ATLAS CONF Note

ATLAS-CONF-2023-048

August 20, 2023

# Search for pair production of higgsinos in events with two Higgs bosons and missing transverse momentum in $\sqrt{s} = 13$ TeV $pp$ collisions at the ATLAS experiment

The ATLAS Collaboration

This paper presents a search for pair production of higgsinos, the supersymmetric partners of the Higgs bosons, in scenarios with gauge-mediated supersymmetry breaking. Each higgsino is assumed to decay into a Higgs boson and a nearly-massless gravitino. The search targets the decay of the two Higgs bosons to $b\bar{b}$, leading to a reconstructed final state with at least three energetic $b$-jets and missing transverse momentum. Two complementary analysis channels are used, with each channel specifically targeting either low or high values of the higgsino mass. The low (high) mass channel exploits 126 (139) fb$^{-1}$ of $\sqrt{s} = 13$ TeV data collected by the ATLAS detector during Run 2 of the Large Hadron Collider. No significant excess above the Standard Model prediction is found. At 95% confidence level, masses up to 940 GeV are excluded for higgsinos decaying exclusively to Higgs bosons. Exclusion limits as a function of the higgsino decay branching ratio to a $Z$ or a Higgs boson are also reported.

ATLAS-CONF-2023-048
12 September 2023

# 1 Introduction

Supersymmetry (SUSY) [1–6] is an extension of the Standard Model (SM) predicting the existence of a bosonic (fermionic) partner for each fermionic (bosonic) particle of the SM. When $R$-parity is conserved [7], the lightest supersymmetric particle (LSP) is not allowed to decay into lighter SM particles and consequently is stable, making it a potential dark matter candidate. The hierarchy mass problem of the Higgs boson could be resolved by supersymmetry due to the cancellation of the divergent Higgs mass diagrams by their SUSY counterparts [8–11]. This class of "natural" SUSY models requires that the superpartners of the top and bottom quarks (*i.e.* the stop, $\tilde{t}$, and sbottom, $\tilde{b}$), the gluon (*i.e.* the gluino, $\tilde{g}$) and of an extended Higgs sector (*i.e.* the higgsinos, $\tilde{H}$) are light. Such particles would be expected to be abundantly produced in the proton–proton ($pp$) collisions of the Large Hadron Collider (LHC). While the ATLAS and CMS collaborations have set strong limits on the masses of the gluino and the stop particles, the exclusion limits on the higgsino masses are much less stringent [12–17], raising the prospect that the higgsino might be the first SUSY particle to be detected at the LHC.

This paper presents a search for pair production of higgsinos in models of general gauge mediation (GGM) [18–22] or gauge-mediated symmetry breaking (GMSB) [23, 24]. In these models, the lightest neutralino, $\tilde{\chi}_1^0$, the neutral particle resulting from mixing between the SUSY partners of the SM electroweak bosons, is the next-to-lightest SUSY particle (NLSP), while the LSP is the particle associated to the spontaneous breaking of the global supersymmetry. The $\tilde{\chi}_1^0$ in these models is dominated by the higgsino component and treated as a pure higgsino. In many GMSB models where the SUSY breaking is mediated at low energy, the LSP is nearly massless. When SUSY is promoted to a local symmetry, the LSP is absorbed by the superpartner of the graviton, the gravitino, $\tilde{G}$. This allows the higgsino, which is produced via mass degenerate pairs of charginos (the charged particles resulting from mixing between the SUSY partners of the SM electroweak bosons) or neutralinos, to decay into a SM Higgs boson and a nearly massless gravitino. The $\tilde{H} \rightarrow h + \tilde{G}$ mode dominates when $m_{\tilde{H}}$ is greater than the Higgs mass and when $\tan\beta$ (the ratio of the vacuum expectation values of the Higgs doublets) is small [25]. This specific scenario is implemented in the search through the simplified model represented in Figure 1. The $\tilde{G}$ of the resulting model is effectively massless, with the mass set to 1 MeV for this analysis. The only free parameter in the model is the mass of the degenerate higgsino states, $m_{\tilde{H}}$.

