{\mathrm{b}} \geq 2$. This sample is then used as the starting point for the re-application of the same procedure, this time relying on $\varepsilon_3$. To make use of $\varepsilon_3$, a third $b$-tagged jet has to be emulated - in addition to the two with the highest $b$-tagging weight. This is done by randomly choosing a jet from the remaining $N_{\mathrm{j}}$ - 2 jets in the event to be $b$-tagged using a probability-dependent weight [47].

## 6.1 Validation of $\mathrm{TRF_{MJ}}$ method

The $\mathrm{TRF_{MJ}}$ method is validated using two different comparisons to data: in the VR-MJ regions defined in Section 5 and in a separate set of $Z$ + jets-enriched events. Figure 4 shows a comparison between measured and estimated event rates in VR-MJ. Data and predictions are found to be in good agreement within systematic uncertainties (described in Section 7), although differences at the level of 1 standard deviation are observed for high $N_{\mathrm{j}}$.

An independent test of the method is performed in $Z$ + jets-enriched events, referred as "VR-ZJ", where additional jets are produced by radiation similar to multijet events and where $b\bar{b}$ pairs arise from gluon splitting. In order to select events where a $Z$-boson decays into two electrons or muons, events are required to pass a single-lepton trigger. Two opposite-sign, same-flavour, *tight* leptons are required each to have $p_{\mathrm{T}} > 27$ GeV and a pair mass larger than 60 GeV. Events are required to have at least four jets with $p_{\mathrm{T}} > 25$ GeV and $|\eta| < 2.5$, of which at least two must be $b$-tagged. The tagging probabilities $\varepsilon_2$ and $\varepsilon_3$ are

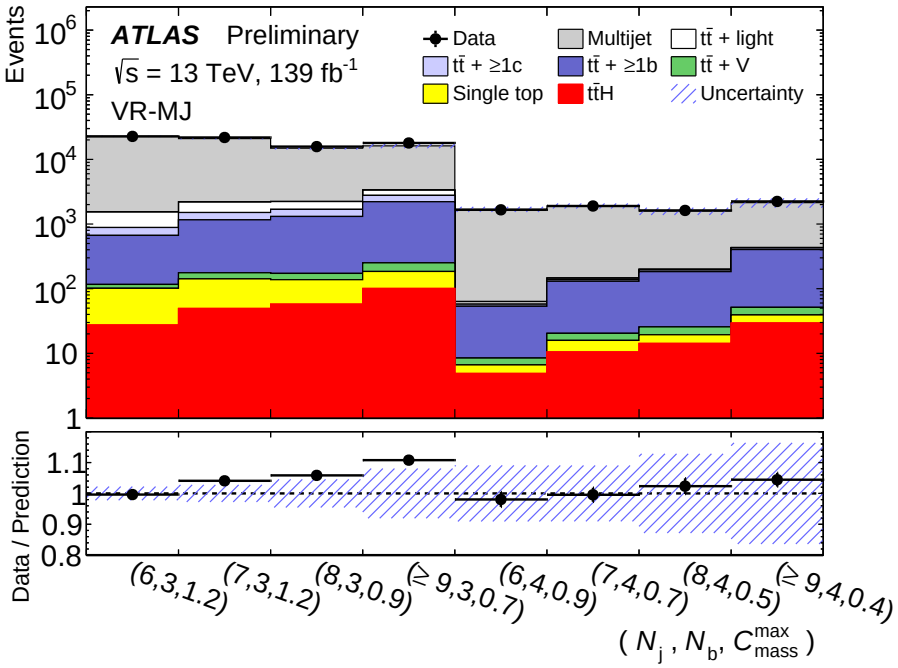



Figure 4: Comparison between data and the predicted number of events with $N_j$ = 6, 7, 8 and ≥ 9 and $N_b$ = 3 and 4 in the VR-MJ validation region, based on an upper value on the centrality mass $C_{mass}$ ($C_{mass}^{max}$). Systematic uncertainties are represented by the blue hatched area.

derived from the five-jet VR-ZJ events and used to predict the number of events in $N_j$ = 6, 7, 8, ≥ 9 and $N_b$ = 4, ≥ 5. The final result, illustrated in Figure 5, shows that, despite the large statistical uncertainties, data and predictions agree within systematic uncertainties whose evaluation is described in Section 7.

## 7 Systematic uncertainties

Several sources of systematic uncertainty are considered that can affect the overall normalization 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 the $TRF_{MJ}$ background estimates. Uncertainties related to the theoretical modelling and due to the description of the detector response in simulated events are only relevant for the signal and background MC samples, and for the estimates of the signal yields after selections.

