ATLAS CONF Note

ATLAS-CONF-2023-043

17th August 2023

Search for new phenomena with top-quark pairs in final states with one lepton, jets and missing transverse momentum using 140 fb^{-1} of data at $\sqrt{s}$ = 13 TeV with the ATLAS detector

The ATLAS Collaboration

A search for new phenomena with top quark pairs in final states with one isolated lepton, either an electron or a muon, jets, and large missing transverse momentum is performed. The search is performed using the Large Hadron Collider proton–proton collision dataset at a centre-of-mass energy of $\sqrt{s}$ = 13 TeV recorded by the ATLAS detector and corresponding to an integrated luminosity of 140 fb^{-1}. Event classifiers based on neural networks are optimised to search for directly produced pairs of supersymmetric partners of the top quark (stop), as well as to search for spin-0 mediators produced in association with a pair of top quarks and decaying into dark-matter particles. A novel analysis approach is developed to achieve high sensitivity over a wide region of the parameter space of the tested signal models. No significant excess above the Standard Model background is observed, and limits at 95% confidence level are set in the stop–neutralino mass plane and as a function of the mediator mass or the dark-matter particle mass. Models with neutralinos from the direct stop decay with masses up to 600 GeV are excluded, while at small neutralino masses models are excluded for stop masses up to 1090 GeV. Scalar (pseudoscalar) dark matter mediator masses as large as 250 (300) GeV are excluded when the coupling strengths of the mediator to Standard Model and dark-matter particles are both equal to one. At lower mediator masses, models with production cross sections as small as 0.22 (0.26) times the nominal predictions are excluded.

ATLAS-CONF-2023-043
04 September 2023

1 Introduction

Being the heaviest known particles in the Standard Model (SM), top quarks are thought to be key probes for testing the existence of new particles in several extensions of the SM. Prominent examples of such extensions are Supersymmetry (SUSY) [1–7] and models with Dark Matter (DM) particles which interact with top quarks via a spin-0 mediator [8, 9]. In the proton–proton (pp) collisions at the Large Hadron Collider (LHC), both scenarios may lead to final states with a pair of top quarks and missing transverse momentum produced by undetected new particles ($t\bar{t} + E_{\mathrm{T}}^{\mathrm{miss}}$). This final state is investigated in this note in the search for new phenomena. Analysed events feature exactly one isolated charged lepton (electron or muon,[1] henceforth referred to as 'lepton' or 'ℓ') from the decay of an on-shell W boson, jets, and a significant amount of missing transverse momentum ($\vec{p}_T^{\,\mathrm{miss}}$), the magnitude of which is denoted by $E_{\mathrm{T}}^{\mathrm{miss}}$.

Since the discovery of the Higgs boson at the LHC [10, 11], SUSY has been searched for with particular interest as it may offer a solution to the hierarchy problem [12–15]. SUSY extends the SM by introducing a supersymmetric partner for each SM particle, the two having identical quantum numbers except for a half-unit difference in spin. In this theory, the hierarchy problem may be solved by the presence of a light supersymmetric partner of the top quark, referred to as top squark or 'stop', which largely cancels out divergent loop corrections to the Higgs boson mass [16–23]. The superpartners of the left- and right-handed top quarks mix to form the mass eigenstates $\tilde{t}_1$ and $\tilde{t}_2$, with $\tilde{t}_1$ being the lighter. A generic R-parity-conserving[2] minimal SUSY extension of the SM (MSSM) [7, 16, 24–26] predicts pair production of SUSY particles and the existence of a stable lightest SUSY particle (LSP). In the MSSM models considered in this search the LSP is assumed to be the lightest mass eigenstate of the superpartners of the Higgs and neutral electroweak gauge bosons. This LSP, also referred to as the lightest neutralino or $\tilde{\chi}_1^0$, may be a DM candidate as it is stable and only interacts weakly with SM particles [27, 28]. This note presents a search for direct pair production of $\tilde{t}_1$ particles ($\tilde{t}_1\tilde{t}_1$) in events with significant amount of $E_{\mathrm{T}}^{\mathrm{miss}}$ from the two weakly interacting LSPs escaping detection and the decay products of two on- or off-shell top quarks. The search for a spin-0 mediator produced in association with top quarks and subsequently decaying into a pair of DM particles ($t\bar{t}$+DM) is motivated by SM extensions which respect the principle of minimal flavour violation resulting in the interaction strength between the spin-0 mediator and the SM quarks being proportional to the quark masses via Yukawa-like couplings [29].

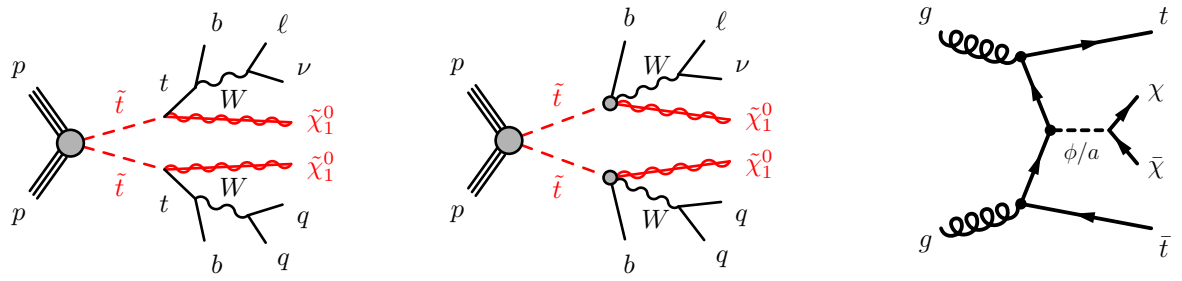
Searches for direct $\tilde{t}_1$ pair production and for spin-0 mediators produced in associations with top quarks and decaying into DM particles were reported by the ATLAS [30–33] and CMS [34–37] collaborations using the full LHC Run-2 pp recorded data set. This note presents an improved search in final states with exactly one lepton which have been already analysed by ATLAS [31] with the same data set. This search employs improved object reconstruction and identification algorithms, improved background simulations, and new neural-network-based classifiers for the reconstruction of the hadronically-decaying top quark and for the event discrimination.

2 Signal models and search strategy

Leading-order diagrams of the signal processes targetted in this search are shown in Fig. 1.

[1] Electrons and muons from τ-lepton decays are included.

[2] A multiplicative quantum number, referred to as R-parity, is introduced to conserve baryon and lepton numbers, where R-parity is $1(-1)$ for all SM (SUSY) particles.



(a) $\tilde{t}_1\tilde{t}_1$ in two-body decays (b) $\tilde{t}_1\tilde{t}_1$ in three-body decays (c) $t\bar{t}$ + DM

Figure 1: Diagrams illustrating direct $\tilde{t}_1$ pair production with (a) two-body ($\tilde{t}_1 \to t\tilde{\chi}_1^0$) or (b) three-body ($\tilde{t}_1 \to bW\tilde{\chi}_1^0$) decays and (c) the production of DM fermions (χ) in association with a pair of top quarks via a (pseudo) scalar mediator ϕ (a). In the SUSY diagrams, the charge-conjugate symbols are omitted for simplicity.

The $\tilde{t}_1\tilde{t}_1$ production is considered in a simplified SUSY model [38, 39] where the only light sparticles are $\tilde{t}_1$ and $\tilde{\chi}_1^0$. Each stop can decay promptly either via the two-body process $\tilde{t}_1 \to t\tilde{\chi}_1^0$ producing an on-shell top quark or via a three-body decay $\tilde{t}_1 \to bW\tilde{\chi}_1^0$ with an intermediate off-shell top quark. These are assumed to be the only possible stop decays depending on the mass splitting $\Delta m(\tilde{t}_1, \tilde{\chi}_1^0) = m(\tilde{t}_1) - m(\tilde{\chi}_1^0)$. The two-body decays are allowed for $\Delta m(\tilde{t}_1, \tilde{\chi}_1^0) > m(t)$, while three-body decays are allowed for $m(W) < \Delta m(\tilde{t}_1, \tilde{\chi}_1^0) < m(t)$. Decays at smaller $\Delta m(\tilde{t}_1, \tilde{\chi}_1^0)$ values, where both the top quark and the W boson are off-shell have been explored in Ref. [31]. In this search, the $\tilde{t}_1\tilde{t}_1$ production is tested with a two-dimensional scan of the parameter space $(m(\tilde{t}_1), m(\tilde{\chi}_1^0))$.

The simplified benchmark model for the $t\bar{t}$+DM production [29, 40] assumes the existence of a (pseudo)scalar mediator ϕ (a) which can be produced in association with two top quarks and decay into a pair of SM-singlet DM particles (χ). The interactions of the mediator with SM particles are Yukawa-like. This model has four parameters: the mass of the mediator $m(\phi/a)$, the mass of the DM particle $m(\chi)$, the strength of the DM–mediator interaction g_χ, and a universal scaling factor of the interactions of the mediator with SM fermions g_q. To reduce the number of free parameters, g_χ and g_q are assumed to have the same value $g = g_\chi = g_q$. Results are here presented as one dimensional scans of either $m(\phi/a)$ or $m(\chi)$. The $t\bar{t} + E_{\mathrm{T}}^{\mathrm{miss}}$ final state is mainly sensitive to the mediator mass range $m(\phi/a) < 2m(t)$, where the $\phi/a \to t\bar{t}$ decay is kinematically not allowed. Within the same simplified model, DM particles may also be produced in association with a top quark and a W boson (tW+DM). This production mode is expected to yield a small additional contribution in the mass range $m(\phi/a) < 2m(t)$. It is not considered in this search, but it was studied in Ref. [33].

The search strategy is designed to retain sensitivity to a wide region of the signal parameter space by partitioning the $t\bar{t} + E_{\mathrm{T}}^{\mathrm{miss}}$ final state into orthogonal and inclusive event categories and by classifying events with neural networks (NNs). Event categories differ in the multiplicity of jets with b-hadrons and in the presence or lack of jets produced by hadronic decays of high-p_{T} top quarks. Events from signals in different regions of the parameter space populate these categories at different rates. For instance, events with a pair of top squarks with high mass are predominantly accepted in categories with high-p_{T} top quarks. Moreover, events accepted in the same category from different signal models feature similar kinematic properties. Background processes also populate event categories at different rates, leading to a varying background composition across categories. These features are exploited by employing several NNs to classify signal from background events. In each event category, one dedicated NN is trained with signal

events from the full parameter space accepted in that category. With this analysis strategy, a large fraction of signal events is retained and a good discovery sensitivity is achieved via a template fit to the data binned as a function of the NN output values. As no category or classifier is optimised for a specific signal model, the same template fit to data is sufficient to determine the presence of signal events from any point of the parameter space.

The most important background processes for this search are top-quark pair production ($t\bar{t}$), the associated production of a top quark and a W boson (tW), the W+jets production and the production of a Z boson in association with a pair of top quarks and decaying into neutrinos ($t\bar{t}Z(\rightarrow \nu\nu)$). Except for the irreducible $t\bar{t}Z$ background, which has a small production cross section and cannot be easily discriminated from signal, all other major backgrounds are determined in situ in each event category with events classified at low NN output values. The resulting background model is then validated with events classified at an intermediate range in NN output values where no significant potential signal contribution is expected. Each event category is therefore split into three regions via orthogonal selections on the NN output value: a Control Region (CR) at low values, a Validation Region (VR) at intermediate values, and a Signal Region (SR) at high values. A combined binned fit to the data in all CRs and SRs is then used to assess the presence of signal events. In this fit, the distributions of the observed data and of the expected backgrounds remain the same, while the signal templates change according to the point of the parameter space under test. Dedicated NNs are used for the search for top squarks and for the search for DM particles.

3 ATLAS detector and data collection

The ATLAS experiment [41] at the LHC uses a multipurpose particle detector with a forward–backward symmetric cylindrical geometry and a near 4π coverage in solid angle.[3] The detector consists of an inner tracking detector (ID) surrounded by a thin superconducting solenoid providing a 2 T axial magnetic field, electromagnetic (EM) and hadron calorimeters, and a muon spectrometer (MS). The ID covers the pseudorapidity range $|\eta| < 2.5$. The high-granularity silicon pixel detector covers the vertex region and typically provides four measurements per track, the first hit normally being in the insertable B-layer (IBL) installed before Run 2 [42, 43]. It is surrounded by a silicon microstrip detector and a straw tube transition-radiation tracking detector. Lead/liquid-argon (LAr) sampling calorimeters provide EM energy measurements with high granularity. A steel/scintillator-tile hadron calorimeter covers the central pseudorapidity range ($|\eta| < 1.7$). The endcap and forward regions are instrumented with LAr calorimeters for both the EM and hadronic energy measurements up to $|\eta| = 4.9$. The MS surrounds the calorimeters and is based on three large superconducting air-core toroidal 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 MS includes a system of precision tracking chambers and fast detectors for triggering. A two-level trigger system is used to select events [44]. The first-level trigger is implemented in hardware and uses a subset of the detector information to accept events at a rate below 100 kHz. This is followed by a software-based trigger that reduces the accepted event rate to 1 kHz on average depending on the data-taking conditions. An extensive software suite [45] is used in data simulation, in the reconstruction and analysis of real and simulated data, in detector operations, and in the trigger and data acquisition systems of the experiment.

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

The analysed data set consists of the full Run-2 pp data set delivered by the LHC at $\sqrt{s} = 13$ TeV in the period between 2015 and 2018. After the application of data-quality requirements [46], the total integrated luminosity is $(140.1 \pm 1.2)\,\text{fb}^{-1}$. The uncertainty of 0.83% [47] is obtained using the LUCID-2 detector [48] for the primary luminosity measurements, complemented by measurements using the inner detector and calorimeters. All analysed events were recorded by triggers that accepted events with large $E_{\text{T}}^{\text{miss}}$ [49] or with an electron [50] or muon [51]. The primary vertex in these events, defined as the reconstructed vertex with the highest $\sum_{\text{tracks}} p_{\text{T}}^2$, must have at least two associated tracks with $p_{\text{T}} > 500\,\text{MeV}$.

4 Simulated events

Samples of Monte Carlo (MC) simulated events are used for the description of the SM processes, as well as of the expected signals. Details of the simulation samples used, including the matrix element (ME) event generator and its accuracy in the strong coupling constant (QCD accuracy), the parton distribution function (PDF) set, the parton shower (PS) and hadronisation model, and the accuracy of the cross-section calculation, are summarised in Table 1.

Table 1: Overview of the configurations used to simulate signal and background processes.

Process	ME event generator	ME QCD accuracy	ME PDF	Parton shower and hadronisation	Cross-section calculation
SUSY signals	MadGraph 2.8.1, 2.9.9 [52]	0,1,2j@LO	NNPDF2.3lo	Pythia 8.240, 8.307	NNLO+NNLL [53–57]
DM signals	MadGraph 2.7.3	0,1j@LO	NNPDF2.3lo	Pythia 8.244	NLO [58, 59]
$t\bar{t}$	Sherpa 2.2.12 [60]	0,1j@NLO +2,3,4j@LO [74–77]	NNPDF3.0nnlo [61]	Sherpa [62–66]	NNLO+NNLL [67–73]
Single-top					
tW	Powheg Box v2 [78–81]	NLO	NNPDF3.0nlo	Pythia 8.307 [82]	NLO+NNLL [83, 84]
s- and t-channel	Powheg Box v2	NLO	NNPDF3.0nlo	Pythia 8.230	NLO [85, 86]
V+jets ($V = Z, W$)	Sherpa 2.2.11	0,1j@NLO +2,3,4j@LO	NNPDF3.0nnlo	Sherpa	NNLO [87]
$t\bar{t}V$	MadGraph5_aMC@NLO 2.3.3 [52]	NLO	NNPDF3.0nlo	Pythia 8.210	NLO QCD+EW [88]
VV'	Sherpa 2.2.1, 2.2.2	0,1j@NLO+2,3j@LO	NNPDF3.0nnlo	Sherpa	

The samples produced with MadGraph5_aMC@NLO [52] and Powheg Box [78–81] used EvtGen [89] for the modelling of b-hadron decays. In these samples, Pythia [82] was used for the parton shower and hadronisation with the A14 tune [90] and the NNPDF2.3lo set of PDFs [91]. In all samples, the effect of multiple interactions in the same and neighbouring bunch crossings (pile-up) was modelled by overlaying the simulated hard-scattering event with inelastic pp events generated with Pythia 8.186 [92] using the NNPDF2.3lo set of PDFs and the A3 set of tuned parameters [93]. All samples were processed with the full simulation of the ATLAS detector [94] based on Geant4 [95].

SUSY samples were generated with MadGraph [52] at leading order (LO) with up to two additional partons. The stop decays were simulated with MadSpin [96] which emulates kinematic distributions to a good approximation without calculating the full ME. The simplified SUSY model was the same used in Ref. [31]. The $\tilde{t}_1\tilde{t}_1$ production cross-sections were calculated to approximate next-to-next-to-leading order in QCD, adding the resummation of soft gluon emission at next-to-next-to-leading-logarithm accuracy (approximate NNLO+NNLL) [53–57]. The nominal cross-section and its uncertainty were derived using the PDF4LHC15_mc PDF set, following the recommendations of Ref. [97].

DM samples were generated with $g = g_\chi = g_q = 1$ using MadGraph at LO with up to one additional parton and with MadSpin for the decay of top quarks. The kinematics of the mediator decay were found

to not depend strongly on the values of the couplings; however, the particle kinematic distributions are sensitive to the scalar or pseudoscalar nature of the mediator and to the mediator and DM particle masses. The cross-sections were computed at next-to-leading order (NLO) in QCD [58, 59].

The productions of $t\bar{t}$ and of W+jets were simulated with the Sherpa 2.2.12 and Sherpa 2.2.11 [60] generators, respectively, using NLO MEs for up to two additional partons, and LO MEs for up to four additional partons calculated with the Comix [74] and OpenLoops [75–77] libraries. They were matched with the Sherpa parton shower [62] using the MEPS@NLO prescription [63–66] using the set of tuned parameters developed by the Sherpa authors. Production of the tW was modelled by the Powheg Box v2 [78–81] generator at NLO in QCD using the five-flavour scheme. The diagram removal scheme [98] was used to remove interference and overlap with $t\bar{t}$ production. The uncertainty related to this scheme was estimated by comparison with an alternative sample generated using the diagram subtraction scheme [98, 99]. The production of a gauge boson with a top quark pair ($t\bar{t}W$ and $t\bar{t}Z$) was modelled using the MadGraph5_aMC@NLO 2.3.3 [52] generator at NLO. Other sources of background events that are considered in this search are diboson production (VV', where $V = W, Z$), Higgs boson production in association with a top quark pair, tZ and tWZ productions, and the production of three and four top quarks.

Further details on the simulation are provided in Table 1. One main difference from the simulations used in Ref. [31] is the use of Sherpa instead of Powheg Box v2 for the $t\bar{t}$ production as it provides an improved modelling of events at high transverse momentum which is relevant for this search. A second difference is the use of dynamic rather than fixed scales in the MEs used for the simulation of the tW production.

5 Event reconstruction

The event categorisation and discrimination is based on reconstructed objects in the event, leptons and jets, including those containing b-hadrons, and on the measured $E_{\mathrm{T}}^{\mathrm{miss}}$. Depending on the quality and kinematic requirements, reconstructed electron or muons are labelled as either *baseline* or *signal*, where the latter is a subset of the former, passing tighter selection criteria. Baseline leptons are used when classifying overlapping objects, to compute the missing transverse momentum, and for background rejection. Signal leptons are used to construct discriminating variables. The momenta of hadronically-decaying top quarks are reconstructed via dedicated algorithms.

Electron candidates are reconstructed from an isolated electromagnetic calorimeter energy deposit which is matched to a track in the ID [100]. The pseudorapidity of the calorimeter energy cluster must satisfy $|\eta_{\mathrm{cluster}}| < 2.47$. Baseline electrons are required to have $p_{\mathrm{T}} > 4.5\,\mathrm{GeV}$ and to satisfy loose likelihood identification criteria. Furthermore, their longitudinal impact parameter (z_0), defined as the distance along the beam direction between the primary vertex and the track's point of closest approach to the beam axis, must satisfy $|z_0 \sin\theta| < 0.5\,\mathrm{mm}$, where θ is the polar angle of the track. Signal electrons must satisfy all the baseline requirements, have $p_{\mathrm{T}} > 20\,\mathrm{GeV}$ and have a transverse impact parameter (d_0) with $|d_0|/\sigma_{d_0} < 5$, where σ_{d_0} is the uncertainty in d_0. Signal electrons are also required to be isolated and satisfy tighter identification criteria. The isolation is defined as the sum of the transverse energy or momentum reconstructed in a cone of size $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$ around the electron, excluding the energy of the electron itself. The isolation criteria depend on the electron's p_{T}.