In this search, higgsinos are assumed to be produced in pairs, resulting in an experimental signature including two SM Higgs bosons and missing transverse momentum ($E_{\mathrm{T}}^{\mathrm{miss}}$). Due to its high branching ratio, the $h \rightarrow b\bar{b}$ decay channel is an ideal target for this model, resulting in a final state with multiple $b$-jets and $E_{\mathrm{T}}^{\mathrm{miss}}$. The search is conducted in two complementary channels, each specifically targeting the production of either two high- or low-mass higgsinos. Depending on the mass of the higgsino, the experimental signature can differ significantly, which motivates different experimental approaches. The high-mass channel is characterised by significant $E_{\mathrm{T}}^{\mathrm{miss}}$ in the final state, and relies on $E_{\mathrm{T}}^{\mathrm{miss}}$-based triggers [26]. The low-mass channel employs a combination of $b$-jet triggers [27] due to the significantly reduced $E_{\mathrm{T}}^{\mathrm{miss}}$ produced by low-mass higgsinos. The transition between which of the two channels is more sensitive for a higgsino decay branching ratio of $\mathrm{BR}(\tilde{H} \rightarrow h + \tilde{G}) = 100\%$ occurs near $m_{\tilde{H}} = 250\,\mathrm{GeV}$. The low-mass channel is used for results below this mass point while the high-mass channel is used for results at or above this mass point. This strategy is used as the channels are not orthogonal and the sensitivity of the high-mass channel decreases rapidly as $m_{\tilde{H}}$ falls below 250 GeV. A similar search using the full Run 2 dataset was performed by the CMS Collaboration [13]. Compared to the previous ATLAS search using 24.3 –36.1 $\mathrm{fb}^{-1}$ of Run 2 data [12], this search includes multiple improvements beyond the increase of the size of the dataset. These include a new method for pairing $b$-jets into Higgs-boson candidates, significantly

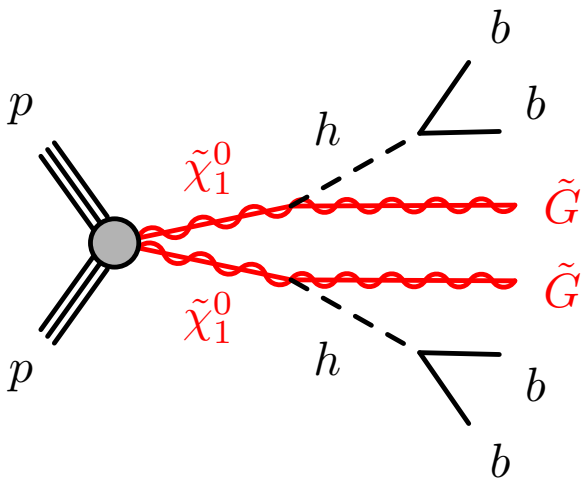



Figure 1: The signal diagram of the simplified SUSY model targeted by this analysis.

improved jet reconstruction and $b$-tagging, optimized $b$-tagging requirements in the low-mass channel and improved signal-to-background discrimination in the high-mass channel through the use of multivariate techniques.

This paper is organized as follows: Section 2 describes the ATLAS detector, while Section 3 describes the data and simulated samples. Section 4 explains the objects and reconstruction inputs of the channels, Section 5 the event selection and background estimation and Section 6 the systematic uncertainties associated with this search. The results and statistical interpretation are reported in Sections 7 and 8 respectively. The conclusions of the search are described in Section 9.

## 2 ATLAS detector

The ATLAS experiment [28] at the LHC is a multipurpose particle detector with a forward–backward symmetric cylindrical geometry and a near $4\pi$ coverage in solid angle.[1] It consists of an inner detector (ID) surrounded by a thin superconducting solenoid providing a 2 T axial magnetic field, electromagnetic and hadron calorimeters, and a muon spectrometer. The ID covers the pseudorapidity range $|\eta| < 2.5$ and it consists of silicon pixel, silicon microstrip, and transition radiation tracking detectors. Lead/liquid-argon (LAr) sampling calorimeters provide electromagnetic (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 muon spectrometer (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.

[1] ATLAS uses a right-handed coordinate system with its origin at the nominal interaction point (IP) in the center of the detector and the $z$-axis along the beam pipe. The $x$-axis points from the IP to the center of the LHC ring, and the $y$-axis points upwards. Cylindrical coordinates $(r, \phi)$ are used in the transverse plane, $\phi$ being the azimuthal angle around the $z$-axis. The pseudorapidity is defined in terms of the polar angle $\theta$ as $\eta = -\ln\tan(\theta/2)$. Angular distance is measured in units of $\Delta R \equiv \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$.