The main assumption of the $TRF_{MJ}$ method is that it is possible to define per-jet $b$-tagging probabilities ($\varepsilon_2$ and $\varepsilon_3$ ) in events with at least two and three $b$-tagged jets and, in particular, that the variables used for the parametrization are sensitive to the heavy-flavour composition of the jet sample. The second assumption is that the per-jet probabilities are universal 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 ≥ 9 jets. The validity of these assumptions is verified using MC simulations. The $TRF_{MJ}$ method is applied to MC Pythia 8 dijet events, and the largest of a) the residual non-closure and b) the statistical uncertainty on the number of events with a given $b$-tagged jet multiplicity, is taken to be the systematic uncertainty associated to the method. Table 2 shows the final $TRF_{MJ}$ systematic uncertainty on the multijet background estimation in each ($N_j$, $N_b$) region.

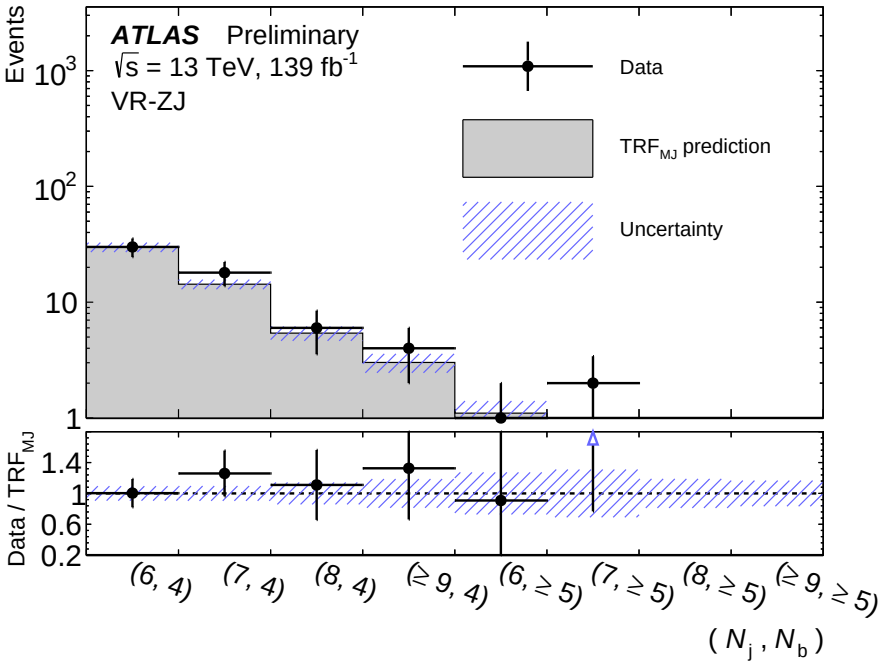



Figure 5: Comparison between data and the number of events with $N_j$ = 6, 7, 8 and ≥ 9 and $N_b$ =4 and ≥5 predicted by the $TRF_{MJ}$ method (grey histogram) in the VR-ZJ region, defined by the requirement of two isolated leptons with invariant mass larger than 60 GeV. Systematic uncertainties on the $TRF_{MJ}$ prediction are represented by the blue hatched area.

For $N_b = 4$ the $TRF_{MJ}$ uncertainties are dominated by the non-closure component, while for $N_j \geq 5$, the statistical component dominates. The $TRF_{MJ}$ uncertainties represent the largest source of uncertainty for the analysis.

| $TRF_{MJ}$ uncertainty | | $N_b$ 4 | $N_b$ ≥5 |
|---|---|---|---|
| $N_j$ | 6 | 9% | 27% |
| | 7 | 9% | 30% |
| | 8 | 13% | 18% |
| | ≥9 | 16% | 14% |

Table 2: Systematic uncertainties on the data-driven multijet background estimation $TRF_{MJ}$ method assessed in Pythia 8 dijet MC events for each value of jet multiplicity ($N_j$) and $b$-tagged jet multiplicity ($N_b$) used in the final fit.

The second largest contribution to the total systematic uncertainty arises from the modelling of the $t\bar{t}$+jets background. Since the diagrams that contribute to $t\bar{t}$+≥1$b$, $t\bar{t}$+≥1$c$, and $t\bar{t}$ +light production are different, and the uncertainties may affect these processes in different regions in different ways, all uncertainties on $t\bar{t}$+jets background modelling, except the uncertainty on the inclusive cross-section, are considered to be uncorrelated among $t\bar{t}$+≥1$b$, $t\bar{t}$+≥1$c$, and $t\bar{t}$+light.

An uncertainty of ±6% is applied for the inclusive $t\bar{t}$ NNLO+NNLL production cross-section [49]. This uncertainty includes effects from varying the factorisation and renormalisation scales, the PDF, $\alpha_S$, and

the top-quark mass. The normalisations of $t\bar{t}$+≥1$c$ and $t\bar{t}$+≥1$b$ yields are taken from their fractional contribution to the nominal $t\bar{t}$+jets sample as generated using the Powheg-Box. In addition to the uncertainty on the inclusive $t\bar{t}$ cross section, an additional uncertainty of 50%, based on the measure of the $t\bar{t}$+≥1$b$ and $t\bar{t}$+≥1$c$ normalization factors performed in Ref. [50], is assigned to $t\bar{t}$+≥1$c$ and $t\bar{t}$+≥1$b$ production cross-sections.