Muon candidates are reconstructed by combining tracks in the ID with tracks in the MS or as stand-alone MS tracks [101, 102] and are required to have $|\eta| < 2.7$. Baseline muons are required to have $p_{\mathrm{T}} > 4\,\mathrm{GeV}$,

$|z_0 \sin\theta| < 0.5\,\text{mm}$, and to satisfy the 'Medium' identification criterion. In addition to all baseline requirements, signal muons must have $p_{\text{T}} > 20\,\text{GeV}$ and $|d_0|/\sigma_{d_0} < 3$. Signal muons must also be isolated, based on criteria similar to those used for electrons.

Particle-flow (PFlow) jets within $|\eta| < 4.5$ are reconstructed using the anti-k_t algorithm [103] with a radius parameter $R = 0.4$ [104], using neutral PFlow constituents and charged constituents associated with the primary vertex as input [105]. They are referred to as 'small-R jets' hereafter. These jets are then calibrated to the particle level by applying a jet energy scale derived from simulation with in situ corrections based on collected data [106]. Jets with $p_{\text{T}} > 20\,\text{GeV}$ are selected. A cleaning procedure is used to identify and remove jets arising from calorimeter noise or non-collision backgrounds. To suppress pile-up jets within $|\eta| < 2.4$, a discriminant called the 'jet vertex tagger' (JVT) is constructed using a likelihood method [107]. A similar discriminant, the 'forward JVT' (fJVT), is used for jets with $|\eta| > 2.5$ [108]. PFlow jets within $|\eta| < 2.5$ identified as containing b-hadrons are referred to as b-tagged jets. Their identification is based on the DL1r b-tagging algorithm [109, 110] with a working point corresponding to a b-tagging efficiency of 77%, measured in a sample of simulated $t\bar{t}$ events. The corresponding rejection factors are approximately 200 and 6 for light-quark and gluon jets and c-jets, respectively.

To reconstruct high-p_{T} hadronic top quark decays, so-called 'large-R jets' are used [111]. These jets are reconstructed using the anti-k_t algorithm with $R = 1.0$ using topological clusters of calorimeter cells calibrated to the hadronic scale [112]. These jets are trimmed to remove energy deposits from pile-up and the underlying event [113]. The energy and the mass of the large-R jets are calibrated with data [114, 115]. Large-R jets in $|\eta| < 2$ are selected if they have a transverse momentum in the range $[600, 2500]$ GeV and a mass in the range $[40, 600]$ GeV. Large-R jets with $p_{\text{T}} < 600\,\text{GeV}$ are not considered, as low-p_{T} hadronic top quark decays are reconstructed from small-R jets, as described in Section 5.1. A multivariate classifier is used to identify the large-R jets from hadronic top decays using a working point with a top-tagging efficiency of 80% [116, 117]. This classifier uses jet substructure observables as input. To improve the identification of hadronic top decays, jets containing b-hadrons are reconstructed inside the large-R jet using a track-based jet reconstruction with a variable cone size [118] and are tagged with the DL1r b-tagging algorithm.

The reconstruction of jets from hadronic decays of τ-leptons ($\tau_{\text{had-vis}}$) [119, 120] is seeded from jets reconstructed from topological clusters by the anti-k_t algorithm with $R = 0.4$. Tracks originating from the primary vertex and satisfying impact parameter requirements, as well as selections based on a multivariate discriminant, are associated to the $\tau_{\text{had-vis}}$ if they are within a cone of $\Delta R < 0.2$ around the jet axis. Reconstructed $\tau_{\text{had-vis}}$ are selected if they have one or three associated tracks with a total charge equal to ± 1. They also need to have $p_{\text{T}} > 20\,\text{GeV}$ and to be reconstructed in the pseudorapidity ranges $|\eta| < 1.37$ or $1.52 < |\eta| < 2.5$. Furthermore, reconstructed $\tau_{\text{had-vis}}$ need to satisfy loose identification criteria aimed at rejecting jets originating from quarks, gluons and electrons.

The missing transverse momentum $\vec{p}_T^{\,\text{miss}}$ [121] is reconstructed as the negative vector sum of the p_{T} of all the selected electrons, muons and small-R jets. An extra track-based 'soft term' is built using additional tracks associated with the primary vertex, but not with any reconstructed object, to improve the performance of the $\vec{p}_T^{\,\text{miss}}$ reconstruction at high pile-up. An object-based missing transverse momentum significance [122] is used to identify events in which the reconstructed $\vec{p}_T^{\,\text{miss}}$ comes from undetected or weakly interacting particles rather than from detected particles with mis-measured energies.

Correction factors are applied to reconstructed objects to take account of differences in trigger, reconstruction, identification and isolation efficiencies, as well as differences in energy scale and resolution, between data and simulation. Ambiguities among selected objects are resolved via an overlap removal procedure similar

to the one used in Ref. [31]. Given a set of baseline objects, the procedure checks for overlaps based on shared tracks, ghost-matching [123], or a minimum distance ΔR_y between pairs of objects.[4] Small-R jets with $\Delta R < 1.1$ from a large-R jet are discarded and no overlap removal is performed between leptons and large-R jets.

5.1 Resolved hadronic top decays

An NN-based algorithm, referred to as 'top-NN' hereafter, is used to identify multiplets of small-R jets associated to hadronic decays of top quarks out of the combinatorics of selected jets, which may contain also jets from additional radiation, pile-up and the underlying event. This algorithm targets resolved decays in which the top quark p_{T} is such that the resulting jets are spatially separated and can be reconstructed as individual small-R jets. The target p_{T} range is below the one for which the top quark decay can be reconstructed as a large-R jet. The purpose of the top-NN is the reconstruction of the momentum of the hadronically-decaying top quark in the event, which is later used for the event classification.

The top-NN evaluates all pairs and triplets of jets in an event. While a fully reconstructed resolved decay would be associated to a triplet of jets, pairs are also considered to allow for cases in which one jet is outside acceptance or for cases in which the two jets from a high-p_{T} W boson are reconstructed as a single small-R jet. Pairs and triplets are built from combinations of up to two b-tagged jets and up to four not-b-tagged jets, referred to as 'light jets' hereafter, in the event. Each multiplet must contain exactly one b-tagged jet. The top-NN takes as input the four momentum components of each jet in the multiplet after transformations which remove symmetries and reduce the dimensionality of the inputs. These transformations include a Lorentz boost into the multiplet rest frame and a rotation such that the b-tagged jet is aligned along the z-axis and the leading-p_{T} light jet lays on the xz-plane with a positive p_x component. After these transformations, six non-trivial jet momentum components together with the original multiplet p_{T} are used as NN inputs. The top-NN is a binary feed-forward NN trained with Keras [124] with the Tensorflow backend. The top-NN is trained with SUSY signal events from all simulated points of the parameter space in order to cover a large spectrum in top quark p_{T}. The signal multiplets are those with jets kinematically matched to the partons from the top quark decays, while all other multiplets from the combinatorics of jets are background multiplets. In the training, weights are assigned such that both signal and background multiplets are uniformly distributed in p_{T}. In each event, the top-NN evaluates all multiplets out of all combinations of selected jets and assigns an output value to each of them. This output value is related to the likelihood of a multiplet to be associated with a hadronic decay of a top quark. The multiplet with the highest output value in each event is chosen as the hadronic top quark candidate. According to simulation, the top-NN reaches the highest efficiency in finding the right multiplet in the range of top quark p_{T} between 200 and 600 GeV. In this p_{T} range, the top-NN identifies the correct multiplet in about 70% of the events in which a b-tagged jet, correctly tagged, and at least one light jet are kinematically matched to the hadronically decaying top quark. Above this range, the performance worsens as the decay products are less likely to be reconstructed as individual small-R jets. These decays are better reconstructed with large-R jets. Therefore, the top-NN is used to reconstruct the hadronic top candidate only in events without a large-R jet with $p_{\mathrm{T}} > 600$ GeV. These events are required to have the multiplet with highest output value above a minimum threshold. This selection rejects events without any multiplet which resembles a hadronically-decaying top quark while retaining close to all events with a real hadronically-decaying top quark. This loose selection ensures that a good hadronic top quark candidate is reconstructed and can be used for the event classification.

[4] Rapidity $y \equiv \frac{1}{2} \ln \frac{E+p_z}{E-p_z}$ is used instead of pseudorapidity (η) when computing the distance ΔR_y.

6 Event classification

Events are collected by $E_{\mathrm{T}}^{\mathrm{miss}}$ and single lepton triggers and different thresholds on the reconstructed $E_{\mathrm{T}}^{\mathrm{miss}}$ and signal leptons are applied in order to select events in the kinematic range where these triggers reach high efficiency. Reconstructed events with $E_{\mathrm{T}}^{\mathrm{miss}} > 230\,\mathrm{GeV}$ are required to have been collected by the $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger. Events with $E_{\mathrm{T}}^{\mathrm{miss}}$ in the range $[70, 230]$ GeV are instead required to have been collected by a single lepton trigger. For these events, the signal electron (muon) associated to the lepton selected at trigger level must have $p_{\mathrm{T}} > 27\,(27.5)$ GeV and $|\eta| < 2.47(2.4)$. All events are also required to have exactly one or three signal leptons and no additional baseline lepton or $\tau_{\mathrm{had-vis}}$. Events with three signal leptons are only used for the validation of the $t\bar{t}Z$ background.

Events are classified into orthogonal categories: the high-$E_{\mathrm{T}}^{\mathrm{miss}}$ events with $E_{\mathrm{T}}^{\mathrm{miss}} > 230\,\mathrm{GeV}$ and the 'boosted' events with a large-R jet. These categories target events with top quark decays at different transverse momenta. In case of ambiguity, the selection for boosted events takes priority. Events are then further categorised according to the number of b-tagged jets and the presence of a top-tagged large-R jet. The selections for these categories are summarised in Table 2. Events not accepted in any of these categories are not further used in the analysis.

Table 2: Summary of the selections for each event category together with the definitions of the hadronic and leptonic top quark candidates. Trigger, lepton and $E_{\mathrm{T}}^{\mathrm{miss}}$ selections are described in the text. Large-R jets are referred to as 'LR' for brevity.

Analysis Category	High-$E_{\mathrm{T}}^{\mathrm{miss}}$		Boosted					
	1b	2b	1b-lep-0t	1b-had-0t	2b-0t	1b-lep-1t	1b-had-1t	2b-1t
N(LR jet)	0		≥ 1					
N(top-tagged LR jet)	-		0			≥ 1		
$N_{b-\mathrm{jet}}$ with $\Delta R(b, \mathrm{LR\ jet}) < 1.1$	-		0	≥ 1	≥ 1	0	≥ 1	≥ 1
$N_{b-\mathrm{jet}}$ with $\Delta R(b, \mathrm{LR\ jet}) > 1.1$	-		≥ 1	0	≥ 1	≥ 1	0	≥ 1
top-NN-tagged multiplet	✓		-					
$N_{b-\mathrm{jet}}$	1	≥ 2	-					
$N_{\mathrm{light-jet}}$	≥ 2	≥ 1	-					
$\mathrm{top}_{\mathrm{had}}$ candidate	top-NN multiplet		LR jet					
$\mathrm{top}_{\mathrm{lep}}$ candidate	ℓ + j	$\ell + b$	$\ell + b$	ℓ(+j)	$\ell + b$	$\ell + b$	ℓ(+j)	$\ell + b$
Event NN selection	See Table 3							

High-$E_{\mathrm{T}}^{\mathrm{miss}}$ events are categorised into '1b' and '2b' events according to the number of b-tagged jets. A minimum number of light jets is required such that a top-NN-tagged multiplet can be defined by at least a pair of jets and an additional jet can be used for the reconstruction of the leptonically-decaying top quark. In the 2b (1b) categories, the hadronic top quark candidate is the vectorial sum of the jets of the multiplet with highest top-NN output value, while the leptonic top quark candidate is the vectorial sum of the momenta of the lepton and the b-tagged (light) leading-p_{T} jet that is not associated to the hadronic top quark candidate. In the 1b category, the leading-p_{T} jet is used to build the leptonic top quark candidate to retain a good kinematic description of the events in which the jet from the b-quark in the leptonic top quark decay is not correctly tagged.

Boosted events must have at least one large-R jet. In case of more, only the leading-p_{T} one is considered for the event selection and classification. These events are first categorised according to whether this large-R jet is top-tagged ('1t') or not ('0t') and then according to the numbers of b-tagged jets inside and outside the

large-R jet. The '2b' events have at least one b-tagged jet inside the large-R jet ($\Delta R(b, \text{large}-R \text{ jet}) < 1.1$) and at least one b-tagged jet outside. The '1b-had' events have at least one b-tagged jet inside the large-R jet and none outside. Finally, the '1b-lep' events have at least one b-tagged jet outside the large-R jet and none inside. This classification defines six orthogonal boosted categories. In these events, the hadronic top quark candidate is defined by the large-R jet momentum, while the leptonic top quark candidate is the vectorial sum of the momenta of the lepton and the leading-p_{T} b-tagged jet outside the large-R jet. In case this b-tagged jet is not present, the leading-p_{T} light jet, if present, is used.

Overall, eight orthogonal categories (two high-$E_{\mathrm{T}}^{\mathrm{miss}}$ and six boosted categories) in different kinematic regimes and with different reconstructed objects are used to maximise the signal acceptance. In each category, an NN-based event classifier either for the $\tilde{t}_1\tilde{t}_1$ search ('stop-NN') or for the $t\bar{t}$+DM search ('DM-NN') is employed to discriminate signal from background events. In the boosted categories, only stop-NNs are used since these categories are not expected to yield a sizeable fraction of $t\bar{t}$+DM events. Each NN is trained with all simulated events accepted in a given category from background processes and signals from across the parameter space. For the stop-NNs, only signal events with two-body decays are used in the training. For the DM-NNs, signal events with both scalar and pseudoscalar mediators are used together. All events are weighted by their predicted cross section. The distribution of the output value of either the stop-NN or the DM-NN in each category is then used to fit the observed data and assess the presence of signal events across the parameter space. As signals with different parameters populate event categories at different rates, events accepted in the same category have similar kinematic properties, but are enhanced with events from signals in specific regions in the parameter space. For instance, $\tilde{t}_1\tilde{t}_1$ events with large $\Delta m(\tilde{t}_1, \tilde{\chi}_1^0)$ are accepted at similar rates in the boosted and high-$E_{\mathrm{T}}^{\mathrm{miss}}$ categories, while events with small $\Delta m(\tilde{t}_1, \tilde{\chi}_1^0)$ are accepted almost exclusively in the high-$E_{\mathrm{T}}^{\mathrm{miss}}$ categories. When compared to more complex approaches with NNs trained for specific sets of signals, this approach retains high classification power across the parameter space with the simplicity of having only one distribution per event category to fit.

The stop- and DM-NNs are binary feed-forward NNs. In each event category the sets of inputs are the same for the stop- and DM-NNs, but these sets differ across categories. A set of 'low-level' inputs is used by all NNs in all categories. It includes the momentum components of the hadronic and leptonic top quark candidates, the momentum components of $\vec{p}_T^{\,\mathrm{miss}}$, and the momentum components of the lepton and of the b-tagged jets associated to the two top quark candidates.[5] Transformations on the plane transverse to the beam direction are used to remove rotational symmetries and reduce the input dimensionality. Additional 'high-level' variables include the $E_{\mathrm{T}}^{\mathrm{miss}}$ significance, the top-NN output value (only for the high-$E_{\mathrm{T}}^{\mathrm{miss}}$ categories), and kinematic variables such as the transverse mass $m_T(\vec{p}_T^{\,\ell}, \vec{p}_T^{\,\mathrm{miss}}) = \sqrt{2|\vec{p}_T^{\,\ell}||\vec{p}_T^{\,\mathrm{miss}}|(1-\cos\Delta\phi(\vec{p}_T^{\,\ell}, \vec{p}_T^{\,\mathrm{miss}}))}$ and the derived stransverse mass [125]:

$$m_{T2}(\vec{p}_T^{\,1}, \vec{p}_T^{\,2}, \vec{p}_T^{\,\mathrm{miss}}) = \min_{\vec{q}_T^{\,1}+\vec{q}_T^{\,2}=\vec{p}_T^{\,\mathrm{miss}}} \left\{\max\left[m_T(\vec{p}_T^{\,1}, \vec{q}_T^{\,1}), m_T(\vec{p}_T^{\,2}, \vec{q}_T^{\,2})\right]\right\} \ ,$$

where $\vec{p}_T^{\,i}$ is the transverse momentum of a visible particle and $\vec{q}_T^{\,i}$ is the transverse momentum of an undetected particle, assumed to be massless, which contributes to the $E_{\mathrm{T}}^{\mathrm{miss}}$. The variants of m_{T2} used in this search are $m_{T2}(b, b, E_{\mathrm{T}}^{\mathrm{miss}})$ and $m_{T2,\mathrm{min}}(b+\ell, b, E_{\mathrm{T}}^{\mathrm{miss}}) = \min\left[m_{T2}(b_1+\ell, b_2, E_{\mathrm{T}}^{\mathrm{miss}}), m_{T2}(b_1, b_2+\ell, E_{\mathrm{T}}^{\mathrm{miss}})\right]$, where b indicates the b-tagged (or light) jet associated to one of the two top quark candidates. The distributions of these variables are expected to have endpoints around $m(W)$ or $m(t)$ in $t\bar{t}$ events, while

[5] In the high-$E_{\mathrm{T}}^{\mathrm{miss}}$ 1b category, the momentum components of the light jet associated to the leptonic top candidate are used. In the boosted 1b-lep categories, the b-tagged jet associated to the hadronic top quark candidate is omitted.

reaching higher values in signal events with more undetected particles. The m_{T2} variables are not used as NN inputs in the boosted categories. In total, depending on the event category, the number of NN inputs ranges between 19 and 26.

When comparing the classification power among different classes of signals in each event category, SUSY signals at large stop masses are always the ones best classified and with output values close to unity, while signals with small $\Delta m(\tilde{t}_1, \tilde{\chi}_1^0)$ or in the three-body regime are worst classified. Among DM signals, events with a pseudoscalar mediators are always better classified than events with a scalar mediator for the same mediator mass. Moreover, when considering mediators with the same parity, signals with small $m(\phi/a)$ are classified less optimally than signals with large $m(\phi/a)$. To illustrate these features, Fig. 2 shows the the expected distributions of events as a function of the NN output value in the high-$E_{\mathrm{T}}^{\mathrm{miss}}$ 2b category for SM background processes as well as for classes of signal models. Given their different distributions, the analysis retains sensitivity to all classes of signal events by performing a binned fit to the data distributed as a function of the NN output value, as described in Section 7.

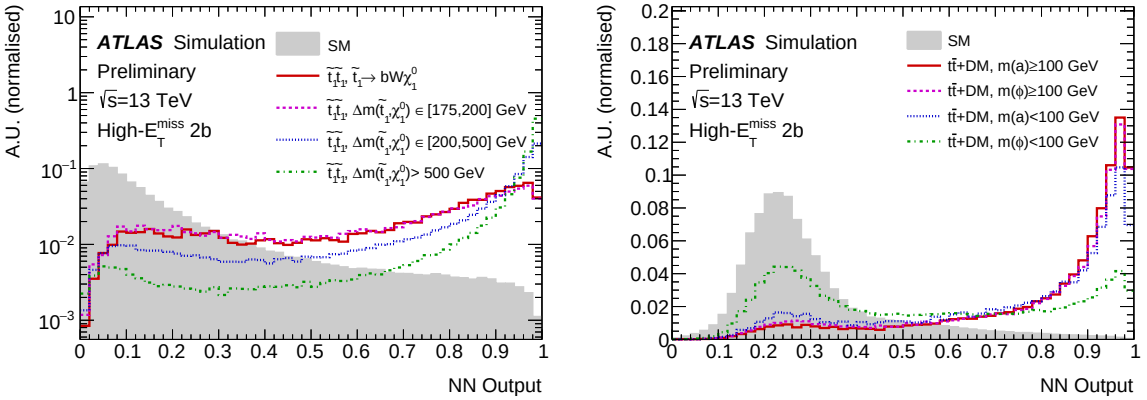


Figure 2: Expected distributions of events as a function of the output value of the stop-NN (left) and of the DM-NN (right) in the high-$E_{\mathrm{T}}^{\mathrm{miss}}$ 2b category. The distribution of the SM background processes is compared to classes of signal models. In the left plot, $\tilde{t}_1\tilde{t}_1$ models are grouped according to $\Delta m(\tilde{t}_1, \tilde{\chi}_1^0)$ and $\tilde{t}_1$ decay mode. In the right plot, $t\bar{t}$+DM models with a scalar ϕ or pseudoscalar a mediator are shown for mediator masses larger or smaller than 100 GeV and $m(\chi) = 1$ GeV. Distributions are normalised to the same integral.