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 "high-level" trigger that reduces the accepted event rate to 1 kHz on average depending on the data-taking conditions. An extensive software suite [29] is used in the reconstruction and analysis of real and simulated data, in detector operations, and in the trigger and data acquisition systems of the experiment.

## 3 Data and simulated samples

The data used in this search were collected by the ATLAS detector during 2015-2018 from the $pp$ collisions produced by the LHC during its Run 2. During this period, the LHC collided proton bunches at a center-of-mass energy of $\sqrt{s} = 13$ TeV with a bunch-crossing separation of 25 ns. The high-mass channel uses the complete Run 2 dataset of $pp$ collisions, corresponding to an integrated luminosity of 139 fb$^{-1}$ after requiring all detector subsystems to be operational and recording good quality data [30]. The data for this channel were collected with a combination of $E_{\mathrm{T}}^{\mathrm{miss}}$ triggers. The low-mass channel uses data collected through a combination of $b$-jet triggers, enabling it to focus on signals that produce lower $E_{\mathrm{T}}^{\mathrm{miss}}$ compared to the high-mass channel, equating to a lower luminosity of 126 fb$^{-1}$ due to operational issues associated with the $b$-jet trigger selections during part of the period. The trigger selections, referred to as 'online' selections, and the fully reconstructed selections, referred to as 'offline', are presented in Table 1. The offline thresholds are tighter than the online ones in order to select events with a well understood trigger efficiency. Additional offline selections are specifically employed in the low-mass channel to provide independent event samples for each trigger employed. This approach decouples between triggers any efficiency corrections that may be needed due to differences between data and Monte Carlo (MC) performance.

MC simulations are used to model the signals and the high-mass channel background processes of this search. In addition, MC simulations of dijet processes are used to validate the background modeling for the high-mass channel. The samples are detailed further in Table 2.

The effect of multiple interactions in the same and neighboring bunch crossings (pileup) was modeled by overlaying the simulated hard-scattering event with inelastic $pp$ events generated with Pythia 8.186 [31] using the NNPDF2.3lo set of PDFs [32] and the A3 set of tuned parameters [33]. All samples using the Pythia parton shower have the decays of bottom and charm hadrons performed by EvtGen [34]. In the production of $t\bar{t}$ events, the $h_{\mathrm{damp}}$ parameter[2] is set to $1.5\,m_{\mathrm{top}}$ [35]. In samples produced with the Sherpa generator, the matrix element calculations were matched and merged with the Sherpa parton shower based on Catani–Seymour dipole factorization [36, 37] using the MEPS@NLO prescription [38–41]. The virtual QCD corrections were provided by the OpenLoops library [42–44]. In the signal simulations, matrix elements (ME) for higgsino pairs were generated with up to two additional partons. Signal cross sections are calculated to next-to-leading order (NLO) in the strong coupling constant, adding the resummation of soft gluon emission at next-to-leading-logarithmic accuracy (NLO+NLL) [45–50].

While simulations are used to estimate the high-mass channel backgrounds, the low-mass channel background is dominated by multijet processes that are not reliably modeled in simulation. A fully data-driven technique is used for the background estimation of the low-mass channel as described in Section 5.

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

Table 1: Online and offline selections used for the high- and low-mass channels of the analysis. The second column, entitled 'Period', refers to the year the data was recorded. The offline selections listed for the low-mass channel are only those required to ensure orthogonality between different trigger selections. The $H_{\mathrm{T}}$ variable corresponds to the scalar sum of the $p_{\mathrm{T}}$ of jets in the event. When a $p_{\mathrm{T}}$ selection is listed for multiple jets, it is applied to each jet. $p_{\mathrm{T},j1}$ is the $p_{\mathrm{T}}$ of the leading jet in an event.