The impact of the parton shower and hadronisation model uncertainties on $t\bar{t}$+jets, $t\bar{t}H$ and $Wt$ single top quark yields is evaluated by comparing the nominal generator setup with a sample produced with the NLO PowhegBox v2 generator using the NNPDF3.0NLO PDF set. The events are interfaced with Herwig7.04 [51, 52], using the H7UE set of tuned parameters [52] and the MMHT2014LO PDF set [53] and, then, processed using fast simulation of the detector response. The difference between the two predictions of $t\bar{t}$+≥1$b$ events ranges from 20% (33%) for $N_{\text{j}} = 6$ and $N_{\text{b}} = 4$ (5) to 46% (60%) in the region with $N_j \geq 9$ and $N_{\text{b}} = 4$ ($\geq 5$).

To assess the uncertainty due to the choice of the matching scheme, the Powheg-Box sample is compared to a generator combining MadGraph5_aMC@NLO and Pythia 8. For the calculation of the hard scattering, MadGraph5_aMC@NLO v2.6.0 with the NNPDF3.0NLO PDF set is used. The events are interfaced with Pythia 8.230, using the A14 set of tuned parameters and the NNPDF23LO PDF and, then, simulated using fast simulation of the detector response. This uncertainty which is obtained from the difference in yields between the two models affects both the normalization and the $N_{\text{j}}$ and $N_{\text{b}}$ dependence of background rates. It is maximal for large values of the jet and $b$-tagged jet multiplicities. For $t\bar{t}$+≥1$b$, it reaches 25% for $N_{\text{j}} = 8$, $\geq 9$ and $N_{\text{b}} = 4$, and 41% (32%) for $N_{\text{j}} = 8$, $\geq 9$ and $N_{\text{b}} \geq 5$.

For $t\bar{t}H$ and $t\bar{t}V$ events, MC weights are used to evaluate the renormalization and factorization scale uncertainties and PDF uncertainties. For the former, the scales are varied simultaneously by common factors of 2.0 and 0.5. For the latter, the 100 variations for NNPDF3.0NLO are taken into account. An uncertainty of ±5% is used on the total cross-section for single-top production [54–56]. For both $t\bar{t}H$ and single top events, an additional uncertainty on initial and final-state radiation and choice of generator is evaluated in a manner similar to that used for $t\bar{t}$ + jets. The uncertainty on the interference between $Wt$ and $t\bar{t}$ production at NLO is assessed by comparing the default "diagram removal" scheme to an alternative "diagram subtraction" scheme [32].

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 $\tilde{t} \rightarrow b\tilde{\chi}_1^{\pm}(\tilde{\chi}_1^{\pm} \rightarrow bbs$ and c.c.) signal model, the $b$-tagging uncertainties in the region $N_j \geq 9$ and $N_{\text{b}} = 4$ are ≈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 factorization, renormalization, 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 on the signal cross-section ranges between 8% and 11% for top-squark mass in the range 600 – 1000 GeV.

## 8 Results

The events are divided up into ($N_{\text{j}}$, $N_{\text{b}}$) regions with different signal-to-background ratios in order to constrain systematic uncertainties and for improved 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.

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

Table 3: Event yields from background predictions (pre-fit) and data in the regions with $N_j$ = 6, 7, 8 or ≥ 9 and $N_b$ = 4. Combined systematic and statistical uncertainties are shown.

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

Table 4: Event yields from background predictions (pre-fit) and data in the regions with $N_j$ = 6, 7, 8 or ≥ 9 and $N_b$ ≥ 5. Combined systematic and statistical uncertainties are shown.

For interpretation, hypothesis testing is performed using a modified frequentist method as implemented in RooStats [57] and 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 high-statistics, background-dominated ($N_j$, $N_b$) bins included in the likelihood fit. The signal-strength parameter, $\mu_{\tilde{t}\tilde{t}^*}$, defined for positive values and corresponding to the signal normalisation, is included and free-floating in the profile-likelihood fit. The normalisation of each component of the background and $\mu_{\tilde{t}\tilde{t}^*}$ are determined simultaneously from the fit to the data.

Individual sources of systematic uncertainty are taken as uncorrelated. Contributions from $t\bar{t}+ \geq 1b$,

$t\bar{t}+ \geq 1c$, $t\bar{t}$ + 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 Section 3), and experimental data. The $\mathrm{TRF_{MJ}}$ 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 [58] prescription. The likelihood is configured differently for the model-independent and model-dependent hypothesis tests.