The stop-NN and DM-NN output values are used to split each event category into an SR, a CR and a VR. The SR contains events at high output values with a potentially high purity of signal events. The CR contains events at low output values with negligible contributions from potential signals. These events are used to improve the modelling of the background normalisations and templates in situ in each event category. The VR contains events in the intermediate range which are used to check the validity of the background model derived at low output values when applied to background events at higher values. For each SR, the lower threshold on the NN output value is set based on its binned distribution and on the maximum of the signal-to-background ratio ($s/b_{\mathrm{bin}}^{\mathrm{max}}$) expected in each bin for all signals not excluded by Ref. [31].[6] As the $s/b_{\mathrm{bin}}^{\mathrm{max}}$ is monotonously increasing, the minimum threshold for the SR is chosen as the maximum output value below which all bins have $s/b_{\mathrm{bin}}^{\mathrm{max}} < 0.1$. The lower threshold for the VR is chosen in the same way with $s/b_{\mathrm{bin}}^{\mathrm{max}} < 0.05$, while ensuring an expected number of events suitable to validate the background model. Events with NN output values below the VR lower threshold are accepted in the CR. In

[6] For the DM signals, given the previously excluded cross sections, the expected $s/b_{\mathrm{bin}}^{\mathrm{max}}$ in each bin is evaluated assuming $g = 0.5$.

some categories, a lower threshold is set for the CR such that events below that threshold are rejected and not used for further analysis. These rejected events have background compositions not representative of the background events in the SR and may contain multi-jet events. With this approach, all CRs, VRs and SRs are orthogonal and span the full range of the NN output value above a minimum threshold. Table 3 reports a summary of these selections, together with the signal acceptance efficiencies in the SRs.

Table 3: Summary of the selections on the stop-NN and DM-NN output values which define CRs, VRs and SRs. Signal efficiency in SRs are also reported. They are computed as the fraction of signal events in a given category with a NN output value in the range accepted in the SR. The quoted range is based on efficiencies estimated for all signals across the simulated parameter space. In boosted categories, only efficiencies for $\tilde{t}_1\tilde{t}_1$ signals with $\Delta m(\tilde{t}_1, \tilde{\chi}_1^0) > 500\,\text{GeV}$ are considered.

Category	stop-NN				DM-NN			
	CR Range	VR Range	SR Range	Eff.	CR Range	VR Range	SR Range	Eff.
High-$E_{\text{T}}^{\text{miss}}$ 1b	$[0.2, 0.64[$	$[0.64, 0.79[$	$[0.79, 1.0]$	0.4-0.9	$[0.3, 0.69[$	$[0.69, 0.87[$	$[0.87, 1.0]$	0.3-0.4
High-$E_{\text{T}}^{\text{miss}}$ 2b	$[0.1, 0.56[$	$[0.56, 0.7[$	$[0.7, 1.0]$	0.5-0.9	$[0.3, 0.6[$	$[0.6, 0.76[$	$[0.76, 1.0]$	0.6-0.8
Boosted 1b-lep-1t	$[0, 0.65[$	$[0.65, 0.8[$	$[0.8, 1.0]$	0.5-0.9				
Boosted 1b-had-1t	$[0, 0.65[$	$[0.65, 0.85[$	$[0.85, 1.0]$	0.6-0.9				
Boosted 2b-1t	$[0, 0.75[$	$[0.75, 0.95[$	$[0.95, 1.0]$	0.6-0.8				
Boosted 1b-lep-0t	$[0, 0.7[$	$[0.7, 0.85[$	$[0.85, 1.0]$	0.6-0.8				
Boosted 1b-had-0t	$[0.1, 0.75[$	$[0.75, 0.95[$	$[0.95, 1.0]$	0.4-0.8				
Boosted 2b-0t	$[0, 0.65[$	$[0.65, 0.8[$	$[0.8, 1.0]$	0.6-0.9				

7 Background estimation and statistical model

Dominant sources of background events are processes producing $t\bar{t}$, single-top (including tW) and W+jets. Despite its small cross section, the $t\bar{t}Z(\rightarrow \nu\nu)$ production constitutes another important background as it yields events with signal-like kinematic properties. All these backgrounds, together with others with minor contributions, are estimated from simulation with corrections determined from the fit of their predictions to data. According to simulation, the fraction of background events in which the selected signal lepton is not a real prompt lepton from a W-boson or τ-lepton decay is small. These events are estimated from simulation with a conservative systematic uncertainty. Multi-jet events are expected to yield a negligible contribution in any category after the selection on the minimum NN output value and are therefore neglected in the background estimation.

Events from $t\bar{t}$ production are modelled via two classes: the semi-leptonic $t\bar{t}$ events with up to one prompt lepton ('$t\bar{t}$-1L') and the fully-leptonic $t\bar{t}$ events with at least two prompt leptons ('$t\bar{t}$-2L'). The $t\bar{t}$-1L events have the same final state as the signal events, but a lower $E_{\text{T}}^{\text{miss}}$ due to the presence of only one neutrino from W-boson decays. The $t\bar{t}$-2L events have an amount of $E_{\text{T}}^{\text{miss}}$ similar to that in signal events, but the final state is different as the hadronic top quark candidate is produced by additional jet activity in the event. Given the different kinematics and the higher susceptibility of the $t\bar{t}$-2L events to the modelling of additional radiation, the $t\bar{t}$-1L and the $t\bar{t}$-2L components are estimated separately. Their prediction is based on the same simulation, but their in-situ determinations are independent.

The predicted event yields from $t\bar{t}$, single-top and W+jets are corrected in situ by fitting the observed events in CRs, while testing the presence of signal events in the SRs. This fit is done with a binned maximum-likelihood approach using the statistical analysis packages RooFit [126], RooStats [127] and

HistFitter [128]. The statistical model is composed of observed and expected binned distributions in CRs and SRs. In SRs, events are binned by the output values of either the stop-NNs or the DM-NNs, depending on the type of signal being tested. In CRs, events are binned by m_{T} multiplied by the lepton electric charge ($m_{\mathrm{T}} \times q(\ell)$) as this variable allows the individual contributions of the $t\bar{t}$-1L, $t\bar{t}$-2L, and W+jets backgrounds to be determined simultaneously. The m_{T} variable is used because events with one leptonically-decaying W boson (eg, $t\bar{t}$-1L and W+jets events) are distributed differently from events with more (eg, $t\bar{t}$-2L) as the former feature an endpoint around m_W, while the latter extend to higher values. The lepton charge is used to discriminate between $t\bar{t}$-1L and W+jets events by exploiting the difference in production cross section for W^+ and W^- bosons. The contribution from single-top events is determined mainly in dedicated CRs. The boosted-2b-1t and boosted-2b-0t CRs are split with a selection on $m_{T2,\mathrm{min}}(b+\ell, b, E_{\mathrm{T}}^{\mathrm{miss}})$ at 300 GeV into 'low-m_{T2}' and 'high-m_{T2}' regions, with the high-m_{T2} region being enhanced in single-top events. Given the small number of events in the high-m_{T2} regions, only their total event yields are considered in the fit.

Overall, the statistical model used for the $\tilde{t}_1\tilde{t}_1$ search includes ten CRs and eight SRs, as summarised in Table 4. In the $t\bar{t}$+DM search, only the high-$E_{\mathrm{T}}^{\mathrm{miss}}$ SRs are considered since no sizeable signal yield is expected in the boosted SRs. Nonetheless, the boosted categories are still used to improve the determination of the single-top background. Therefore the statistical model for the $t\bar{t}$+DM includes the two high-$E_{\mathrm{T}}^{\mathrm{miss}}$ SRs with their associated CRs and also the eight boosted CRs, the same used in the $\tilde{t}_1\tilde{t}_1$ search. In these fits, the normalisations of the $t\bar{t}$-1L, $t\bar{t}$-2L, single-top and W+jets contributions are scaled by independent unconstrained fit parameters (normalisation factors or 'NFs'), as reported in Table 4. Via these NFs and the experimental and theoretical systematic uncertainties described in Section 8, the background predictions based on simulation are corrected in situ with the data in CRs and then used to estimate the background contributions in VRs and SRs. The samples of events in VRs are useful to validate the background predictions since they have background compositions similar to those in SRs and only small potential signal contributions. The observed and predicted distributions in CRs, VRs and SRs are shown in Section 9.

Table 4: Summary of the statistical model in terms of event categories, regions into which categories are divided, observables used in the template fit and normalisation factors ('NFs') applied to background contributions in each region. 'NN' stands for the NN output value. The background process on which these NFs act is indicated by the name of the NF. All ten CRs and eight SRs are used in the fit for the $\tilde{t}_1\tilde{t}_1$ search. The statistical model for the $t\bar{t}$+DM search differs as the SRs in the boosted categories are not used.

Category	Fitted observable		Normalisation Factors						
	CR	SR	$\mathrm{NF}_{\mathrm{top-1L}}^{\mathrm{High-met}}$	$\mathrm{NF}_{\mathrm{top-1L}}^{\mathrm{Boosted}}$	$\mathrm{NF}_{\mathrm{top-2L}}^{\mathrm{High-met}}$	$\mathrm{NF}_{\mathrm{top-2L}}^{\mathrm{Boosted}}$	$\mathrm{NF}_{W}^{\mathrm{High-met}}$	$\mathrm{NF}_{W}^{\mathrm{Boosted}}$	$\mathrm{NF}_{\mathrm{singletop}}$
High-$E_{\mathrm{T}}^{\mathrm{miss}}$ 1b	$m_{\mathrm{T}} \times q(\ell)$	NN	✓		✓		✓		✓
High-$E_{\mathrm{T}}^{\mathrm{miss}}$ 2b	$m_{\mathrm{T}} \times q(\ell)$	NN	✓		✓		✓		✓
Boosted 1b-lep-1t	$m_{\mathrm{T}} \times q(\ell)$	NN		✓		✓		✓	✓
Boosted 1b-had-1t	$m_{\mathrm{T}} \times q(\ell)$	NN		✓		✓		✓	✓
Boosted 2b-1t	$m_{\mathrm{T}} \times q(\ell)$ (low-m_{T2}) Yield (high-m_{T2})	NN		✓		✓		✓	✓
Boosted 1b-lep-0t	$m_{\mathrm{T}} \times q(\ell)$	NN		✓		✓		✓	✓
Boosted 1b-had-0t	$m_{\mathrm{T}} \times q(\ell)$	NN		✓		✓		✓	✓
Boosted 2b-0t	$m_{\mathrm{T}} \times q(\ell)$ (low-m_{T2}) Yield (high-m_{T2})	NN		✓		✓		✓	✓

The validation of the predictions for the irreducible $t\bar{t}Z(\to \nu\nu)$ background is done using events enhanced in $t\bar{t}Z(\to \ell\ell)$. These events are selected with the same selection described in Section 6, except that exactly three signal leptons and no additional baseline lepton are required. Out of these three, two leptons are required to be of the same flavour and with opposite electric charge. The invariant mass of these two leptons is also required to be compatible with the leptonic decay of a Z boson, $m_{\ell\ell} \in [81, 101]$ GeV. In order to select $t\bar{t}Z(\to \ell\ell)$ events with kinematic properties similar to those of the $t\bar{t}Z(\to \nu\nu)$ events accepted in

the SRs, the three-lepton events are categorised and classified as the one-lepton events using a modified missing transverse momentum which includes the vectorial sum of the two leptons from the Z-boson decay.[7] As shown in Fig. 3, the observed distributions of events as a function of the stop-NN output value in the high-$E_{\mathrm{T}}^{\mathrm{miss}}$ three-leptons regions are in good agreement with predictions. These predictions are purely based on theoretical predictions without data-driven corrections. Similar agreement is observed also for the DM-NN output values. The three-lepton boosted regions are not considered as they do not yield enough events to provide useful validation regions.

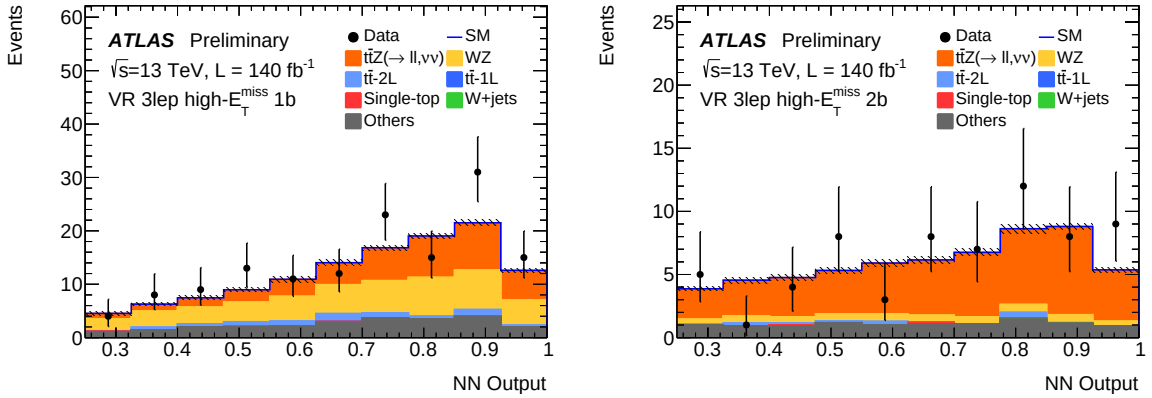


Figure 3: Observed and expected distributions of the stop-NN output values in the high-$E_{\mathrm{T}}^{\mathrm{miss}}$ three-lepton events with one (left) and two or more (right) b-tagged jets. Only statistical uncertainties are shown and no data-driven correction is applied to background predictions.

Data in the three-lepton categories are also useful to validate the accuracy of the simulations used in these searches in modelling the distributions of the NN output values in signal-like events. This assessment would have been otherwise not possible in samples of events without a potentially significant signal contamination.

8 Systematic uncertainties

The main experimental systematic uncertainties on the signal and background predictions include uncertainties on the determination of the energy scale and resolution of jets and of the $E_{\mathrm{T}}^{\mathrm{miss}}$ soft-term, as well as on the modelling of the efficiencies in b-tagging and top-tagging and in the (f)JVT selection. Other systematic uncertainties include uncertainties on the modelling of leptons in terms of energy scale and reconstruction, trigger, identification and isolation efficiencies. Uncertainties on the integrated luminosity and on the reweighting of the simulation as a function of the number of interactions per bunch crossing in data are also considered. The prediction of the small contribution of background events with a non-prompt lepton is estimated with a 50% uncertainty on the total yield, although the estimation of this small background bears minimal effect on the results of the searches.

Theoretical uncertainties on the differential cross section of the main sources of background are also considered. For $t\bar{t}$ and W+jets events, these include uncertainties related to factorisation and renormalisation

[7] The $E_{\mathrm{T}}^{\mathrm{miss}}$ significance used to classify three-lepton events is not computed with the modified missing transverse momentum, but rather with the original $\vec{p}_T^{\,\mathrm{miss}}$.

scales, PDF uncertainties, uncertainties on the matching of partons between ME and PS, uncertainties on the resummation scale, and uncertainties related to electroweak NLO corrections. For $t\bar{t}$, uncertainties on the fragmentation model are also considered. For W+jets, a relative uncertainty of 30% between the W+jets predictions in event categories with one and two or more b-tagged jets is assigned to cover uncertainties on the production of a W boson in association with heavy-flavour quarks [129]. To match the NF scheme used in the statistical model, these theoretical uncertainties are considered uncorrelated between $t\bar{t}$-1L and $t\bar{t}$-2L events, as well as between the predictions in the high-$E_{\mathrm{T}}^{\mathrm{miss}}$ and boosted categories. Theoretical uncertainties on the single-top production include uncertainties on the amount of initial-state radiation, on PDF, and on factorisation and renormalisation scales. An uncertainty on the removal of the on-shell contributions in the tW process is also accounted for as the difference in predicted yields when using the diagram removal or the diagram subtraction prescriptions. For $t\bar{t}Z$, uncertainties related to PDF and to factorisation and normalisation scales are considered. Other minor background contributions, mostly VV and ttV besides $t\bar{t}Z(\to \nu\nu, \ell\ell)$, are predicted with a conservative uncertainty of 30% on the total cross section [130, 131]. Theoretical uncertainties on signal predictions from factorisation and renormalisation scales and PDF uncertainties are also accounted for.

In the statistical model, systematic uncertainties are associated to nuisance parameters (NPs) which are constrained with Gaussian probability density functions. Statistical uncertainties on the simulated events are also accounted for. The impact of these uncertainties on the sensitivity of the searches is discussed in Section 10.

9 Results

Multiple types of fit, either for the $\tilde{t}_1\tilde{t}_1$ or the $t\bar{t}$+DM search, have been performed to analyse the observed data: first the 'background-only' fit with only data in CRs to validate the background modelling in VRs; then the background-only fit with data in CRs and SRs to check the compatibility of the data in SRs with only background contributions; and finally the 'full' fit in which data in CRs and SRs are fit with a model which includes contributions from background processes as well as from a signal model. These fits are performed separately for the $\tilde{t}_1\tilde{t}_1$ and the $t\bar{t}$+DM searches with their dedicated regions, all fit simultaneously.

In the first background-only fit, background predictions were fit to data in all CRs and then applied via the fit model to VRs, where predictions were found in agreement with data. After this validation of the background modelling, the background-only fit with also data in SRs was performed. Observed and expected distributions in CRs, VRs and SRs after such fit are shown in Figs. 4–9.[8] Expected distributions in SRs from benchmark signal models are also shown. In the $\tilde{t}_1\tilde{t}_1$ search, the NFs of the $t\bar{t}$, W+jets and single-top contributions are all compatible with unity. The NFs for the $t\bar{t}$-1L (-2L) background are 1.10 ± 0.07 (1.00 ± 0.05) and 1.00 ± 0.05 (1.21 ± 0.12) in the high-$E_{\mathrm{T}}^{\mathrm{miss}}$ and boosted regions, respectively. The NFs for the W+jets backgrounds are 1.1 ± 0.2 and 1.2 ± 0.1 in the same regions. The single-top NF is 1.4 ± 0.4, with the large uncertainty being in part due to the correlation with the uncertainty related to the removal of the on-shell tW contributions. Similar values are obtained in the fit for the $t\bar{t}$+DM search. Compared to Ref. [31], the overall agreement between data and predictions is better due to improved background modelling, namely the simulation of the $t\bar{t}$ production with SHERPA and the simulation of the

[8] Distributions in the boosted CRs as determined in the fit for the $t\bar{t}$+DM search are not shown as they are similar to those obtained in the fit for the stop search.

tW production with dynamic scales. The results of the background-only fits with and without data in SRs are statistically compatible with each other.

A good agreement between data and background predictions is seen in all CRs, VRs and SRs, although an excess of events in data compared to predictions is seen in SRs with two b-tagged jets in bins at high NN output values. To quantify the level of agreement between data and predictions and the significance of this excess, a profile-likelihood-ratio test [132] is performed on the result of the full fit, in which the statistical model includes signal predictions. In the $\tilde{t}_1\tilde{t}_1$ search, no point in the considered parameter space leads to a potential signal contribution determined to be different from zero at more than 2.3 standard deviations. The points with highest significance are those at high $m(\tilde{t}_1)$ and low $m(\tilde{\chi}_1^0)$ as their distributions are expected to be strongly peaked at high NN output values and predominantly accepted in the boosted 2b-1t category. For the $t\bar{t}$+DM search, no potential signal contribution is observed with significances from zero at more than 2.1 standard deviations. The significance does not show a clear dependence on the mass or parity of the mediator.

10 Interpretations

As no significant excess is observed, exclusion limits are calculated using the CL_s prescription [133] at 95% confidence level (CL). Figure 10 shows the expected and observed exclusion contours as a function of $m(\tilde{t}_1)$ and $m(\tilde{\chi}_1^0)$ for the $\tilde{t} \to t\tilde{\chi}_1^0$ or $\tilde{t} \to bW\tilde{\chi}_1^0$ scenarios. The difference between the observed and the expected limits at high $m(\tilde{t}_1)$ is due to the excess of events in data at high NN output values, as discussed in Section 9. The $\pm 1\sigma_{\text{exp}}$ uncertainty band indicates how much the expected limit is affected by systematic and statistical uncertainties. The results of this search are similar to those from Ref. [31] at high stop masses where the sensitivity is driven by the statistical power of the recorded data set. However, this search significantly improves the sensitivity for high neutralino masses covering the parameter region at $\Delta m(\tilde{t}_1, \tilde{\chi}_1^0) \sim m(t)$ which was not probed by the previous search. For signals with three-body decays ($m(W) < \Delta m(\tilde{t}_1, \tilde{\chi}_1^0) < m(t)$), this analysis achieves comparable sensitivity to Ref. [31] even though the stop-NNs have not been trained for these signal events and in Ref. [31] an optimised event selection and an NN-based discriminant were used. This result shows the benefit of the unified approach used in this search to cover the full parameter space without designing SRs targetting specific parameter regions. Table 5 shows the main sources of uncertainties on the precision in determining a potential signal contribution. Statistical uncertainties are the dominant source, followed by the theoretical uncertainties on the background predictions.