| Category | Period | Online selections | Offline selections |
|---|---|---|---|
| Low-mass channel | | | |
| $2b1\mathrm{j}$ | 2016 | 1 jet ($p_{\mathrm{T}} > 100\,\mathrm{GeV}$), 2 $b$-jets (60% $b$-jet efficiency, $p_{\mathrm{T}} > 55\,\mathrm{GeV}$) | $p_{\mathrm{T},j1} > 150\,\mathrm{GeV}$ |
| | 2017 | 1 jet ($p_{\mathrm{T}} > 150\,\mathrm{GeV}$), 2 $b$-jets (70% $b$-jet efficiency, $p_{\mathrm{T}} > 55\,\mathrm{GeV}$) | $p_{\mathrm{T},j1} > 350\,\mathrm{GeV}$ |
| | 2018 | | $p_{\mathrm{T},j1} > 500\,\mathrm{GeV}$ |
| $2bH_{\mathrm{T}}$ | 2017 | $H_{\mathrm{T}} > 300\,\mathrm{GeV}$, 2 $b$-jets (50% $b$-jet efficiency, $p_{\mathrm{T}} > 55\,\mathrm{GeV}$) | $p_{\mathrm{T},j1} < 350\,\mathrm{GeV}$, $H_{\mathrm{T}} > 850\,\mathrm{GeV}$ |
| | 2018 | | $p_{\mathrm{T},j1} < 500\,\mathrm{GeV}$, $H_{\mathrm{T}} > 700\,\mathrm{GeV}$ |
| $2b2\mathrm{j}$ | 2016 | 2 jets ($p_{\mathrm{T}} > 35\,\mathrm{GeV}$), 2 $b$-jets (60% $b$-jet efficiency, $p_{\mathrm{T}} > 35\,\mathrm{GeV}$) | $p_{\mathrm{T},j1} < 150\,\mathrm{GeV}$ |
| | 2017 | 2 jets ($p_{\mathrm{T}} > 35\,\mathrm{GeV}$), 2 $b$-jets (40% $b$-jet efficiency, $p_{\mathrm{T}} > 35\,\mathrm{GeV}$) | $p_{\mathrm{T},j1} < 350\,\mathrm{GeV}$, $H_{\mathrm{T}} < 850\,\mathrm{GeV}$ |
| | 2018 | 2 jets ($p_{\mathrm{T}} > 35\,\mathrm{GeV}$), 2 $b$-jets (60% $b$-jet efficiency, $p_{\mathrm{T}} > 35\,\mathrm{GeV}$) | $p_{\mathrm{T},j1} < 500\,\mathrm{GeV}$, $H_{\mathrm{T}} < 700\,\mathrm{GeV}$ |
| High-mass channel | | | |
| $E_{\mathrm{T}}^{\mathrm{miss}}$ | 2015 | $E_{\mathrm{T}}^{\mathrm{miss}} > 70\,\mathrm{GeV}$ | $E_{\mathrm{T}}^{\mathrm{miss}} > 150\,\mathrm{GeV}$ |
| | 2016 | $E_{\mathrm{T}}^{\mathrm{miss}} > 90\,\mathrm{GeV}$ | |
| | 2017 | $E_{\mathrm{T}}^{\mathrm{miss}} > 100\,\mathrm{GeV}$ | |
| | 2018 | $E_{\mathrm{T}}^{\mathrm{miss}} > 110\,\mathrm{GeV}$ | |

# 4 Object reconstruction

Charged particle tracks are required to have $p_{\mathrm{T}} > 0.5\,\mathrm{GeV}$. Primary vertices are reconstructed from at least two charged-particle tracks [81]. The hard-scatter interaction is identified as the vertex with the largest sum of squared track $p_{\mathrm{T}}$ ($\sum p_{\mathrm{T,track}}^2$).

Small-radius jets are reconstructed using the anti-$k_t$ algorithm [82] with a radius parameter of $R = 0.4$ and using particle-flow objects as inputs. These objects are created from calorimeter clusters and ID tracks through the subtraction of the energy deposited by well-measured tracks in the calorimeter to enhance the resolution of the combined energy measurement [83]. Jets produced from collisions other than the hard-scatter vertex (pileup jets) are removed by testing their compatibility with the hard-scatter interaction using the Jet Vertex Tagger (JVT) discriminant [84]. Jets with $p_{\mathrm{T}} < 60\,\mathrm{GeV}$ and $|\eta| < 2.4$ are required to pass the "Medium" JVT working point. Different MC-based calibration steps are applied to the reconstructed jets [85], including an area-based correction to account for pileup jet energy dependency, a $p_{\mathrm{T}}$- and $\eta$-dependent calibration to match the truth energy scale of the jets and the Global Sequential Calibration (GSC) to minimize energy calibration differences between quark- and gluon-initiated jets. Finally, an *in-situ* calibration is applied to jets in data to match the energy scale in simulation. Signal scenarios where jets are close to each other due to large higgsino-gravitino mass splittings gain sensitivity with the use of large-radius jets produced by reclustering $R = 0.4$ jets [86] through another iteration of the anti-$k_t$ algorithm with a radius parameter of $R = 0.8$. Calibrations are propagated through that iteration.