For the model-independent test, a profile-likelihood fit is performed independently in two $\mathrm{SR_{discovery}}$ regions with $N_\mathrm{j} \geq 8$, $N_\mathrm{b} \geq 5$ and $N_\mathrm{j} \geq 9$, $N_\mathrm{b} \geq 5$. 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 $\mathrm{SR_{discovery}}$ region.

For the model-dependent test, assuming a specific top-squark model, exclusion tests of the signal-plus-background hypothesis, i.e. $\mu_{\tilde{t}\tilde{t}^*} = 1$, are formed for each $m_{\tilde{t}}$ and $m_{\chi_1^\pm,\chi_{1,2}^0}$. These are used to derive exclusion limits for the specific top-squark model. A full set of regions with $N_\mathrm{j} = 6$, 7, 8 and $\geq 9$ and $N_\mathrm{b} = 4$ and $\geq 5$ is employed in the likelihood. The signal contribution, as predicted by the given model, is considered in all bins and is scaled by $\mu_{\tilde{t}\tilde{t}^*}$.

Figure 6 shows the observed numbers of data events compared to the fitted background model. The likelihood fit is configured using the model-dependent set-up where all bins are input to the fit, and $\mu_{\tilde{t}\tilde{t}^*}$ is set to zero. This configuration is also referred to as *background-only* fit. An example signal model is also shown in the figure to illustrate the achieved separation between the signal and the background.

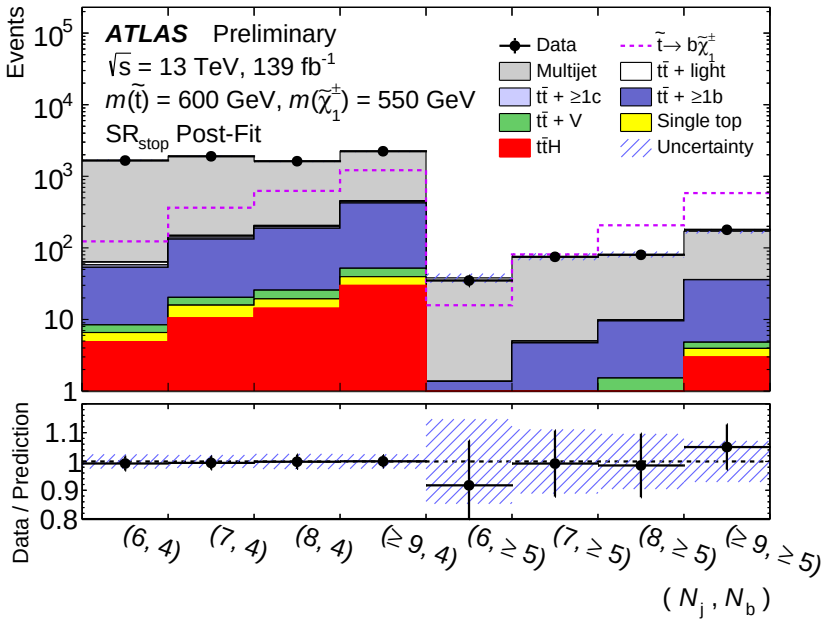



Figure 6: Expected background and observed number of events in different jet and $b$-tag multiplicity bins. The background is estimated by including all bins in a background-only fit and is plotted separately for each contribution. An example signal yield for $\tilde{t} \to \bar{b}\tilde{\chi}_1^\pm$ ($\tilde{\chi}_1^\pm \to bbs$ and c.c.) production with $m_{\tilde{t}} = 600$ GeV and $m_{\tilde{\chi}_1^\pm} = 550$ GeV is overlaid. All uncertainties, which can be correlated across bins, are included in the error bands (hatched regions).

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

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.

### 8.1 Model-independent interpretation

The model-independent results are calculated from the observed number of events and background predictions in the two independent $SR_{discovery}$ regions. The observed number of events and backgrounds obtained from the fits are shown for both $SR_{discovery}$ regions in Table 5.

Model-independent upper limits at 95% CL on the expected and observed number of BSM events, $N^{95}_{exp}$ and $\sigma^{95}_{obs}$, that may contribute to the signal regions are computed from the observed number of events and the fitted background. Normalizing these results by the integrated luminosity, $L$, of the data sample, they can be interpreted as upper limits on the visible BSM cross-section $\sigma^{95}_{obs}$, defined as:

$$\sigma^{95}_{obs} = \sigma_{prod} \times A \times \epsilon = \frac{N^{95}_{obs}}{L}, \tag{1}$$

where $\sigma_{prod}$ 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 $Z$.

| Signal region | $\sigma^{95}_{obs}$ [fb] | $N^{95}_{obs}$ | $N^{95}_{exp}$ | $p_0$ ($Z$) |
|---|---|---|---|---|
| $N_j \geq 8$, $N_b \geq 5$ | 0.76 | 105 | $85^{+30}_{-24}$ | 0.24 (0.7) |
| $N_j \geq 9$, $N_b \geq 5$ | 0.54 | 75 | $52^{+20}_{-15}$ | 0.11 (1.2) |

Table 6: Observed 95% CL model-independent upper limits on $\sigma^{95}_{obs}$, expressed in fb, obtained from the product of cross-section, acceptance and efficiency for the $N_j \geq 8$, $N_b \geq 5$ and $N_j \geq 9$, $N_b \geq 5$ signal regions and the relative number of observed (expected) number of BSM events, $N^{95}_{obs}$ ($N^{95}_{exp}$). The $p_0$ value quantifies 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, and $Z$ represents the significance.