The exclusion limits for the $t\bar{t}$+DM search are shown in Fig. 11. These limits are computed as upper limits at 95% CL on the ratio of the $t\bar{t}$+DM production cross-section for the spin-0 mediator model to the theoretical cross-section. Limits are shown either as a function of $m(\phi/a)$ assuming $m(\chi) = 1$ GeV, or as a function of $m(\chi)$ assuming $m(\phi/a) = 10$ GeV. All limits are computed under the hypothesis that $g = 1$. All signal uncertainties are included in the fit. The $\pm 1\sigma$ and $\pm 2\sigma$ uncertainty bands around the expected upper limits account for all statistical and systematic uncertainties. The upper limits plotted as a function of $m(\chi)$ show a sharp weakening at $m(\chi) = 5$ GeV as this corresponds to the transition from the $t\bar{t}$+DM production with an on-shell to an off-shell mediator with $m(\phi/a) = 10$ GeV. The expected upper limits set by this search supersede those from the previous search in the one-lepton final state [31] and also improve on the combined limits from Ref. [33], in which the sensitivity is largely dominated by the search in the two-lepton final state. The differences between expected and observed limits are due to the excess of events in data in the high-$E_{\text{T}}^{\text{miss}}$ 2b SR in the bins at highest NN output values, as previously discussed. As shown

Table 5: Ratios of the signal cross sections as determined in the fit to the expected signal cross sections $\mu = \sigma_{\text{fit}}^{\text{sig}}/\sigma_{\text{Th}}^{\text{sig}}$ for different signal models in the $\tilde{t}_1\tilde{t}_1$ and $t\bar{t}$+DM searches. Components of the total uncertainty $\sigma(\mu)$ from statistical and major systematic uncertainties are also shown. The component from statistical uncertainties in data is estimated as the uncertainty on μ when all nuisance parameters in the fit are fixed. Components from systematic uncertainties are estimated as $\sigma_{\text{sys}}(\mu) = \sqrt{\sigma^2(\mu) - \sigma_{\text{fix}}^2(\mu)}$, where $\sigma_{\text{fix}}(\mu)$ is the uncertainty when the nuisance parameters associated to the systematic uncertainties are fixed. The components are reported as percentage relative to the total uncertainty $\sigma(\mu)$. Correlations across components are not accounted for.

	$\tilde{t}_1\tilde{t}_1$, $m(\tilde{t}_1, \tilde{\chi}_1^0)$ GeV		$t\bar{t}$+DM, $m(a, \chi)$ GeV	
	(1000, 600)	(1200, 200)	(50, 1)	(200, 1)
$\mu \pm \sigma(\mu)$ (total uncertainty)	0.27 ± 0.4	0.9 ± 0.5	0.14 ± 0.10	0.22 ± 0.18
Data statistical uncertainty	85 %	91 %	72 %	72 %
Background modelling	42 %	33 %	52 %	50 %
MC statistical uncertainty	24 %	23 %	31 %	36 %
Jet energy scale and resolution	21 %	13 %	29 %	28 %
Flavour tagging efficiency	17 %	11 %	20 %	19 %

in Table 5, the impact of statistical and systematic uncertainties on the sensitivity of the $t\bar{t}$+DM search is similar to that of the $\tilde{t}_1\tilde{t}_1$ search.

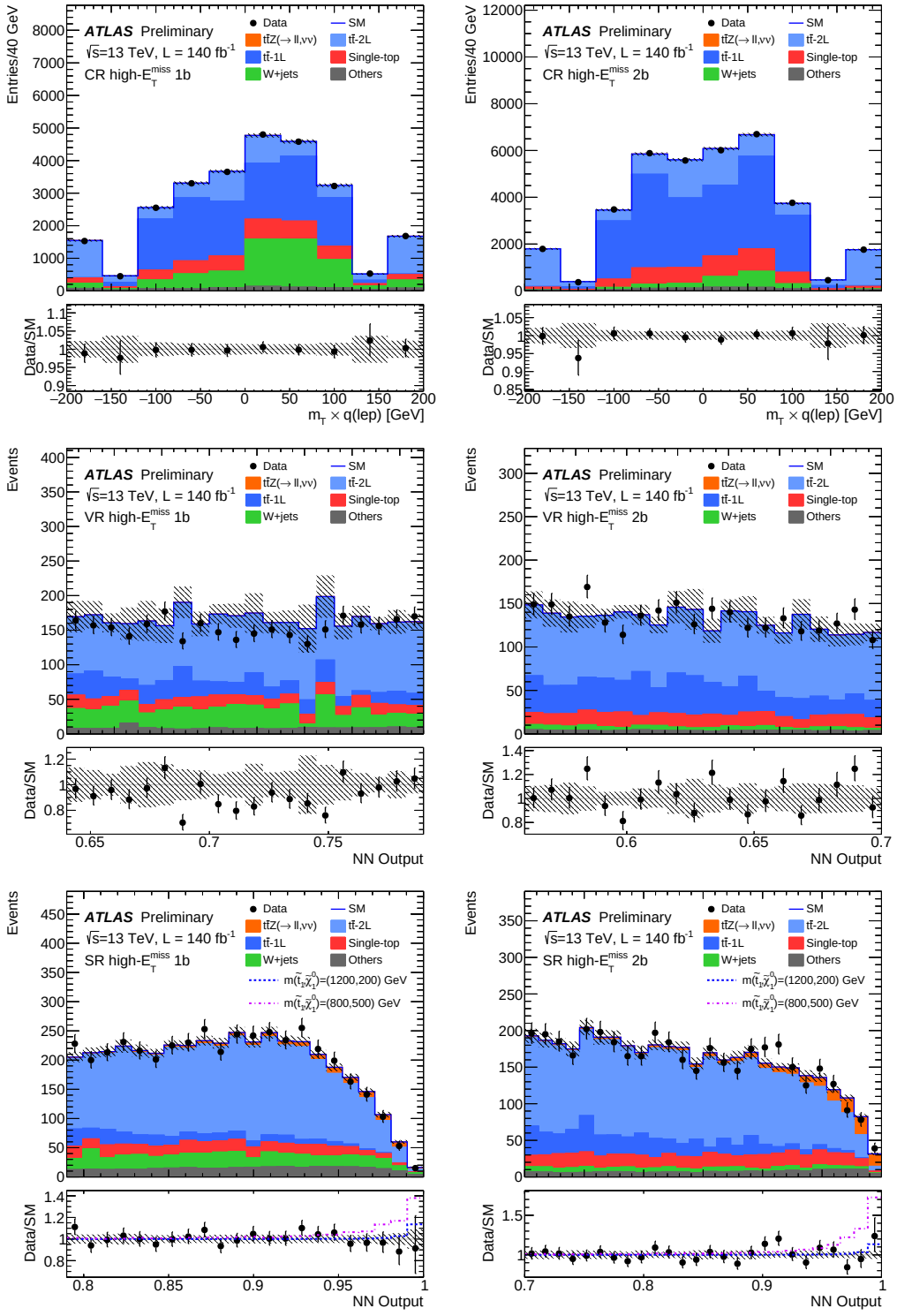


Figure 4: Observed and expected events as a function of $m_{\mathrm{T}} \times q(\ell)$ in the CR (top) and of the stop-NN output value in the VR (middle row) and the SR (bottom) in the high-$E_{\mathrm{T}}^{\mathrm{miss}}$ 1b (left) and 2b (right) stop regions. The expected distributions are as determined in the background-only fit with data in CRs and SRs for the $\tilde{t}_1\tilde{t}_1$ search. The hatched area around the total SM prediction includes statistical and systematic uncertainties. The lower panels show the ratio of the observed data to the expected SM events, as well as the ratio of the sum of the expected SM and signal events to the SM events. In CRs, the last (first) bin contains overflows (underflows).

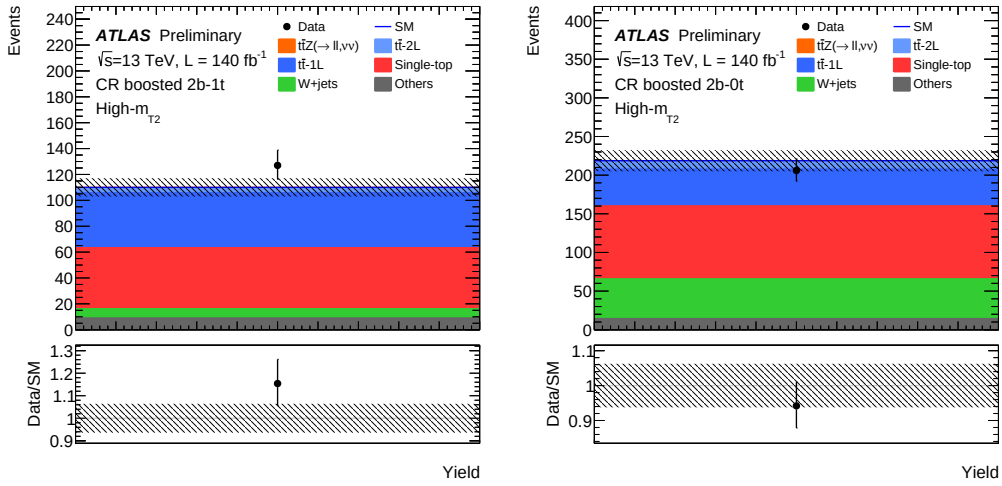


Figure 5: Observed and expected event yields in the high-m_{T2} CR in the boosted 2b-1t (left) and 2b-0t (right) stop regions. The expected distributions are as determined in the background-only fit with data in CRs and SRs for the $\tilde{t}_1\tilde{t}_1$ search. The hatched area around the total SM prediction includes statistical and systematic uncertainties. The lower panels show the ratio of the observed data to the expected SM events.

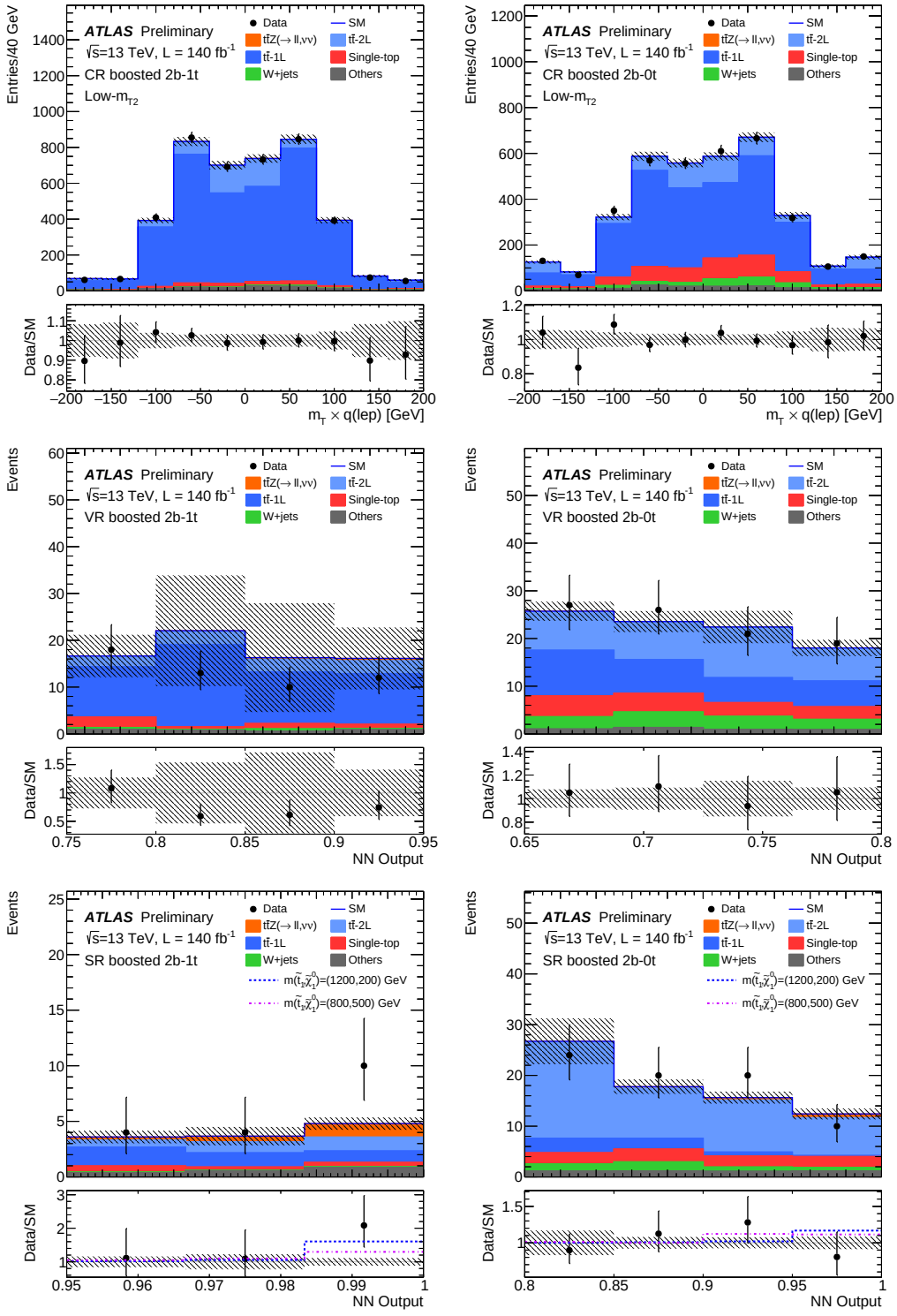


Figure 6: Observed and expected events distributed as a function of $m_{\mathrm{T}} \times q(\ell)$ in the low-m_{T2} CR (top) and of the stop-NN output value in the VR (middle row) and SR (bottom) in the boosted 2b-1t (left) and 2b-0t (right) stop regions. The expected distributions are as determined in the background-only fit with data in CRs and SRs for the $\tilde{t}_1\tilde{t}_1$ search. The hatched area around the total SM prediction includes statistical and systematic uncertainties. The lower panels show the ratio of the observed data to the expected SM events, as well as the ratio of the sum of the expected SM and signal events to the SM events. In CRs, the last (first) bin contains overflows (underflows).

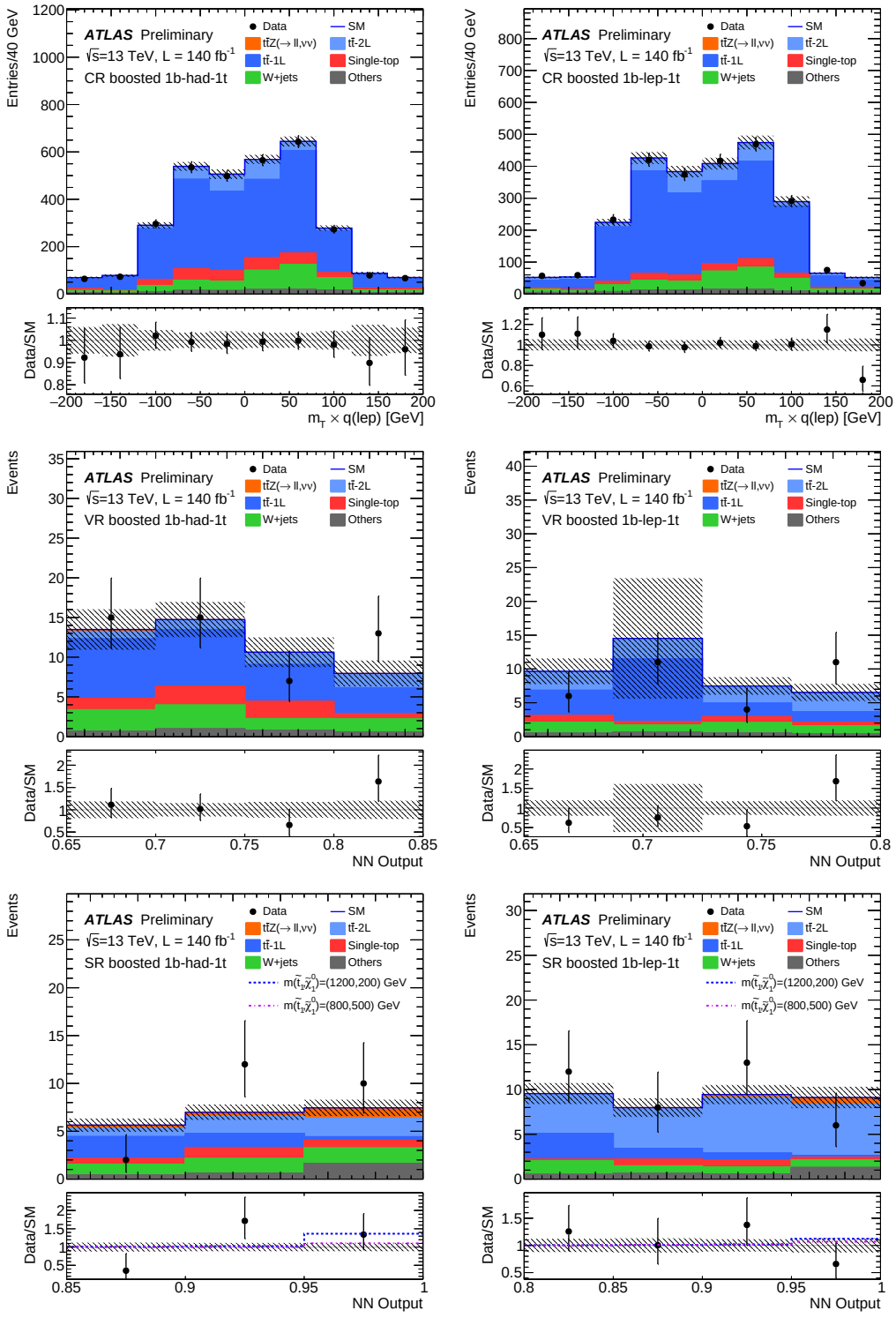


Figure 7: Observed and expected events as a function of $m_{\mathrm{T}} \times q(\ell)$ in the CR (top) and of the stop-NN output value in the VR (middle row) and the SR (bottom) in the boosted 1b-had-1t (left) and 1b-lep-1t (right) stop regions. The expected distributions are as determined in the background-only fit with data in CRs and SRs for the $\tilde{t}_1\tilde{t}_1$ search. The hatched area around the total SM prediction includes statistical and systematic uncertainties. The lower panels show the ratio of the observed data to the expected SM events, as well as the ratio of the sum of the expected SM and signal events to the SM events. In CRs, the last (first) bin contains overflows (underflows).