### 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 Section 8. Figure 7 shows exclusion limits in the top-squark production model when BR($\tilde{t} \to b\chi_1^{\pm}$) 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 this final state to be produced. For lower value of chargino mass, the limit weakens such that for chargino masses of around 200 GeV, the top-squark mass is constrained to be less than 800 GeV. In this phase-space, the signal is concentrated at lower $N_{\mathrm{j}}$ and $N_{\mathrm{b}}$ values where the background is larger.

The limit for pure higgsino LSPs are shown in Figure 8. In the region, $m_{\tilde{t}} - m_{\tilde{\chi}^0_{1,2},\tilde{\chi}^{\pm}_1} \leq m_t$, the sensitivity of the analysis is reduced compared to the pure $\tilde{t} \to b\chi_1^{\pm}$ case, because contributions to the signal with one leptonically-decaying top fail the lepton-veto requirement.

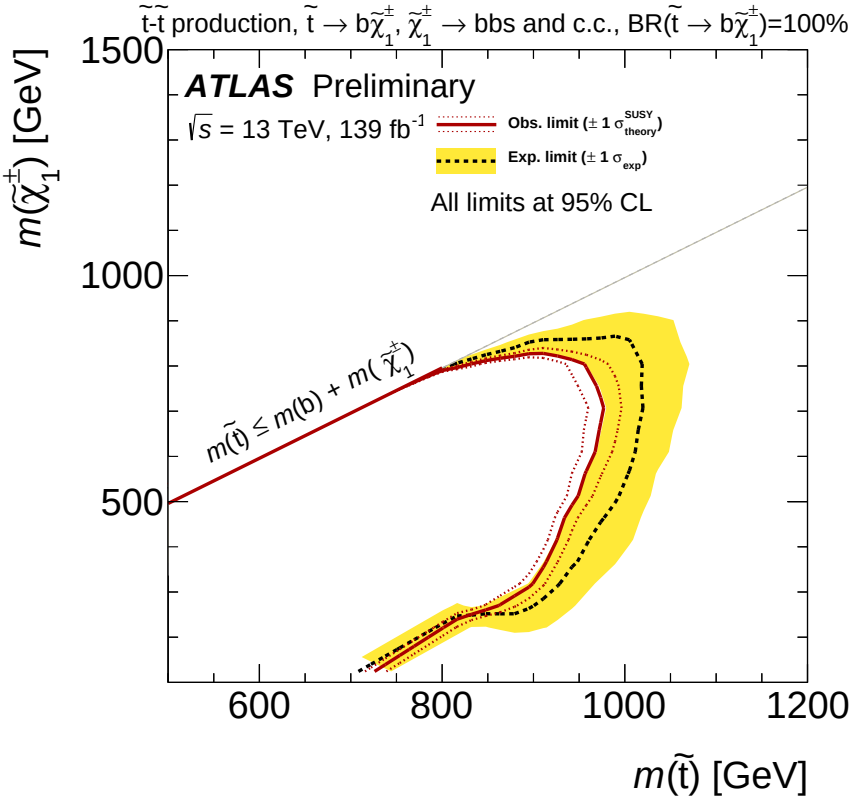



Figure 7: Observed and expected exclusion contours on the $\tilde{t}$ and $\chi_1^{\pm}$ masses in the context of top-squark production model with RPV decays of the $\chi_1^{\pm}$. Limits are shown in the case of BR($\tilde{t} \to b\chi_1^{\pm}$) equal to unity. The contours of the band around the expected 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 diagonal line indicates the kinematic limit for the decays in the considered scenario.

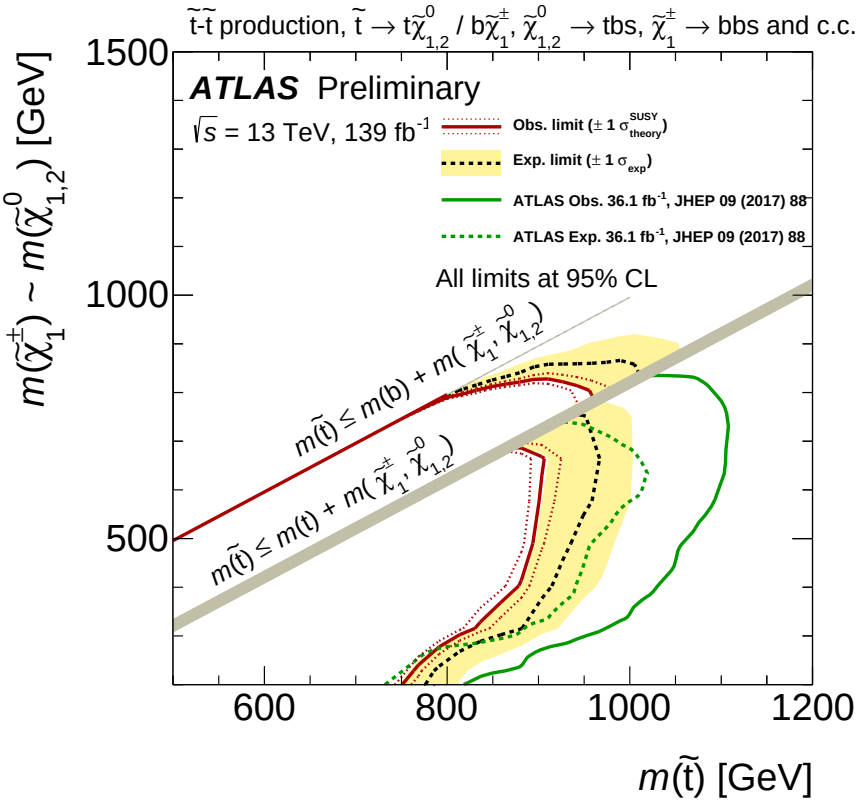



Figure 8: Observed and expected exclusion contours on the $\tilde{t}$ and $\chi_1^\pm$ masses in the context of top-squark production model with RPV decays of the $\chi_1^\pm$. Limits are shown in the case of pure Higgsino ($\tilde{H}$) LSP. The contours of the band around the expected limit are the $\pm 1\sigma$ 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 diagonal line indicates the kinematic limit for the decays in the considered scenario. Also shown are the limits from Ref. [11].

# 9 Conclusion

A search for beyond the Standard Model physics in events with high jet multiplicity and a large number of $b$-tagged jets is described in this note. Unlike many previous searches in similar final states, leptons are vetoed and no requirement on the missing transverse momentum in the event is applied. A data-driven technique is used to estimate the dominant multijet background. With no significant excess over the SM expectation observed, model-independent limits on the contribution of new phenomena to the signal-region yields are computed. In the context of this model with direct top-squark production and RPV decays of the higgsinos, the data exclude top squarks with masses up to 950 GeV in a region where this analysis has exclusive sensitivity. The results represent the first limits for the top-squark production decaying exclusively into a chargino and a $b$-quark.

## References


[1] J. A. Evans, *A Swarm of Bs*, JHEP **1408** (2014) 073, arXiv: 1402.4481 [hep-ph].

[2] M. R. Buckley, D. Feld, S. Macaluso, A. Monteux and D. Shih, *Cornering Natural SUSY at LHC Run II and Beyond*, JHEP **08** (2017) 115, arXiv: 1610.08059 [hep-ph].

[3] S. Diglio, L. Feligioni and G. Moultaka, *Stashing the stops in multijet events at the LHC*, Phys. Rev. **D96** (2017) 055032, arXiv: 1611.05850 [hep-ph].

[4] P. Fayet, *Supergauge Invariant Extension of the Higgs Mechanism and a Model for the Electron and Its Neutrino*, Nucl. Phys. B **90** (1975) 104.

[5] A. Salam and J. A. Strathdee, *Supersymmetry and Fermion Number Conservation*, Nucl. Phys. **B 87** (1975) 85.

[6] R. Barbier et al., *R-parity violating supersymmetry*, Phys. Rept. **420** (2005) 1, arXiv: hep-ph/0406039.

[7] R. Barbieri and G. F. Giudice, *Upper Bounds on Supersymmetric Particle Masses*, Nucl. Phys. B **306** (1988) 63.

[8] B. de Carlos and J. A. Casas, *One loop analysis of the electroweak breaking in supersymmetric models and the fine tuning problem*, Phys. Lett. B **309** (1993) 320, arXiv: hep-ph/9303291.

[9] ATLAS Collaboration, *A search for pair-produced resonances in four-jet final states at $\sqrt{s}$ = 13 TeV with the ATLAS detector*, Eur. Phys. J. C **78** (2018) 250, arXiv: 1710.07171 [hep-ex].

[10] CMS Collaboration, *Search for pair-produced resonances decaying to quark pairs in proton-proton collisions at $\sqrt{s}$ = 13 TeV*, Phys. Rev. D **98** (2018) 112014, arXiv: 1808.03124 [hep-ex].