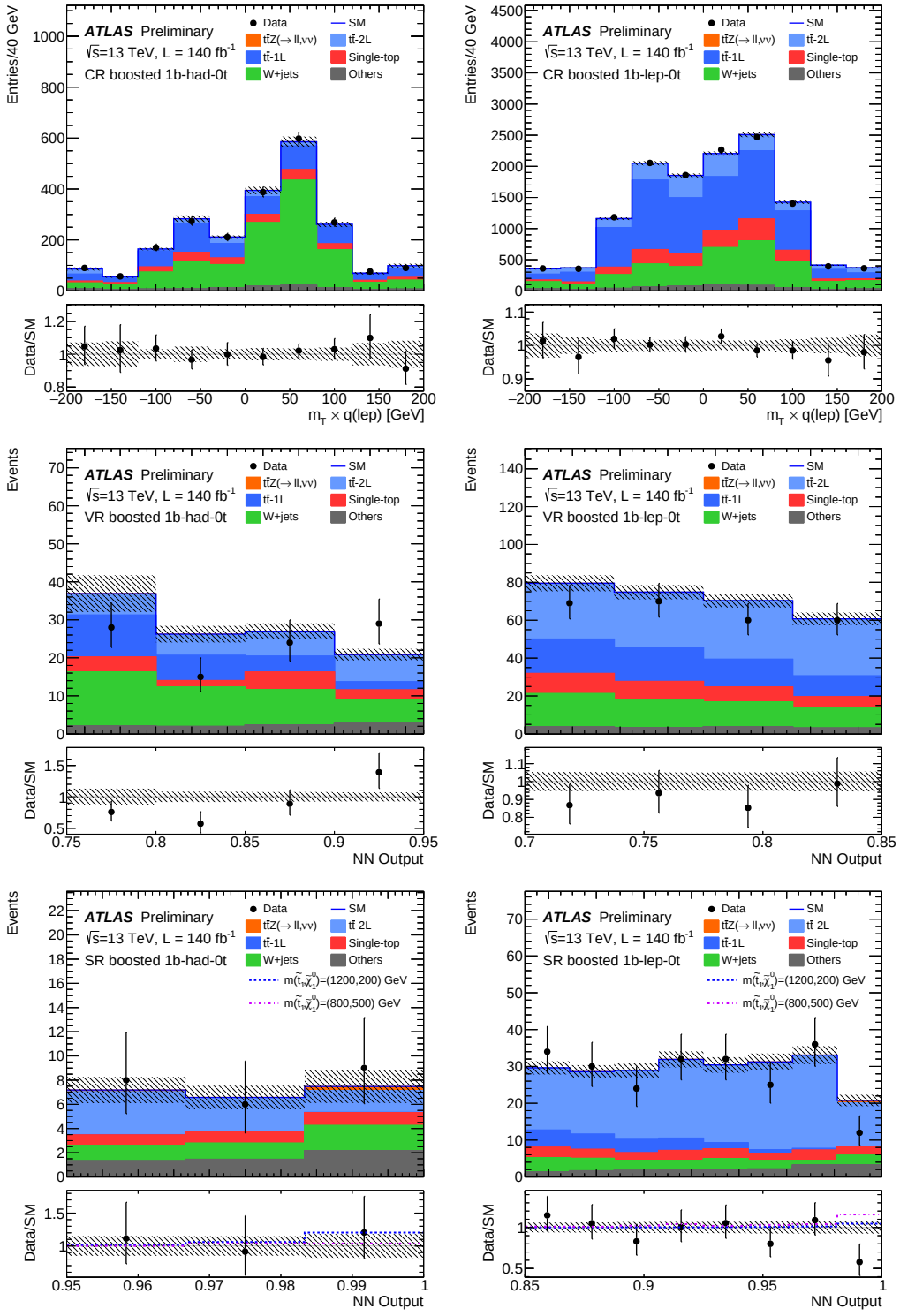


Figure 8: Observed and expected events as a function of $m_{\mathrm{T}} \times q(\ell)$ in the CR (top) and of the stop-NN output value in the VR (middle row) and the SR (bottom) in the boosted 1b-had-0t (left) and 1b-lep-0t (right) stop regions. The expected distributions are as determined in the background-only fit with data in CRs and SRs for the $\tilde{t}_1\tilde{t}_1$ search. The hatched area around the total SM prediction includes statistical and systematic uncertainties. The lower panels show the ratio of the observed data to the expected SM events, as well as the ratio of the sum of the expected SM and signal events to the SM events. In CRs, the last (first) bin contains overflows (underflows).

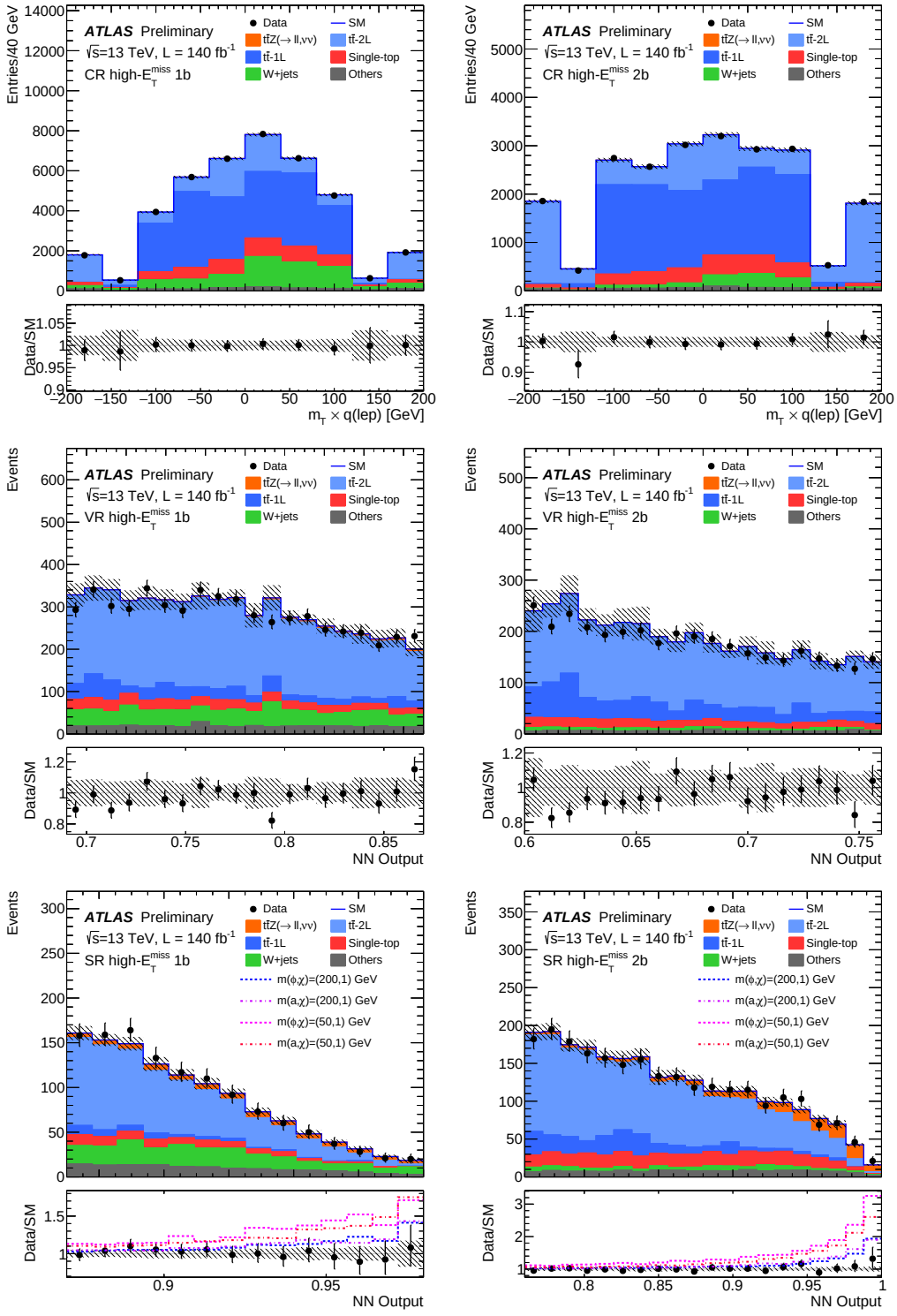


Figure 9: Observed and expected events as a function of $m_{\mathrm{T}} \times q(\ell)$ in the CR (top) and of the DM-NN output value in the VR (middle row) and the SR (bottom) in the high-$E_{\mathrm{T}}^{\mathrm{miss}}$ 1b (left) and 2b (right) DM regions. The expected distributions are as determined in the background-only fit with data in CRs and SRs for the $t\bar{t}$+DM search. The hatched area around the total SM prediction includes statistical and systematic uncertainties. The lower panels show the ratio of the observed data to the expected SM events, as well as the ratio of the sum of the expected SM and signal events to the SM events. In CRs, the last (first) bin contains overflows (underflows).

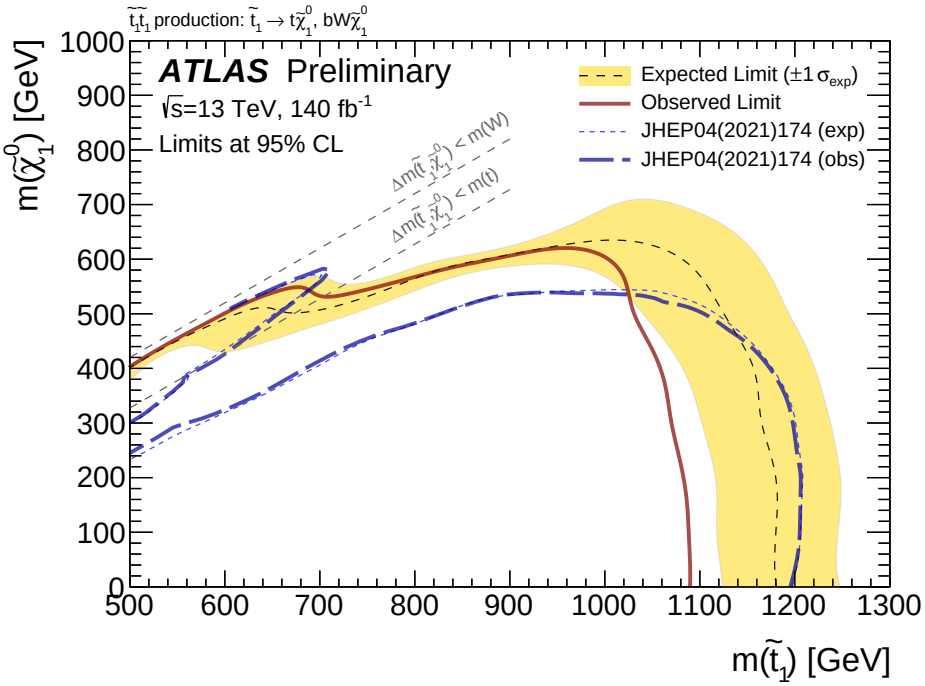


Figure 10: Expected and observed 95% CL excluded regions in the plane of $m(\tilde{\chi}_1^0)$ and $m(\tilde{t}_1)$ for $\tilde{t}_1\tilde{t}_1$ production assuming either a $\tilde{t} \to t\tilde{\chi}_1^0$ or $\tilde{t} \to bW\tilde{\chi}_1^0$ decay with a branching ratio of 100%. Models that lie within the contours are excluded. Uncertainty bands corresponding to the $\pm 1\sigma$ variation of the expected limit (yellow band) are also indicated. The diagonal dashed lines indicate the kinematic threshold of the stop decay modes.

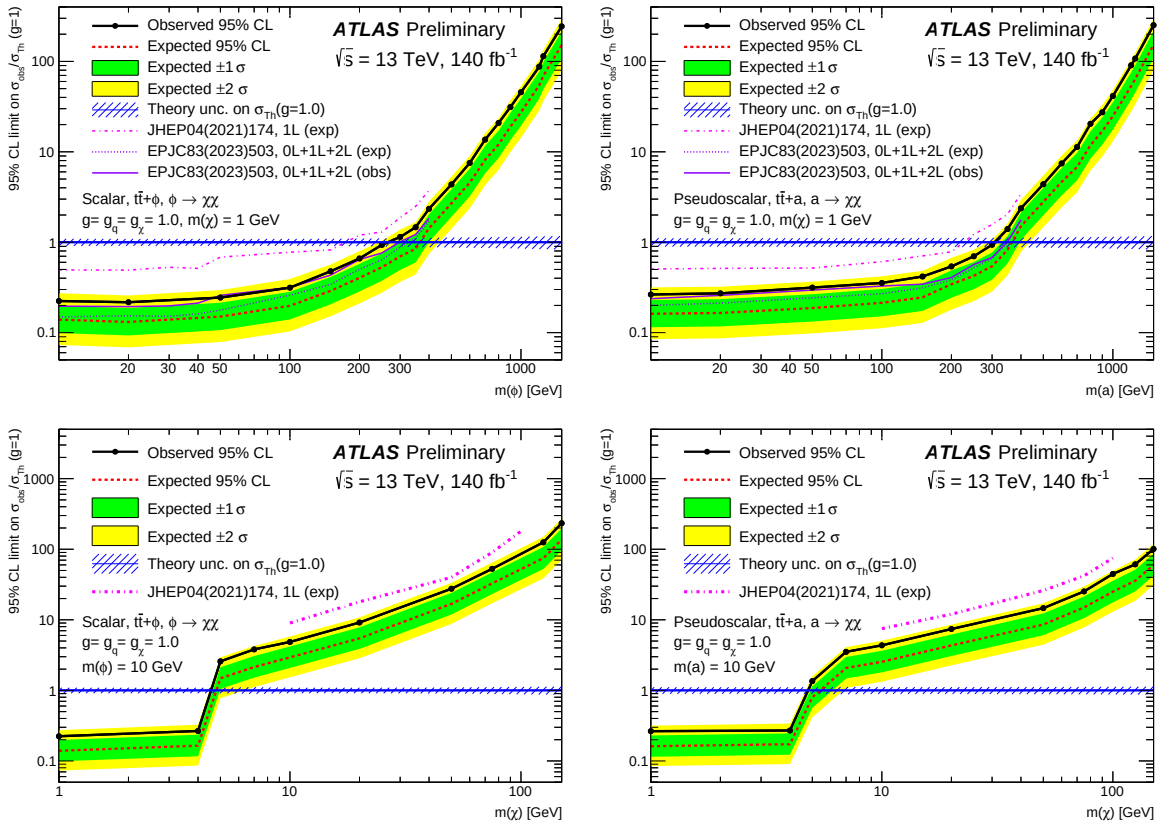


Figure 11: Upper limit at 95% CL on the ratio of the $t\bar{t}$+DM production cross-section to the theoretical cross-section under the hypothesis of (left) a scalar or (right) a pseudoscalar mediator. Limits are shown as a function of $m(\phi/a)$ assuming $m(\chi) = 1\,\text{GeV}$ (top), or as a function of $m(\chi)$ assuming $m(\phi/a) = 10\,\text{GeV}$ (bottom). All limits are computed assuming $g = 1$. Limits from previous publications [31, 33] are also shown. Limits from this search as well as those from Ref. [33] disregard tW+DM contributions to the signal yield.

11 Discussion and conclusions

This note presents searches for the direct production of a pair of top squarks, and for the production of a spin-0 mediator in association with a pair of top quarks and decaying into DM particles, using the full Run-2 pp data set recorded by ATLAS at $\sqrt{s} = 13$ TeV. These searches are conducted with events compatible with the presence of a semileptonic decay of a pair of top quarks and with large $E_{\mathrm{T}}^{\mathrm{miss}}$.

This note improves on the searches presented in Ref. [31] as a result of the development of a novel inclusive analysis approach based on neural networks. In this approach, inclusive event categories are defined using the multiplicities and kinematic properties of jets, b-tagged jets and $E_{\mathrm{T}}^{\mathrm{miss}}$. In each category, neural networks are trained by considering potential signals across the parameter space under study, so that the resulting signal regions are sensitive to wide ranges of the parameter space. A novel approach using neural networks is also developed to reconstruct hadronic top quark decays.

Compared to Ref. [31], which uses the same data set, these searches are sensitive to wide ranges of parameter space. In particular, the sensitivity to stop pair production for splitting between the stop and the neutralino masses at about the top mass is significantly improved. The sensitivity to spin-0 mediators is also improved across all the parameter space. In addition to the broader sensitivity, this novel analysis approach offers the benefit that no dedicated signal region need be optimised for specific regions of the parameter space. Moreover, with only one discriminant per category for all signal models under test the analysis is simpler because the data and background distributions remain the same while different signal predictions are tested. This also means the validation of the background modelling is more straightforward and the interpretability of results for new signals is likely to be simpler and more effective. Another important development included in this note is the improved modelling of top quark production both in pairs and via single-top processes.

No statistically significant deviations from the SM expectation are observed and exclusion limits at 95% confidence level are set. In the simplified models considered in this note, stops are excluded for masses up to 1090 GeV, while neutralinos are excluded for masses up to 600 GeV. Compared to the parameter space excluded by Ref. [31], the region with stop-neutralino mass splitting slightly above the top mass and neutralino masses below 500 GeV is now excluded, while the maximum excluded neutralino mass is increased by about 80 GeV. In the $t\bar{t}$+DM search, when assuming the strengths of the couplings of the mediator to DM and SM particles $g = g_q = g_\chi = 1$, models with scalar (pseudoscalar) mediators with masses up to 250 (300) GeV are excluded, while at lower mediator masses, models with production cross sections as small as 0.22 (0.26) times the nominal predictions are excluded. These cross sections are about half as large as the cross sections excluded in Ref. [31].

References

[1] Y. Golfand and E. Likhtman,
Extension of the Algebra of Poincare Group Generators and Violation of P Invariance,
JETP Lett. **13** (1971) 323, [Pisma Zh. Eksp. Teor. Fiz. **13** (1971) 452] (cit. on p. 2).

[2] D. Volkov and V. Akulov, *Is the neutrino a goldstone particle?*, Phys. Lett. B **46** (1973) 109 (cit. on p. 2).

[3] J. Wess and B. Zumino, *Supergauge transformations in four dimensions*,
Nucl. Phys. B **70** (1974) 39 (cit. on p. 2).

[4] J. Wess and B. Zumino, *Supergauge invariant extension of quantum electrodynamics*, Nucl. Phys. B **78** (1974) 1 (cit. on p. 2).

[5] S. Ferrara and B. Zumino, *Supergauge invariant Yang-Mills theories*, Nucl. Phys. B **79** (1974) 413 (cit. on p. 2).

[6] A. Salam and J. Strathdee, *Super-symmetry and non-Abelian gauges*, Phys. Lett. B **51** (1974) 353 (cit. on p. 2).

[7] G. R. Farrar and P. Fayet, *Phenomenology of the production, decay, and detection of new hadronic states associated with supersymmetry*, Phys. Lett. B **76** (1978) 575 (cit. on p. 2).

[8] G. D'Ambrosio, G. Giudice, G. Isidori and A. Strumia, *Minimal flavour violation: an effective field theory approach*, Nucl. Phys. B **645** (2002) 155, arXiv: hep-ph/0207036 (cit. on p. 2).

[9] G. Isidori and D. M. Straub, *Minimal flavour violation and beyond*, Eur. Phys. J. C **72** (2012), arXiv: 1202.0464 [hep-ph] (cit. on p. 2).

[10] ATLAS Collaboration, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Phys. Lett. B **716** (2012) 1, arXiv: 1207.7214 [hep-ex] (cit. on p. 2).

[11] CMS Collaboration, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, Phys. Lett. B **716** (2012) 30, arXiv: 1207.7235 [hep-ex] (cit. on p. 2).

[12] S. Weinberg, *Implications of dynamical symmetry breaking*, Phys. Rev. D **13** (4 1976) 974 (cit. on p. 2).

[13] E. Gildener, *Gauge-symmetry hierarchies*, Phys. Rev. D **14** (6 1976) 1667 (cit. on p. 2).

[14] S. Weinberg, *Implications of dynamical symmetry breaking: An addendum*, Phys. Rev. D **19** (4 1979) 1277 (cit. on p. 2).

[15] L. Susskind, *Dynamics of spontaneous symmetry breaking in the Weinberg-Salam theory*, Phys. Rev. D **20** (10 1979) 2619 (cit. on p. 2).

[16] S. Dimopoulos and H. Georgi, *Softly broken supersymmetry and SU(5)*, Nucl. Phys. B **193** (1981) 150 (cit. on p. 2).

[17] E. Witten, *Dynamical breaking of supersymmetry*, Nucl. Phys. B **188** (1981) 513 (cit. on p. 2).

[18] M. Dine, W. Fischler and M. Srednicki, *Supersymmetric technicolor*, Nucl. Phys. B **189** (1981) 575 (cit. on p. 2).

[19] S. Dimopoulos and S. Raby, *Supercolor*, Nucl. Phys. B **192** (1981) 353 (cit. on p. 2).

[20] N. Sakai, *Naturalnes in supersymmetric GUTS*, Z. Phys. C **11** (1981) 153 (cit. on p. 2).

[21] R. K. Kaul and P. Majumdar, *Cancellation of quadratically divergent mass corrections in globally supersymmetric spontaneously broken gauge theories*, Nucl. Phys. B **199** (1982) 36 (cit. on p. 2).

[22] R. Barbieri and G. Giudice, *Upper bounds on supersymmetric particle masses*, Nucl. Phys. B **306** (1988) 63 (cit. on p. 2).

[23] B. de Carlos and J. Casas, *One-loop analysis of the electroweak breaking in supersymmetric models and the fine-tuning problem*, Phys. Lett. B **309** (1993) 320 (cit. on p. 2).

[24] P. Fayet, *Supersymmetry and weak, electromagnetic and strong interactions*, Phys. Lett. B **64** (1976) 159 (cit. on p. 2).

[25] P. Fayet,
Spontaneously broken supersymmetric theories of weak, electromagnetic and strong interactions,
Phys. Lett. B **69** (1977) 489 (cit. on p. 2).

[26] P. Fayet, *Relations between the masses of the superpartners of leptons and quarks, the goldstino coupling and the neutral currents*, Physics Letters B **84** (1979) 416 (cit. on p. 2).

[27] H. Goldberg, *Constraint on the Photino Mass from Cosmology*, Phys. Rev. Lett. **50** (1983) 1419 (cit. on p. 2), Erratum: Phys. Rev. Lett. **103** (2009) 099905.

[28] J. Ellis, J. Hagelin, D. V. Nanopoulos, K. A. Olive and M. Srednicki,
Supersymmetric relics from the big bang, Nucl. Phys. B **238** (1984) 453 (cit. on p. 2).