[11] ATLAS Collaboration, *Search for new phenomena in a lepton plus high jet multiplicity final state with the ATLAS experiment using $\sqrt{s}$ = 13 TeV proton-proton collision data*, JHEP **09** (2017) 088, arXiv: 1704.08493 [hep-ex].

[12] CMS Collaboration, *Search for pair-produced resonances each decaying into at least four quarks in proton-proton collisions at $\sqrt{s}$ = 13 TeV*, Phys. Rev. Lett. **121** (2018) 141802, arXiv: 1806.01058 [hep-ex].

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

[14] ATLAS Collaboration, *ATLAS Insertable B-Layer Technical Design Report*, tech. rep. CERN-LHCC-2010-013. ATLAS-TDR-19, 2010, URL: https://cds.cern.ch/record/1291633.

[15] ATLAS Collaboration, *Luminosity determination in pp collisions at $\sqrt{s}$ = 13 TeV using the ATLAS detector at the LHC*, ATLAS-CONF-2019-021, 2019, URL: https://cds.cern.ch/record/2677054.

[16] G. Avoni et al., *The new LUCID-2 detector for luminosity measurement and monitoring in ATLAS*, Journal of Instrumentation **13** (2018) P07017.

[17] T. Sjöstrand et al., *An introduction to PYTHIA 8.2*, Comput. Phys. Commun. **191** (2015) 159, arXiv: 1410.3012 [hep-ph].

[18] *The Pythia 8 A3 tune description of ATLAS minimum bias and inelastic measurements incorporating the Donnachie-Landshoff diffractive model*, tech. rep. ATL-PHYS-PUB-2016-017, CERN, 2016, URL: https://cds.cern.ch/record/2206965.

[19] NNPDF Collaboration, R.D. Ball et al., *Parton distributions for the LHC Run II*, JHEP **04** (2015) 040, arXiv: 1410.8849 [hep-ph].

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

[21] GEANT4 Collaboration, S. Agostinelli et al., *GEANT4: A Simulation toolkit*, Nucl. Instrum. Meth. A **506** (2003) 250.

[22] W. Lukas, *Fast Simulation for ATLAS: Atlfast-II and ISF*, tech. rep. ATL-SOFT-PROC-2012-065, CERN, 2012, URL: https://cds.cern.ch/record/1458503.

[23] R. D. Ball et al., *Parton distributions with LHC data*, Nucl. Phys. B **867** (2013) 244, arXiv: 1207.1303 [hep-ph].

[24] ATLAS Collaboration, *ATLAS Pythia 8 tunes to 7 TeV data*, ATL-PHYS-PUB-2014-021, 2014, URL: https://cds.cern.ch/record/1966419.

[25] S. Frixione, P. Nason and G. Ridolfi, *A positive-weight next-to-leading-order Monte Carlo for heavy flavour hadroproduction*, JHEP **09** (2007) 126, arXiv: 0707.3088 [hep-ph].

[26] P. Nason, *A new method for combining NLO QCD with shower Monte Carlo algorithms*, JHEP **11** (2004) 040, arXiv: hep-ph/0409146.

[27] S. Frixione, P. Nason and C. Oleari, *Matching NLO QCD computations with Parton Shower simulations: the POWHEG method*, JHEP **11** (2007) 070, arXiv: 0709.2092 [hep-ph].

[28] S. Alioli, P. Nason, C. Oleari and E. Re, *A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX*, JHEP **06** (2010) 043, arXiv: 1002.2581 [hep-ph].

[29] ATLAS Collaboration, *Studies on top-quark Monte Carlo modelling for Top2016*, ATL-PHYS-PUB-2016-020, 2016, URL: https://cds.cern.ch/record/2216168.

[30] D. J. Lange, *The EvtGen particle decay simulation package*, Nucl. Instrum. Meth. A **462** (2001) 152.

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

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

[33] J. Alwall et al., *The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations*, JHEP **07** (2014) 079, arXiv: 1405.0301 [hep-ph].

[34] D. de Florian et al., *Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector*, (2016), arXiv: 1610.07922 [hep-ph].

[35] W. Beenakker, C. Borschensky, M. Krämer, A. Kulesza and E. Laenen, *NNLL-fast: predictions for coloured supersymmetric particle production at the LHC with threshold and Coulomb resummation*, JHEP **12** (2016) 133, arXiv: 1607.07741 [hep-ph].

[36] ATLAS Collaboration, *Electron reconstruction and identification efficiency measurements with the ATLAS detector using the 2011 LHC proton–proton collision data*, Eur. Phys. J. C **74** (2014) 2941, arXiv: 1404.2240 [hep-ex].

[37] ATLAS Collaboration, *Electron efficiency measurements with the ATLAS detector using the 2015 LHC proton–proton collision data*, ATLAS-CONF-2016-024, 2016, URL: https://cds.cern.ch/record/2157687.

[38] ATLAS Collaboration, *Muon reconstruction performance of the ATLAS detector in proton–proton collision data at* $\sqrt{s}$ = 13 *TeV*, Eur. Phys. J. C **76** (2016) 292, arXiv: 1603.05598 [hep-ex].

[39] W. Lampl et al., *Calorimeter Clustering Algorithms: Description and Performance*, tech. rep. ATL-LARG-PUB-2008-002. ATL-COM-LARG-2008-003, CERN, 2008, URL: https://cds.cern.ch/record/1099735.

[40] ATLAS Collaboration, *Jet calibration and systematic uncertainties for jets reconstructed in the ATLAS detector at* $\sqrt{s}$ =*13 TeV*, ATLAS-PHYS-PUB-2015-015 (2015), URL: http://cdsweb.cern.ch/record/2037613.

[41] ATLAS Collaboration, *Monte Carlo Calibration and Combination of In-Situ Measurements of Jet Energy Scale, Jet Energy Resolution and Jet Mass in ATLAS*, ATLAS-CONF-2015-037 (2015), URL: http://cdsweb.cern.ch/record/2044941.

[42] ATLAS Collaboration, *Performance of pile-up mitigation techniques for jets in pp collisions at* $\sqrt{s}$ = 8 *TeV using the ATLAS detector*, (2015), arXiv: 1510.03823 [hep-ex].

[43] ATLAS Collaboration, *Selection of jets produced in 13 TeV proton-proton collisions with the ATLAS detector*, ATLAS-CONF-2015-029 (2015), URL: http://cdsweb.cern.ch/record/2037702.

[44] ATLAS Collaboration, *Optimisation of the ATLAS b-tagging performance for the 2016 LHC Run*, tech. rep. ATL-PHYS-PUB-2016-012, CERN, 2016, URL: https://cds.cern.ch/record/2160731.

[45] ATLAS Collaboration, *Commissioning of the ATLAS b-tagging algorithms using* $t\bar{t}$ *events in early Run-2 data*, ATLAS-PHYS-PUB-2015-039 (2015), URL: http://cdsweb.cern.ch/record/2047871.

[46] ATLAS Collaboration, *Search for the Standard Model Higgs boson decaying into* $b\bar{b}$ *produced in association with top quarks decaying hadronically in pp collisions at* $\sqrt{s}$ = 8 *TeV with the ATLAS detector*, JHEP **05** (2016) 160, arXiv: 1604.03812 [hep-ex].

[47] ATLAS Collaboration, *Search for four-top-quark production in the single-lepton and opposite-sign dilepton final states in pp collisions at* $\sqrt{s}$ = *13 TeV with the ATLAS detector*, Phys. Rev. **D99** (2019) 052009, arXiv: 1811.02305 [hep-ex].

[48] D0 Collaboration, *Measurement of the t anti-t production cross section in p anti-p collisions at* $\sqrt{s}$ = 1.96 *TeV using secondary vertex b tagging*, Phys. Rev. D **74** (2006) 112004, arXiv: hep-ex/0611002.

[49] M. Czakon and A. Mitov, *Top++: A program for the calculation of the top-pair cross-section at hadron colliders*, Comput. Phys. Commun. **185** (2014) 2930, arXiv: 1112.5675 [hep-ph].

[50] ATLAS Collaboration, *Search for the standard model Higgs boson produced in association with top quarks and decaying into a* $b\bar{b}$ *pair in pp collisions at* $\sqrt{s}$ = *13 TeV with the ATLAS detector*, Phys. Rev. **D97** (2018) 072016, arXiv: 1712.08895 [hep-ex].

[51] M. Bahr et al., *Herwig++ physics and manual*, Eur. Phys. J. C **58** (2008) 639, arXiv: 0803.0883 [hep-ph].

[52] J. Bellm et al., *Herwig 7.0/Herwig++ 3.0 release note*, Eur. Phys. J. C **76** (2016) 196, arXiv: 1512.01178 [hep-ph].

[53] L. Harland-Lang, A. Martin, P. Motylinski and R. Thorne, *Parton distributions in the LHC era: MMHT 2014 PDFs*, Eur. Phys. J. C **75** (2015) 204, arXiv: 1412.3989 [hep-ph].

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

[55] N. Kidonakis, *Two-loop soft anomalous dimensions for single top quark associated production with a W- or H-*, Phys. Rev. **D82** (2010) 054018, arXiv: 1005.4451 [hep-ph].

[56] N. Kidonakis, *NNLL resummation for s-channel single top quark production*, Phys. Rev. D **81** (2010) 054028, arXiv: 1001.5034 [hep-ph].

[57] W. Verkerke and D. Kirkby, *RooFit Users Manual*, http://roofit.sourceforge.net/.

[58] A. L. Read, *Presentation of search results: The CL(s) technique*, J. Phys. **G28** (2002) 2693.