[29] D. Abercrombie et al., *Dark Matter benchmark models for early LHC Run-2 Searches: Report of the ATLAS/CMS Dark Matter Forum*, Phys. Dark Univ. **27** (2020) 100371 (cit. on pp. 2, 3).

[30] ATLAS Collaboration, *Search for a scalar partner of the top quark in the all-hadronic $t\bar{t}$ plus missing transverse momentum final state at $\sqrt{s} = 13$ TeV with the ATLAS detector*,
Eur. Phys. J. C **80** (2020) 737, arXiv: 2004.14060 [hep-ex] (cit. on p. 2).

[31] ATLAS Collaboration,
Search for new phenomena with top quark pairs in final states with one lepton, jets, and missing transverse momentum in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector,
JHEP **04** (2020) 174, arXiv: 2012.03799 [hep-ex] (cit. on pp. 2, 3, 5, 6, 8, 11, 15, 16, 25, 26).

[32] ATLAS Collaboration, *Search for new phenomena in events with two opposite-charge leptons, jets and missing transverse momentum in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*,
JHEP **04** (2021) 165, arXiv: 2102.01444 [hep-ex] (cit. on p. 2).

[33] ATLAS Collaboration,
Constraints on spin-0 dark matter mediators and invisible Higgs decays using ATLAS 13 TeV pp collision data with two top quarks and missing transverse momentum in the final state,
Eur. Phys. J. C **83** (2023) 503, arXiv: 2211.05426 [hep-ex] (cit. on pp. 2, 3, 16, 25).

[34] CMS Collaboration, *Search for top squark production in fully-hadronic final states in proton–proton collisions at $\sqrt{s} = 13$ TeV*, Phys. Rev. D **104** (2021) 052001,
arXiv: 2103.01290 [hep-ex] (cit. on p. 2).

[35] CMS Collaboration, *Search for direct top squark pair production in events with one lepton, jets, and missing transverse momentum at 13 TeV with the CMS experiment*, JHEP **05** (2020) 032,
arXiv: 1912.08887 [hep-ex] (cit. on p. 2).

[36] CMS Collaboration, *Search for top squark pair production using dilepton final states in pp collision data collected at $\sqrt{s} = 13$ TeV*, Eur. Phys. J. C **81** (2021) 3,
arXiv: 2008.05936 [hep-ex] (cit. on p. 2).

[37] CMS Collaboration, *Combined searches for the production of supersymmetric top quark partners in proton–proton collisions at $\sqrt{s} = 13$ TeV*, Eur. Phys. J. C **81** (2021) 970,
arXiv: 2107.10892 [hep-ex] (cit. on p. 2).

[38] J. Alwall, M.-P. Le, M. Lisanti and J. G. Wacker,
Searching for directly decaying gluinos at the Tevatron, Phys. Lett. B **666** (2008) 34,
arXiv: 0803.0019 [hep-ph] (cit. on p. 3).

[39] J. Alwall, P. Schuster and N. Toro,
Simplified models for a first characterization of new physics at the LHC,
Phys. Rev. D **79** (2009) 075020, arXiv: 0810.3921 [hep-ph] (cit. on p. 3).

[40] ATLAS Collaboration, *Search for dark matter produced in association with bottom or top quarks in $\sqrt{s} = 13$ TeV pp collisions with the ATLAS detector*, Eur. Phys. J. C **78** (2018) 18, arXiv: 1710.11412 [hep-ex] (cit. on p. 3).

[41] ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, JINST **3** (2008) S08003 (cit. on p. 4).

[42] ATLAS Collaboration, *ATLAS Insertable B-Layer: Technical Design Report*, ATLAS-TDR-19; CERN-LHCC-2010-013, 2010, URL: https://cds.cern.ch/record/1291633 (cit. on p. 4), Addendum: ATLAS-TDR-19-ADD-1; CERN-LHCC-2012-009, 2012, URL: https://cds.cern.ch/record/1451888.

[43] B. Abbott et al., *Production and integration of the ATLAS Insertable B-Layer*, JINST **13** (2018) T05008, arXiv: 1803.00844 [physics.ins-det] (cit. on p. 4).

[44] ATLAS Collaboration, *Performance of the ATLAS trigger system in 2015*, Eur. Phys. J. C **77** (2017) 317, arXiv: 1611.09661 [hep-ex] (cit. on p. 4).

[45] ATLAS Collaboration, *The ATLAS Collaboration Software and Firmware*, ATL-SOFT-PUB-2021-001, 2021, URL: https://cds.cern.ch/record/2767187 (cit. on p. 4).

[46] ATLAS Collaboration, *ATLAS data quality operations and performance for 2015–2018 data-taking*, JINST **15** (2020) P04003, arXiv: 1911.04632 [physics.ins-det] (cit. on p. 5).

[47] ATLAS Collaboration, *Luminosity determination in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector at the LHC*, (2022), arXiv: 2212.09379 [hep-ex] (cit. on p. 5).

[48] G. Avoni et al., *The new LUCID-2 detector for luminosity measurement and monitoring in ATLAS*, JINST **13** (2018) P07017 (cit. on p. 5).

[49] ATLAS Collaboration, *Performance of the missing transverse momentum triggers for the ATLAS detector during Run-2 data taking*, JHEP **08** (2020) 080, arXiv: 2005.09554 [hep-ex] (cit. on p. 5).

[50] ATLAS Collaboration, *Performance of electron and photon triggers in ATLAS during LHC Run 2*, Eur. Phys. J. C **80** (2020) 47, arXiv: 1909.00761 [hep-ex] (cit. on p. 5).

[51] ATLAS Collaboration, *Performance of the ATLAS muon triggers in Run 2*, JINST **15** (2020) P09015, arXiv: 2004.13447 [hep-ex] (cit. on p. 5).

[52] 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] (cit. on pp. 5, 6).

[53] M. Beneke, M. Czakon, P. Falgari, A. Mitov and C. Schwinn, *Threshold expansion of the $gg(q\bar{q}) \to Q\bar{Q} + X$ cross section at $O(\alpha_s^4)$*, Phys. Lett. B **690** (2010) 483, arXiv: 0911.5166 [hep-ph] (cit. on p. 5), Erratum: Phys. Lett. B **778** (2018) 464.

[54] 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] (cit. on p. 5).

[55] W. Beenakker, M. Krämer, T. Plehn, M. Spira and P. Zerwas, *Stop production at hadron colliders*, Nucl. Phys. B **515** (1998) 3, arXiv: hep-ph/9710451 (cit. on p. 5).

[56] W. Beenakker et al., *Supersymmetric top and bottom squark production at hadron colliders*, JHEP **08** (2010) 098, arXiv: 1006.4771 [hep-ph] (cit. on p. 5).

[57] W. Beenakker et al., *NNLL resummation for stop pair-production at the LHC*, JHEP **05** (2016) 153, arXiv: 1601.02954 [hep-ph] (cit. on p. 5).

[58] O. Mattelaer and E. Vryonidou, *Dark-matter production through loop-induced processes at the LHC: the s-channel mediator case*, Eur. Phys. J. C **75** (2015) 133 (cit. on pp. 5, 6).

[59] M. Backović et al., *Higher-order QCD predictions for dark matter production at the LHC in simplified models with s-channel mediators*, Eur. Phys. J. C **75** (2015) 482 (cit. on pp. 5, 6).

[60] E. Bothmann et al., *Event generation with Sherpa 2.2*, SciPost Phys. **7** (2019) 034, arXiv: 1905.09127 [hep-ph] (cit. on pp. 5, 6).

[61] The NNPDF Collaboration, R. D. Ball et al., *Parton distributions for the LHC run II*, JHEP **04** (2015) 040, arXiv: 1410.8849 [hep-ph] (cit. on p. 5).

[62] S. Schumann and F. Krauss, *A parton shower algorithm based on Catani–Seymour dipole factorisation*, JHEP **03** (2008) 038, arXiv: 0709.1027 [hep-ph] (cit. on pp. 5, 6).

[63] S. Höche, F. Krauss, M. Schönherr and F. Siegert, *A critical appraisal of NLO+PS matching methods*, JHEP **09** (2012) 049, arXiv: 1111.1220 [hep-ph] (cit. on pp. 5, 6).

[64] S. Höche, F. Krauss, M. Schönherr and F. Siegert, *QCD matrix elements + parton showers. The NLO case*, JHEP **04** (2013) 027, arXiv: 1207.5030 [hep-ph] (cit. on pp. 5, 6).

[65] S. Catani, F. Krauss, B. R. Webber and R. Kuhn, *QCD Matrix Elements + Parton Showers*, JHEP **11** (2001) 063, arXiv: hep-ph/0109231 (cit. on pp. 5, 6).

[66] S. Höche, F. Krauss, S. Schumann and F. Siegert, *QCD matrix elements and truncated showers*, JHEP **05** (2009) 053, arXiv: 0903.1219 [hep-ph] (cit. on pp. 5, 6).

[67] M. Beneke, P. Falgari, S. Klein and C. Schwinn, *Hadronic top-quark pair production with NNLL threshold resummation*, Nucl. Phys. B **855** (2012) 695, arXiv: 1109.1536 [hep-ph] (cit. on p. 5).

[68] M. Cacciari, M. Czakon, M. Mangano, A. Mitov and P. Nason, *Top-pair production at hadron colliders with next-to-next-to-leading logarithmic soft-gluon resummation*, Phys. Lett. B **710** (2012) 612, arXiv: 1111.5869 [hep-ph] (cit. on p. 5).

[69] P. Bärnreuther, M. Czakon and A. Mitov, *Percent-Level-Precision Physics at the Tevatron: Next-to-Next-to-Leading Order QCD Corrections to $q\bar{q} \to t\bar{t} + X$*, Phys. Rev. Lett. **109** (2012) 132001, arXiv: 1204.5201 [hep-ph] (cit. on p. 5).

[70] M. Czakon and A. Mitov, *NNLO corrections to top-pair production at hadron colliders: the all-fermionic scattering channels*, JHEP **12** (2012) 054, arXiv: 1207.0236 [hep-ph] (cit. on p. 5).

[71] M. Czakon and A. Mitov,
NNLO corrections to top pair production at hadron colliders: the quark-gluon reaction,
JHEP **01** (2013) 080, arXiv: 1210.6832 [hep-ph] (cit. on p. 5).

[72] M. Czakon, P. Fiedler and A. Mitov,
Total Top-Quark Pair-Production Cross Section at Hadron Colliders Through $O(\alpha_S^4)$,
Phys. Rev. Lett. **110** (2013) 252004, arXiv: 1303.6254 [hep-ph] (cit. on p. 5).

[73] 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] (cit. on p. 5).

[74] T. Gleisberg and S. Höche, *Comix, a new matrix element generator*, JHEP **12** (2008) 039,
arXiv: 0808.3674 [hep-ph] (cit. on pp. 5, 6).

[75] F. Buccioni et al., *OpenLoops 2*, Eur. Phys. J. C **79** (2019) 866, arXiv: 1907.13071 [hep-ph]
(cit. on pp. 5, 6).

[76] F. Cascioli, P. Maierhöfer and S. Pozzorini, *Scattering Amplitudes with Open Loops*,
Phys. Rev. Lett. **108** (2012) 111601, arXiv: 1111.5206 [hep-ph] (cit. on pp. 5, 6).

[77] A. Denner, S. Dittmaier and L. Hofer,
COLLIER: A fortran-based complex one-loop library in extended regularizations,
Comput. Phys. Commun. **212** (2017) 220, arXiv: 1604.06792 [hep-ph] (cit. on pp. 5, 6).

[78] E. Re,
Single-top Wt-channel production matched with parton showers using the POWHEG method,
Eur. Phys. J. C **71** (2011) 1547, arXiv: 1009.2450 [hep-ph] (cit. on pp. 5, 6).

[79] P. Nason, *A new method for combining NLO QCD with shower Monte Carlo algorithms*,
JHEP **11** (2004) 040, arXiv: hep-ph/0409146 (cit. on pp. 5, 6).

[80] 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] (cit. on pp. 5, 6).

[81] 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] (cit. on pp. 5, 6).

[82] T. Sjöstrand et al., *An introduction to PYTHIA 8.2*, Comput. Phys. Commun. **191** (2015) 159,
arXiv: 1410.3012 [hep-ph] (cit. on p. 5).

[83] N. Kidonakis,
Two-loop soft anomalous dimensions for single top quark associated production with a W^- or H^-,
Phys. Rev. D **82** (2010) 054018, arXiv: 1005.4451 [hep-ph] (cit. on p. 5).

[84] N. Kidonakis, 'Top Quark Production', *Proceedings, Helmholtz International Summer School on Physics of Heavy Quarks and Hadrons (HQ 2013)* (JINR, Dubna, Russia, 15th–28th July 2013) 139, arXiv: 1311.0283 [hep-ph] (cit. on p. 5).

[85] M. Aliev et al., *HATHOR – HAdronic Top and Heavy quarks crOss section calculatoR*,
Comput. Phys. Commun. **182** (2011) 1034, arXiv: 1007.1327 [hep-ph] (cit. on p. 5).

[86] P. Kant et al., *HatHor for single top-quark production: Updated predictions and uncertainty estimates for single top-quark production in hadronic collisions*,
Comput. Phys. Commun. **191** (2015) 74, arXiv: 1406.4403 [hep-ph] (cit. on p. 5).

[87] C. Anastasiou, L. Dixon, K. Melnikov and F. Petriello, *High-precision QCD at hadron colliders: Electroweak gauge boson rapidity distributions at next-to-next-to leading order*, Phys. Rev. D **69** (2004) 094008, arXiv: hep-ph/0312266 (cit. on p. 5).

[88] 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] (cit. on p. 5).

[89] D. J. Lange, *The EvtGen particle decay simulation package*, Nucl. Instrum. Meth. A **462** (2001) 152 (cit. on p. 5).

[90] ATLAS Collaboration, *ATLAS Pythia 8 tunes to 7 TeV data*, ATL-PHYS-PUB-2014-021, 2014, URL: https://cds.cern.ch/record/1966419 (cit. on p. 5).

[91] NNPDF Collaboration, R. D. Ball et al., *Parton distributions with LHC data*, Nucl. Phys. B **867** (2013) 244, arXiv: 1207.1303 [hep-ph] (cit. on p. 5).

[92] T. Sjöstrand, S. Mrenna and P. Skands, *A brief introduction to PYTHIA 8.1*, Comput. Phys. Commun. **178** (2008) 852, arXiv: 0710.3820 [hep-ph] (cit. on p. 5).

[93] ATLAS Collaboration, *The Pythia 8 A3 tune description of ATLAS minimum bias and inelastic measurements incorporating the Donnachie–Landshoff diffractive model*, ATL-PHYS-PUB-2016-017, 2016, URL: https://cds.cern.ch/record/2206965 (cit. on p. 5).

[94] ATLAS Collaboration, *The ATLAS Simulation Infrastructure*, Eur. Phys. J. C **70** (2010) 823, arXiv: 1005.4568 [physics.ins-det] (cit. on p. 5).

[95] S. Agostinelli et al., *GEANT4 – a simulation toolkit*, Nucl. Instrum. Meth. A **506** (2003) 250 (cit. on p. 5).

[96] P. Artoisenet, R. Frederix, O. Mattelaer and R. Rietkerk, *Automatic spin-entangled decays of heavy resonances in Monte Carlo simulations*, JHEP **03** (2013) 015, arXiv: 1212.3460 [hep-ph] (cit. on p. 5).

[97] J. Butterworth et al., *PDF4LHC recommendations for LHC Run II*, J. Phys. G **43** (2016) 023001, arXiv: 1510.03865 [hep-ph] (cit. on p. 5).

[98] S. Frixione, E. Laenen, P. Motylinski, C. White and B. R. Webber, *Single-top hadroproduction in association with a W boson*, JHEP **07** (2008) 029, arXiv: 0805.3067 [hep-ph] (cit. on p. 6).

[99] ATLAS Collaboration, *Studies on top-quark Monte Carlo modelling for Top2016*, ATL-PHYS-PUB-2016-020, 2016, URL: https://cds.cern.ch/record/2216168 (cit. on p. 6).

[100] ATLAS Collaboration, *Electron and photon performance measurements with the ATLAS detector using the 2015–2017 LHC proton–proton collision data*, JINST **14** (2019) P12006, arXiv: 1908.00005 [hep-ex] (cit. on p. 6).

[101] 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] (cit. on p. 6).

[102] ATLAS Collaboration, *Muon reconstruction and identification efficiency in ATLAS using the full Run 2 pp collision data set at* $\sqrt{s} = 13$ *TeV*, Eur. Phys. J. C **81** (2021) 578, arXiv: 2012.00578 [hep-ex] (cit. on p. 6).

[103] M. Cacciari, G. P. Salam and G. Soyez, *The anti-k_t jet clustering algorithm*, JHEP **04** (2008) 063, arXiv: 0802.1189 [hep-ph] (cit. on p. 7).

[104] M. Cacciari, G. P. Salam and G. Soyez, *FastJet user manual*, Eur. Phys. J. C **72** (2012) 1896, arXiv: 1111.6097 [hep-ph] (cit. on p. 7).

[105] ATLAS Collaboration, *Jet reconstruction and performance using particle flow with the ATLAS Detector*, Eur. Phys. J. C **77** (2017) 466, arXiv: 1703.10485 [hep-ex] (cit. on p. 7).

[106] ATLAS Collaboration, *Jet energy scale and resolution measured in proton–proton collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector*, Eur. Phys. J. C **81** (2020) 689, arXiv: 2007.02645 [hep-ex] (cit. on p. 7).

[107] ATLAS Collaboration, *Performance of pile-up mitigation techniques for jets in pp collisions at $\sqrt{s}$ = 8 TeV using the ATLAS detector*, Eur. Phys. J. C **76** (2016) 581, arXiv: 1510.03823 [hep-ex] (cit. on p. 7).

[108] ATLAS Collaboration, *Forward jet vertex tagging using the particle flow algorithm*, ATL-PHYS-PUB-2019-026, 2019, URL: https://cds.cern.ch/record/2683100 (cit. on p. 7).

[109] ATLAS Collaboration, *ATLAS b-jet identification performance and efficiency measurement with $t\bar{t}$ events in pp collisions at $\sqrt{s}$ = 13 TeV*, Eur. Phys. J. C **79** (2019) 970, arXiv: 1907.05120 [hep-ex] (cit. on p. 7).

[110] ATLAS Collaboration, *ATLAS flavour-tagging algorithms for the LHC Run 2 pp collision dataset*, (2022), arXiv: 2211.16345 [physics.data-an] (cit. on p. 7).

[111] ATLAS Collaboration, *Optimisation of large-radius jet reconstruction for the ATLAS detector in 13 TeV proton–proton collisions*, Eur. Phys. J. C **81** (2020) 334, arXiv: 2009.04986 [hep-ex] (cit. on p. 7).

[112] T. Barillari et al., *Local Hadronic Calibration*, ATL-LARG-PUB-2009-001-2, 2008, URL: https://cds.cern.ch/record/1112035 (cit. on p. 7).

[113] D. Krohn, J. Thaler and L.-T. Wang, *Jet Trimming*, JHEP **02** (2010) 084, arXiv: 0912.1342 [hep-ph] (cit. on p. 7).

[114] ATLAS Collaboration, *In situ calibration of large-radius jet energy and mass in 13 TeV proton–proton collisions with the ATLAS detector*, Eur. Phys. J. C **79** (2019) 135, arXiv: 1807.09477 [hep-ex] (cit. on p. 7).

[115] ATLAS Collaboration, *Measurement of the ATLAS Detector Jet Mass Response using Forward Folding with 80 fb^{-1} of $\sqrt{s}$ = 13 TeV pp data*, ATLAS-CONF-2020-022, 2020, URL: https://cds.cern.ch/record/2724442 (cit. on p. 7).

[116] ATLAS Collaboration, *Performance of top-quark and W-boson tagging with ATLAS in Run 2 of the LHC*, Eur. Phys. J. C **79** (2019) 375, arXiv: 1808.07858 [hep-ex] (cit. on p. 7).

[117] ATLAS Collaboration, *Boosted hadronic vector boson and top quark tagging with ATLAS using Run 2 data*, ATL-PHYS-PUB-2020-017, 2020, URL: https://cds.cern.ch/record/2724149 (cit. on p. 7).

[118] D. Krohn, J. Thaler and L.-T. Wang, *Jets with Variable R*, JHEP **06** (2009) 059, arXiv: 0903.0392 [hep-ph] (cit. on p. 7).

[119] ATLAS Collaboration, *Reconstruction, Identification, and Calibration of hadronically decaying tau leptons with the ATLAS detector for the LHC Run 3 and reprocessed Run 2 data*, ATL-PHYS-PUB-2022-044, 2022, URL: https://cds.cern.ch/record/2827111 (cit. on p. 7).

[120] ATLAS Collaboration, *Measurement of the tau lepton reconstruction and identification performance in the ATLAS experiment using pp collisions at $\sqrt{s} = 13$ TeV*, ATLAS-CONF-2017-029, 2017, URL: https://cds.cern.ch/record/2261772 (cit. on p. 7).

[121] ATLAS Collaboration, *E_T^{miss} performance in the ATLAS detector using 2015–2016 LHC pp collisions*, ATLAS-CONF-2018-023, 2018, URL: https://cds.cern.ch/record/2625233 (cit. on p. 7).

[122] ATLAS Collaboration, *Object-based missing transverse momentum significance in the ATLAS Detector*, ATLAS-CONF-2018-038, 2018, URL: https://cds.cern.ch/record/2630948 (cit. on p. 7).

[123] M. Cacciari, G. P. Salam and G. Soyez, *The catchment area of jets*, JHEP **2008** (2008) 005 (cit. on p. 8).

[124] F. Chollet et al., *Keras*, 2015, URL: https://keras.io (cit. on p. 8).

[125] H.-C. Cheng and Z. Han, *Minimal Kinematic Constraints and m(T2)*, JHEP **12** (2008) 063, arXiv: 0810.5178 [hep-ph] (cit. on p. 10).

[126] W. Verkerke and D. Kirkby, *The RooFit toolkit for data modeling*, 2003, arXiv: physics/0306116 [physics.data-an] (cit. on p. 12).

[127] L. Moneta et al., *The RooStats Project*, PoS **ACAT2010** (2010) 057, arXiv: 1009.1003 [physics.data-an] (cit. on p. 12).

[128] M. Baak et al., *HistFitter software framework for statistical data analysis*, (2014), arXiv: 1410.1280 [hep-ex] (cit. on p. 13).

[129] ATLAS Collaboration, *ATLAS simulation of boson plus jets processes in Run 2*, ATL-PHYS-PUB-2017-006, 2017, URL: https://cds.cern.ch/record/2261937 (cit. on p. 15).

[130] ATLAS Collaboration, *Multi-boson simulation for 13 TeV ATLAS analyses*, ATL-PHYS-PUB-2016-002, 2016, URL: https://cds.cern.ch/record/2119986 (cit. on p. 15).

[131] ATLAS Collaboration, *Modelling of the $t\bar{t}H$ and $t\bar{t}V(V = W, Z)$ processes for $\sqrt{s} = 13$ TeV ATLAS analyses*, ATL-PHYS-PUB-2016-005, 2016, URL: https://cds.cern.ch/record/2120826 (cit. on p. 15).

[132] G. Cowan, K. Cranmer, E. Gross and O. Vitells, *Asymptotic formulae for likelihood-based tests of new physics*, Eur. Phys. J. C **71** (2011) 1554, arXiv: 1007.1727 [physics.data-an] (cit. on p. 16), Erratum: Eur. Phys. J. C **73** (2013) 2501.

[133] A. L. Read, *Presentation of search results: the CL_S technique*, J. Phys. G **28** (2002) 2693 (cit. on p. 16).

Appendix

Table 6: Observed and expected event yields before and after fit in the high-$E_{\mathrm{T}}^{\mathrm{miss}}$ 1b regions for the $\tilde{t}_1\tilde{t}_1$ search. Uncertainties include statistical and systematic uncertainties. Uncertainties on the fitted yields are symmetric by construction, where the negative error is truncated at zero.

Channel	CR high-$E_{\mathrm{T}}^{\mathrm{miss}}$ 1b	VR high-$E_{\mathrm{T}}^{\mathrm{miss}}$ 1b	SR high-$E_{\mathrm{T}}^{\mathrm{miss}}$ 1b
Observed events	26301	3070	4327
Fitted bkg events	26300 ± 200	3300 ± 100	4280 ± 60
Fitted $t\bar{t}$ (2L)	6100 ± 300	1810 ± 80	2700 ± 100
Fitted $t\bar{t}$ (1L)	10700 ± 500	470 ± 50	270 ± 20
Fitted single-top	3200 ± 900	300 ± 80	360 ± 90
Fitted W+jets	5600 ± 700	500 ± 100	460 ± 90
Fitted $t\bar{t}Z$	63.0 ± 3.0	33.0 ± 1.0	140.0 ± 5.0
Fitted WZ	150 ± 40	10.0 ± 3.0	9.0 ± 3.0
Fitted Others	500 ± 100	130 ± 30	280 ± 70
Fitted Fakes	100 ± 60	30 ± 10	20 ± 10
Prefit SM events	25000 ± 2000	3000 ± 400	4000 ± 500
Prefit $t\bar{t}$ (2L)	6100 ± 500	1700 ± 200	2600 ± 300
Prefit $t\bar{t}$ (1L)	10000 ± 1000	500 ± 100	270 ± 40
Prefit single-top	2300 ± 400	230 ± 20	270 ± 30
Prefit W+jets	5300 ± 600	400 ± 200	400 ± 200
Prefit $t\bar{t}Z$	65.0 ± 4.0	34.0 ± 2.0	144.0 ± 8.0
Prefit WZ	150 ± 50	11.0 ± 4.0	10.0 ± 3.0
Prefit Others	500 ± 100	120 ± 40	270 ± 80
Prefit Fakes	140 ± 70	40 ± 20	30 ± 20

Table 7: Observed and expected event yields before and after fit in the high-$E_{\mathrm{T}}^{\mathrm{miss}}$ 2b regions for the $\tilde{t}_1\tilde{t}_1$ search. Uncertainties include statistical and systematic uncertainties. Uncertainties on the fitted yields are symmetric by construction, where the negative error is truncated at zero.

Channel	CR high-$E_{\mathrm{T}}^{\mathrm{miss}}$ 2b	VR high-$E_{\mathrm{T}}^{\mathrm{miss}}$ 2b	SR high-$E_{\mathrm{T}}^{\mathrm{miss}}$ 2b
Observed events	35795	2675	4110
Fitted bkg events	35800 ± 200	2630 ± 60	4090 ± 60
Fitted $t\bar{t}$ (2L)	9500 ± 400	1500 ± 60	2620 ± 100
Fitted $t\bar{t}$ (1L)	19200 ± 800	660 ± 40	500 ± 40
Fitted single-top	4000 ± 1000	290 ± 70	400 ± 100
Fitted W+jets	1900 ± 500	90 ± 40	140 ± 40
Fitted $t\bar{t}Z$	82.0 ± 3.0	27.0 ± 1.0	197.0 ± 8.0
Fitted WZ	140 ± 40	10.0 ± 3.0	8.0 ± 3.0
Fitted Others	500 ± 100	60 ± 10	160 ± 40
Fitted Fakes	120 ± 70	10.0 ± 6.0	20 ± 10
Prefit SM events	35000 ± 3000	2500 ± 300	4000 ± 500
Prefit $t\bar{t}$ (2L)	10100 ± 900	1400 ± 100	2500 ± 400
Prefit $t\bar{t}$ (1L)	19000 ± 2000	600 ± 100	510 ± 100
Prefit single-top	3300 ± 300	220 ± 40	400 ± 100
Prefit W+jets	2200 ± 800	100 ± 100	170 ± 100
Prefit $t\bar{t}Z$	83.0 ± 5.0	26.0 ± 2.0	200 ± 10
Prefit WZ	150 ± 50	10.0 ± 4.0	8.0 ± 3.0
Prefit Others	500 ± 100	50 ± 20	150 ± 50
Prefit Fakes	170 ± 90	15.0 ± 7.0	30 ± 10

Table 8: Observed and expected event yields before and after fit in the boosted 2b-1t regions for the $\tilde{t}_1\tilde{t}_1$ search. Uncertainties include statistical and systematic uncertainties. Uncertainties on the fitted yields are symmetric by construction, where the negative error is truncated when reaching to zero event yield.

Channel	CR boosted 2b-1t low-m_{T2}	CR boosted 2b-1t high-m_{T2}	VR boosted 2b-1t	SR boosted 2b-1t
Observed events	4183	127	53	18
Fitted bkg events	4180 ± 80	110.0 ± 7.0	70 ± 20	12.0 ± 1.0
Fitted $t\bar{t}$ (2L)	510 ± 70	4.0 ± 0.7	10.0 ± 3.0	2.9 ± 0.5
Fitted $t\bar{t}$ (1L)	3400 ± 100	42.0 ± 4.0	50 ± 10	4.0 ± 0.8
Fitted single-top	130 ± 50	47.0 ± 9.0	5.0 ± 1.0	1.2 ± 0.3
Fitted W+jets	26.0 ± 7.0	7.0 ± 2.0	1.3 ± 0.4	0.2 ± 0.1
Fitted $t\bar{t}Z$	7.8 ± 0.5	0.4 ± 0.1	1.3 ± 0.3	2.0 ± 0.2
Fitted WZ	1.3 ± 0.4	0.4 ± 0.1	0.01 ± 0.01	0.0 ± 0.0
Fitted Others	100 ± 20	8.0 ± 2.0	2.9 ± 0.9	1.7 ± 0.4
Fitted Fakes	20 ± 10	0.9 ± 0.5	0.3 ± 0.1	0.1 ± 0.0
Prefit SM events	4000 ± 500	110 ± 40	80 ± 30	11.0 ± 2.0
Prefit $t\bar{t}$ (2L)	370 ± 60	2.9 ± 0.6	9.0 ± 3.0	2.0 ± 1.0
Prefit $t\bar{t}$ (1L)	3400 ± 500	42.0 ± 7.0	60 ± 30	4.0 ± 2.0
Prefit single-top	90 ± 40	50 ± 30	3.0 ± 2.0	1.1 ± 0.7
Prefit W+jets	30 ± 10	8.0 ± 3.0	1.5 ± 0.6	0.2 ± 0.2
Prefit $t\bar{t}Z$	8.1 ± 0.6	0.4 ± 0.1	1.3 ± 0.3	1.9 ± 0.2
Prefit WZ	1.4 ± 0.4	0.4 ± 0.1	0.01 ± 0.01	0.0 ± 0.0
Prefit Others	90 ± 30	7.0 ± 2.0	3.0 ± 1.0	1.5 ± 0.5
Prefit Fakes	30 ± 20	1.2 ± 0.6	0.4 ± 0.2	0.1 ± 0.0

Table 9: Observed and expected event yields before and after fit in the boosted 2b-0t regions for the $\tilde{t}_1\tilde{t}_1$ search. Uncertainties include statistical and systematic uncertainties. Uncertainties on the fitted yields are symmetric by construction, where the negative error is truncated at zero.

Channel	CR boosted 2b-0t low-m_{T2}	CR boosted 2b-0t high-m_{T2}	VR boosted 2b-0t	SR boosted 2b-0t
Observed events	3526	206	93	74
Fitted bkg events	3520 ± 60	220 ± 10	90.0 ± 6.0	73.0 ± 6.0
Fitted $t\bar{t}$ (2L)	550 ± 60	11.0 ± 2.0	33.0 ± 4.0	47.0 ± 5.0
Fitted $t\bar{t}$ (1L)	2230 ± 100	48.0 ± 4.0	27.0 ± 3.0	6.0 ± 1.0
Fitted single-top	400 ± 100	90 ± 20	14.0 ± 3.0	9.0 ± 2.0
Fitted W+jets	160 ± 40	50 ± 10	11.0 ± 3.0	5.0 ± 2.0
Fitted $t\bar{t}Z$	4.8 ± 0.3	0.1 ± 0.1	0.5 ± 0.1	1.7 ± 0.2
Fitted WZ	12.0 ± 4.0	2.5 ± 0.8	0.4 ± 0.1	0.2 ± 0.1
Fitted Others	90 ± 20	10.0 ± 3.0	3.5 ± 0.9	3.9 ± 1.0
Fitted Fakes	20 ± 10	1.4 ± 0.8	0.2 ± 0.1	0.4 ± 0.2
Prefit SM events	3300 ± 200	240 ± 100	80 ± 20	60 ± 10
Prefit $t\bar{t}$ (2L)	450 ± 40	10.0 ± 2.0	28.0 ± 6.0	36.0 ± 7.0
Prefit $t\bar{t}$ (1L)	2200 ± 200	49.0 ± 7.0	28.0 ± 7.0	9.0 ± 6.0
Prefit single-top	350 ± 50	100 ± 90	12.0 ± 5.0	9.0 ± 5.0
Prefit W+jets	190 ± 80	60 ± 20	10^{+10}_{-10}	$5.0^{+6.0}_{-5.0}$
Prefit $t\bar{t}Z$	4.8 ± 0.4	0.1 ± 0.1	0.6 ± 0.2	1.9 ± 0.2
Prefit WZ	12.0 ± 4.0	2.4 ± 0.8	0.4 ± 0.2	0.2 ± 0.1
Prefit Others	90 ± 30	10.0 ± 3.0	3.0 ± 1.0	4.0 ± 1.0
Prefit Fakes	30 ± 10	2.0 ± 1.0	0.3 ± 0.2	0.5 ± 0.3

Table 10: Observed and expected event yields before and after fit in the boosted 1b-had-1t regions for the $\tilde{t}_1\tilde{t}_1$ search. Uncertainties include statistical and systematic uncertainties. Uncertainties on the fitted yields are symmetric by construction, where the negative error is truncated at zero.

Channel	CR boosted 1b-had-1t	VR boosted 1b-had-1t	SR boosted 1b-had-1t
Observed events	3093	50	24
Fitted bkg events	3130 ± 50	47.0 ± 6.0	20.0 ± 1.0
Fitted $t\bar{t}$ (2L)	290 ± 40	5.0 ± 3.0	4.7 ± 0.7
Fitted $t\bar{t}$ (1L)	2070 ± 80	23.0 ± 5.0	4.2 ± 0.5
Fitted single-top	270 ± 70	7.0 ± 2.0	2.6 ± 0.7
Fitted W+jets	390 ± 30	8.8 ± 0.9	4.3 ± 0.6
Fitted $t\bar{t}Z$	4.7 ± 0.3	0.8 ± 0.1	1.8 ± 0.2
Fitted WZ	14.0 ± 4.0	0.4 ± 0.1	0.0 ± 0.0
Fitted Others	70 ± 20	2.2 ± 0.5	2.4 ± 0.6
Fitted Fakes	20 ± 10	0.4 ± 0.3	0.2 ± 0.1
Prefit SM events	3000 ± 300	40 ± 10	17.0 ± 3.0
Prefit $t\bar{t}$ (2L)	220 ± 70	$4.0^{+5.0}_{-4.0}$	3.0 ± 2.0
Prefit $t\bar{t}$ (1L)	2100 ± 200	24.0 ± 9.0	5.0 ± 2.0
Prefit single-top	200 ± 10	5.0 ± 3.0	1.9 ± 1.0
Prefit W+jets	350 ± 20	8.0 ± 2.0	3.0 ± 1.0
Prefit $t\bar{t}Z$	4.8 ± 0.4	0.8 ± 0.1	1.8 ± 0.2
Prefit WZ	15.0 ± 5.0	0.4 ± 0.1	0.0 ± 0.0
Prefit Others	70 ± 20	2.0 ± 0.6	2.2 ± 0.7
Prefit Fakes	30 ± 10	0.6 ± 0.3	0.3 ± 0.1

Table 11: Observed and expected event yields before and after fit in the boosted 1b-had-0t regions for the $\tilde{t}_1\tilde{t}_1$ search. Uncertainties include statistical and systematic uncertainties. Uncertainties on the fitted yields are symmetric by construction, where the negative error is truncated at zero.

Channel	CR boosted 1b-had-0t	VR boosted 1b-had-0t	SR boosted 1b-had-0t
Observed events	2223	96	23
Fitted bkg events	2210 ± 50	111.0 ± 8.0	21.0 ± 3.0
Fitted $t\bar{t}$ (2L)	140 ± 20	24.0 ± 4.0	8.0 ± 2.0
Fitted $t\bar{t}$ (1L)	570 ± 30	24.0 ± 6.0	0.1 ± 0.1
Fitted single-top	210 ± 40	13.0 ± 3.0	2.9 ± 0.7
Fitted W+jets	1180 ± 60	40.0 ± 3.0	5.0 ± 2.0
Fitted $t\bar{t}Z$	2.2 ± 0.2	1.0 ± 0.1	0.6 ± 0.1
Fitted WZ	40 ± 10	1.6 ± 0.5	0.1 ± 0.0
Fitted Others	60 ± 10	7.0 ± 2.0	5.0 ± 1.0
Fitted Fakes	4.0 ± 2.0	0.4 ± 0.2	0.1 ± 0.1
Prefit SM events	2000 ± 200	90 ± 20	19.0 ± 8.0
Prefit $t\bar{t}$ (2L)	120 ± 20	21.0 ± 7.0	6.0 ± 4.0
Prefit $t\bar{t}$ (1L)	590 ± 40	20^{+20}_{-20}	0.2 ± 0.1
Prefit single-top	180 ± 70	11.0 ± 5.0	3.0 ± 2.0
Prefit W+jets	1000 ± 100	30 ± 10	$5.0^{+6.0}_{-5.0}$
Prefit $t\bar{t}Z$	2.2 ± 0.2	1.0 ± 0.1	0.6 ± 0.1
Prefit WZ	40 ± 10	1.7 ± 0.5	0.1 ± 0.0
Prefit Others	60 ± 20	7.0 ± 2.0	4.0 ± 1.0
Prefit Fakes	5.0 ± 3.0	0.5 ± 0.3	0.2 ± 0.1

Table 12: Observed and expected event yields before and after fit in the boosted 1b-lep-1t regions for the $\tilde{t}_1\tilde{t}_1$ search. Uncertainties include statistical and systematic uncertainties. Uncertainties on the fitted yields are symmetric by construction, where the negative error is truncated at zero.

Channel	CR boosted 1b-lep-1t	VR boosted 1b-lep-1t	SR boosted 1b-lep-1t
Observed events	2429	32	39
Fitted bkg events	2430 ± 70	40 ± 10	36.0 ± 3.0
Fitted $t\bar{t}$ (2L)	280 ± 40	11.0 ± 3.0	21.0 ± 3.0
Fitted $t\bar{t}$ (1L)	1640 ± 70	17.0 ± 9.0	5.0 ± 1.0
Fitted single-top	140 ± 40	2.9 ± 0.9	2.0 ± 0.5
Fitted W+jets	280 ± 30	5.5 ± 0.9	4.0 ± 0.5
Fitted $t\bar{t}Z$	2.9 ± 0.3	0.2 ± 0.1	1.4 ± 0.1
Fitted WZ	10.0 ± 3.0	0.3 ± 0.1	0.2 ± 0.1
Fitted Others	70 ± 20	1.5 ± 0.5	2.5 ± 0.6
Fitted Fakes	9.0 ± 5.0	0.1 ± 0.1	0.3 ± 0.1
Prefit SM events	2300 ± 200	40 ± 30	30 ± 10
Prefit $t\bar{t}$ (2L)	260 ± 60	9.0 ± 4.0	17.0 ± 8.0
Prefit $t\bar{t}$ (1L)	1600 ± 200	30^{+30}_{-30}	$5.0^{+8.0}_{-5.0}$
Prefit single-top	110 ± 10	3.0 ± 1.0	1.6 ± 0.9
Prefit W+jets	260 ± 20	4.6 ± 0.9	3.4 ± 0.5
Prefit $t\bar{t}Z$	3.0 ± 0.3	0.2 ± 0.1	1.4 ± 0.1
Prefit WZ	11.0 ± 3.0	0.4 ± 0.2	0.2 ± 0.1
Prefit Others	60 ± 20	1.5 ± 0.5	2.3 ± 0.7
Prefit Fakes	13.0 ± 7.0	0.2 ± 0.1	0.4 ± 0.2

Table 13: Observed and expected event yields before and after fit in the boosted 1b-lep-0t regions for the $\tilde{t}_1\tilde{t}_1$ search. Uncertainties include statistical and systematic uncertainties. Uncertainties on the fitted yields are symmetric by construction, where the negative error is truncated at zero.

Channel	CR boosted 1b-lep-0t	VR boosted 1b-lep-0t	SR boosted 1b-lep-0t
Observed events	12705	259	225
Fitted bkg events	12700 ± 200	290 ± 10	230 ± 10
Fitted $t\bar{t}$ (2L)	1900 ± 200	120 ± 10	160 ± 10
Fitted $t\bar{t}$ (1L)	5700 ± 300	61.0 ± 4.0	19.0 ± 1.0
Fitted single-top	1500 ± 400	34.0 ± 8.0	20.0 ± 5.0
Fitted W+jets	3100 ± 200	56.0 ± 7.0	22.0 ± 2.0
Fitted $t\bar{t}Z$	6.3 ± 0.7	0.2 ± 0.1	1.4 ± 0.4
Fitted WZ	120 ± 40	3.1 ± 0.9	1.0 ± 0.3
Fitted Others	340 ± 80	10.0 ± 2.0	15.0 ± 4.0
Fitted Fakes	40 ± 20	1.0 ± 0.6	0.9 ± 0.5
Prefit SM events	12000 ± 1000	240 ± 30	200 ± 20
Prefit $t\bar{t}$ (2L)	1600 ± 200	100 ± 10	130 ± 20
Prefit $t\bar{t}$ (1L)	6100 ± 800	65.0 ± 8.0	19.0 ± 3.0
Prefit single-top	1100 ± 200	26.0 ± 6.0	17.0 ± 7.0
Prefit W+jets	2700 ± 200	40 ± 20	16.0 ± 9.0
Prefit $t\bar{t}Z$	6.5 ± 0.8	$0.2^{+0.2}_{-0.2}$	1.5 ± 0.5
Prefit WZ	120 ± 40	3.1 ± 1.0	1.1 ± 0.4
Prefit Others	320 ± 100	9.0 ± 3.0	15.0 ± 5.0
Prefit Fakes	60 ± 30	1.4 ± 0.7	1.3 ± 0.7

Table 14: Observed and expected event yields before and after fit in the high-$E_{\mathrm{T}}^{\mathrm{miss}}$ 1b regions for the $t\bar{t}$+DM search. Uncertainties include statistical and systematic uncertainties. Uncertainties on the fitted yields are symmetric by construction, where the negative error is truncated at zero.

Channel	CR high-$E_{\mathrm{T}}^{\mathrm{miss}}$ 1b	VR high-$E_{\mathrm{T}}^{\mathrm{miss}}$ 1b	SR high-$E_{\mathrm{T}}^{\mathrm{miss}}$ 1b
Observed events	40315	5644	1222
Fitted bkg events	40300 ± 300	5800 ± 200	1200 ± 30
Fitted $t\bar{t}$ (2L)	9600 ± 500	3600 ± 200	620 ± 40
Fitted $t\bar{t}$ (1L)	19000 ± 1000	530 ± 60	51.0 ± 10.0
Fitted single-top	4000 ± 2000	400 ± 100	90 ± 20
Fitted W+jets	6000 ± 1000	700 ± 200	230 ± 40
Fitted $t\bar{t}Z$	75.0 ± 3.0	97.0 ± 4.0	82.0 ± 5.0
Fitted WZ	160 ± 50	13.0 ± 4.0	2.6 ± 0.8
Fitted Others	700 ± 200	300 ± 80	120 ± 30
Fitted Fakes	200 ± 100	50 ± 30	8.0 ± 4.0
Prefit SM events	39000 ± 4000	5400 ± 700	1100 ± 100
Prefit $t\bar{t}$ (2L)	9100 ± 700	3300 ± 400	570 ± 90
Prefit $t\bar{t}$ (1L)	20000 ± 2000	600 ± 200	60 ± 30
Prefit single-top	3600 ± 700	380 ± 40	80 ± 20
Prefit W+jets	5400 ± 800	600 ± 200	160 ± 70
Prefit $t\bar{t}Z$	75.0 ± 5.0	97.0 ± 6.0	82.0 ± 6.0
Prefit WZ	170 ± 50	13.0 ± 4.0	2.7 ± 0.9
Prefit Others	600 ± 200	280 ± 90	110 ± 30
Prefit Fakes	200 ± 100	70 ± 30	10.0 ± 5.0

Table 15: Observed and expected event yields before and after fit in the high-$E_{\mathrm{T}}^{\mathrm{miss}}$ 2b regions for the $t\bar{t}$+DM search. Uncertainties include statistical and systematic uncertainties. Uncertainties on the fitted yields are symmetric by construction, where the negative error is truncated at zero.

Channel	CR high-$E_{\mathrm{T}}^{\mathrm{miss}}$ 2b	VR high-$E_{\mathrm{T}}^{\mathrm{miss}}$ 2b	SR high-$E_{\mathrm{T}}^{\mathrm{miss}}$ 2b
Observed events	22033	3579	2420
Fitted bkg events	22000 ± 200	3700 ± 100	2430 ± 50
Fitted $t\bar{t}$ (2L)	7600 ± 300	2500 ± 100	1420 ± 70
Fitted $t\bar{t}$ (1L)	10700 ± 500	710 ± 70	260 ± 30
Fitted single-top	2200 ± 700	300 ± 80	310 ± 70
Fitted W+jets	1000 ± 300	100 ± 30	90 ± 40
Fitted $t\bar{t}Z$	62.0 ± 3.0	46.0 ± 2.0	186.0 ± 8.0
Fitted WZ	70 ± 20	5.0 ± 2.0	2.7 ± 0.9
Fitted Others	310 ± 80	80 ± 20	130 ± 30
Fitted Fakes	110 ± 60	30 ± 20	18.0 ± 10.0
Prefit SM events	22000 ± 2000	3600 ± 400	2400 ± 300
Prefit $t\bar{t}$ (2L)	7400 ± 600	2300 ± 300	1400 ± 200
Prefit $t\bar{t}$ (1L)	11000 ± 1000	800 ± 200	300 ± 60
Prefit single-top	1800 ± 200	270 ± 80	300 ± 100
Prefit W+jets	1100 ± 300	120 ± 80	110 ± 60
Prefit $t\bar{t}Z$	61.0 ± 3.0	45.0 ± 4.0	190 ± 10
Prefit WZ	70 ± 20	5.0 ± 2.0	2.8 ± 1.0
Prefit Others	280 ± 90	80 ± 20	120 ± 40
Prefit Fakes	130 ± 70	40 ± 20	20 ± 10

Table 16: Observed and expected event yields before and after fit in the boosted 1t CRs for the $t\bar{t}$+DM search. Uncertainties include statistical and systematic uncertainties. Uncertainties on the fitted yields are symmetric by construction, where the negative error is truncated at zero.

Channel	CR boosted 1b-had-1t	CR boosted 1b-lep-1t	CR boosted 2b-1t low-m_{T2}	CR boosted 2b-1t high-m_{T2}
Observed events	3093	2429	4183	127
Fitted bkg events	3110 ± 50	2430 ± 60	4190 ± 80	108.0 ± 7.0
Fitted $t\bar{t}$ (2L)	290 ± 50	290 ± 50	520 ± 90	4.0 ± 0.9
Fitted $t\bar{t}$ (1L)	2080 ± 80	1640 ± 80	3400 ± 100	44.0 ± 4.0
Fitted single-top	230 ± 70	120 ± 40	110 ± 50	44.0 ± 9.0
Fitted W+jets	400 ± 30	290 ± 30	27.0 ± 7.0	8.0 ± 2.0
Fitted $t\bar{t}Z$	4.6 ± 0.4	2.9 ± 0.3	7.7 ± 0.5	0.4 ± 0.1
Fitted WZ	14.0 ± 4.0	10.0 ± 3.0	1.3 ± 0.4	0.4 ± 0.1
Fitted Others	70 ± 20	70 ± 20	90 ± 20	8.0 ± 2.0
Fitted Fakes	20 ± 10	11.0 ± 6.0	30 ± 20	1.0 ± 0.6
Prefit SM events	3000 ± 300	2300 ± 200	4000 ± 500	110 ± 40
Prefit $t\bar{t}$ (2L)	220 ± 70	260 ± 60	370 ± 60	2.9 ± 0.6
Prefit $t\bar{t}$ (1L)	2100 ± 200	1600 ± 200	3400 ± 500	42.0 ± 7.0
Prefit single-top	200 ± 20	110 ± 10	90 ± 30	50 ± 40
Prefit W+jets	350 ± 20	260 ± 20	30 ± 10	8.0 ± 3.0
Prefit $t\bar{t}Z$	4.8 ± 0.4	3.0 ± 0.3	8.1 ± 0.6	0.4 ± 0.1
Prefit WZ	15.0 ± 5.0	11.0 ± 3.0	1.4 ± 0.4	0.4 ± 0.1
Prefit Others	70 ± 20	60 ± 20	90 ± 30	7.0 ± 2.0
Prefit Fakes	30 ± 10	13.0 ± 7.0	30 ± 20	1.2 ± 0.6

Table 17: Observed and expected event yields before and after fit in the boosted 0t CRs for the $t\bar{t}$+DM search. Uncertainties include statistical and systematic uncertainties. Uncertainties on the fitted yields are symmetric by construction, where the negative error is truncated at zero.

Channel	CR boosted 1b-had-0t	CR boosted 1b-lep-0t	CR boosted 2b-0t low-m_{T2}	CR boosted 2b-0t high-m_{T2}
Observed events	2223	12705	3526	206
Fitted bkg events	2210 ± 50	12700 ± 200	3510 ± 60	220 ± 10
Fitted $t\bar{t}$ (2L)	150 ± 20	1900 ± 300	570 ± 90	12.0 ± 2.0
Fitted $t\bar{t}$ (1L)	570 ± 30	5800 ± 400	2300 ± 100	48.0 ± 4.0
Fitted single-top	180 ± 40	1300 ± 400	400 ± 100	90 ± 20
Fitted W+jets	1210 ± 60	3200 ± 200	170 ± 50	50 ± 10
Fitted $t\bar{t}Z$	2.2 ± 0.2	6.3 ± 0.7	4.8 ± 0.3	0.1 ± 0.1
Fitted WZ	30 ± 10	120 ± 40	12.0 ± 4.0	2.5 ± 0.8
Fitted Others	60 ± 20	340 ± 90	90 ± 20	10.0 ± 3.0
Fitted Fakes	4.0 ± 2.0	50 ± 20	20 ± 10	1.6 ± 0.9
Prefit SM events	2000 ± 200	12000 ± 1000	3300 ± 200	240 ± 100
Prefit $t\bar{t}$ (2L)	120 ± 20	1600 ± 200	450 ± 40	10.0 ± 2.0
Prefit $t\bar{t}$ (1L)	590 ± 40	6100 ± 800	2200 ± 200	49.0 ± 7.0
Prefit single-top	180 ± 80	1100 ± 200	350 ± 70	100 ± 90
Prefit W+jets	1000 ± 100	2700 ± 200	190 ± 80	60 ± 20
Prefit $t\bar{t}Z$	2.2 ± 0.2	6.5 ± 0.8	4.8 ± 0.4	0.1 ± 0.1
Prefit WZ	40 ± 10	120 ± 40	12.0 ± 4.0	2.4 ± 0.8
Prefit Others	60 ± 20	320 ± 100	90 ± 30	10.0 ± 3.0
Prefit Fakes	5.0 ± 3.0	60 ± 30	30 ± 10	2.0 ± 1.0

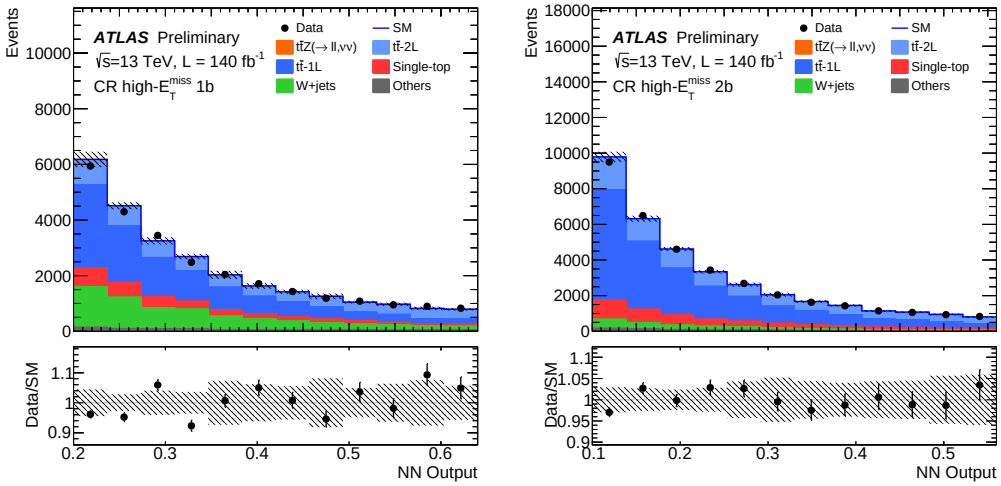


Figure 12: Observed and expected events as a function of the stop-NN output value in the high-$E_{\mathrm{T}}^{\mathrm{miss}}$ 1b (left) and 2b (right) CRs. The expected distributions are as determined in the fit for the $\tilde{t}_1\tilde{t}_1$ search. The hatched area around the total SM prediction includes statistical and systematic uncertainties. The lower panels show the ratio of the observed data to the expected SM events.

Table 18: Lists of input features for the stop- and DM-NNs in each event category. Inputs are defined in text. † The p_{T} of the $t\bar{t} + E_{\mathrm{T}}^{\mathrm{miss}}$ system is before Lorentz boost.

Category		High-$E_{\mathrm{T}}^{\mathrm{miss}}$		Boosted		
		1b	2b	1b-lep-1t (-0t)	1b-had-1t (-0t)	2b-1t (-0t)
$E_{\mathrm{T}}^{\mathrm{miss}}$	E	✓	✓	✓	✓	✓
	Significance	✓	✓	✓	✓	✓
$\mathrm{top_{had}}$	p_x	✓	✓	✓	✓	✓
	p_y	✓	✓	✓	✓	✓
	p_z	✓	✓	✓	✓	✓
	E	✓	✓	✓	✓	✓
	top-NN output	✓	✓			
$\mathrm{top_{lep}}$	p_z	✓	✓	✓	✓	✓
	E	✓	✓	✓	✓	✓
$t\bar{t} + E_{\mathrm{T}}^{\mathrm{miss}}$	p_{T} †	✓	✓	✓	✓	✓
lepton	p_x	✓	✓	✓	✓	✓
	p_y	✓	✓	✓	✓	✓
	p_z	✓	✓	✓	✓	✓
	E	✓	✓	✓	✓	✓
b-jet in $\mathrm{top_{had}}$	p_x	✓	✓		✓	✓
	p_y	✓	✓		✓	✓
	p_z	✓	✓		✓	✓
	E	✓	✓		✓	✓
b-jet in $\mathrm{top_{lep}}$	p_x	✓	✓	✓	✓	✓
	p_y	✓	✓	✓	✓	✓
	p_z	✓	✓	✓	✓	✓
	E	✓	✓	✓	✓	✓
$\Delta R(b, b)$		✓	✓	✓	✓	✓
$m_T(\ell, E_{\mathrm{T}}^{\mathrm{miss}})$		✓	✓	✓	✓	✓
$m_{T2}(b, b, E_{\mathrm{T}}^{\mathrm{miss}})$		✓	✓			
$m_{T2,\mathrm{min}}(b + \ell, b, E_{\mathrm{T}}^{\mathrm{miss}})$		✓	✓			

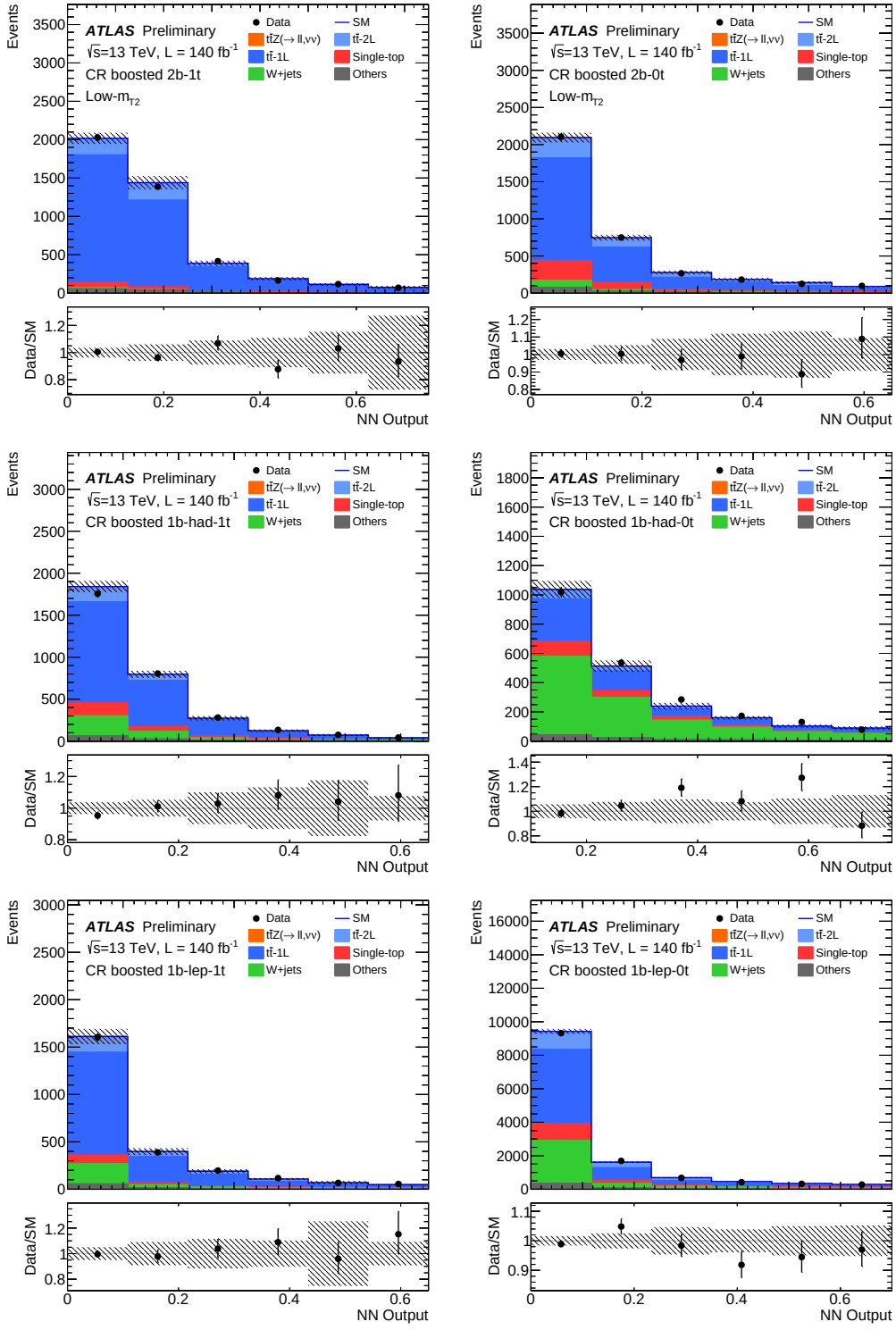


Figure 13: Observed and expected events as a function of the stop-NN output value in the boosted 2b (top) and 1b (middle row) and 1b-lep (bottom) CRs. Plots on the left are for 1t regions and plots on the right are for 0t regions. The expected distributions are as determined in the fit for the $\tilde{t}_1\tilde{t}_1$ search. The hatched area around the total SM prediction includes statistical and systematic uncertainties. The lower panels show the ratio of the observed data to the expected SM events.

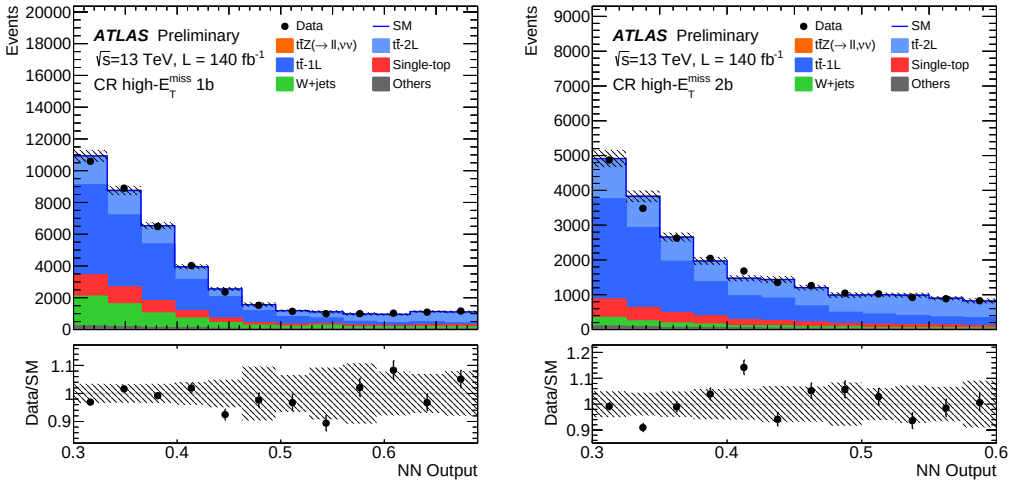


Figure 14: Observed and expected events as a function of the DM-NN output value in the high-$E_{\mathrm{T}}^{\mathrm{miss}}$ 1b (left) and 2b (right) CRs. The expected distributions are as determined in the fit for the $t\bar{t}$+DM search. The hatched area around the total SM prediction includes statistical and systematic uncertainties. The lower panels show the ratio of the observed data to the expected SM events.