Table 2: List of generators used for the different processes considered in this publication. Information is given about the underlying-event tunes, the PDF sets and the pQCD highest-order accuracy used for the normalization of the different samples. Henceforth, the $t\bar{t}W$, $t\bar{t}Z$, 4-top and $t\bar{t}h$ processes are grouped into a single category, denoted $t\bar{t}+X$. The dijet samples are only used for validation.

| Process | Generator + fragmentation/hadronization | Tune | PDF set | Order of cross section |
|---|---|---|---|---|
| **SUSY signals** | MadGraph v2.6.1/v2.6.2 [51] + Pythia v8.230 [62] | A14 [52] | NNPDF2.3 [32] | $\text{NNLO}_{\text{approx}}$+NNLL [53–61] |
| $t\bar{t}$ | Powheg Box v2 [63–66] + Pythia v8.230 | A14 | NNPDF3.0 (ME) [67] NNPDF2.3 (UE) | NNLO+NNLL [68–74] |
| **Single top** | Powheg Box v2 + Pythia v8.230 | A14 | NNPDF3.0 (ME) NNPDF2.3 (UE) | NLO [75] ($t/s$-channel) NLO+NNLL [76, 77] ($Wt$) |
| $t\bar{t}W/t\bar{t}Z$ | MadGraph5_aMC@NLO v2.3.3 [51] + Pythia-8.210 | A14 | NNPDF3.0 (ME) NNPDF2.3 (UE) | NLO [78] |
| $t\bar{t}t\bar{t}$ | MadGraph5_aMC@NLO v2.2.2 + Pythia v8.186 | A14 | NNPDF2.3 | NLO [51] |
| $t\bar{t}h$ | Powheg Box v2 + Pythia v8.230 | A14 | NNPDF3.0 (ME) NNPDF2.3 (UE) | NLO [78] |
| **Dibosons** $WW, WZ, ZZ$ | Sherpa v2.2.1 [36–44, 79] | Default | NNPDF3.0 | NLO [39, 80] |
| $W/Z$**+jets** | Sherpa v2.2.1 | Default | NNPDF3.0 | NNLO [42–44] |
| **Dijet** | Pythia v8.230 | A14 | NNPDF2.3 | LO |

The reclustered jets are trimmed [87] by removing small-radius jets whose $p_{\text{T}}$ falls below $f_{\text{cut}} = 10\%$ of the $p_{\text{T}}$ of the large-radius jet. After this procedure, reclustered jets are required to have $p_{\text{T}} > 100\,\text{GeV}$ and $|\eta| < 2.0$. The high-mass channel requires $R = 0.4$ jets to have $p_{\text{T}} > 25\,\text{GeV}$ and $|\eta| < 2.8$, while the low-mass channel uses jets with $p_{\text{T}} > 40\,\text{GeV}$ and $|\eta| < 2.8$ due to requirements from its jet-based trigger strategy. The high-mass channel uses as a discriminant the total mass of the large-radius jets in the event, as explained later in Section 5.1.1.

Small-radius jets initiated by $b$-quarks and decaying within the ID acceptance ($|\eta| < 2.5$) are identified as $b$-tagged jets using the `DL1r` classifier set to a working point of 77% efficiency for simulated $t\bar{t}$ events [88]. This classification algorithm uses a selection of inputs, including information about the impact parameters of ID tracks, the presence of displaced secondary vertices and the reconstructed flight paths of $b$- and $c$-hadrons inside the jet. At the selected working point, the light-jet (charm-jet) rejection factor measured in $t\bar{t}$ events is approximately 130 (4.9) [88–91]. Correction factors are applied to the simulated samples to account for differences in the $b$-tagging efficiencies between data and simulation. For the low-mass channel, additional correction factors are applied to account for differences in the online $b$-tagging efficiencies. Correlations between the trigger and offline correction factors are taken into account.

Two types of electrons and muons, "loose" and "signal-quality", are defined for this analysis. Electron candidates are built from energy deposits in the EM calorimeter that are matched to ID tracks [92]. Loose electrons are required to pass the *LooseLH* likelihood identification criteria [93] and to have $p_{\text{T}} > 7\,\text{GeV}$

and $|\eta| < 2.47$. Further rejection of fakes and non-prompt sources is achieved by requiring electron tracks to match the hard-scatter vertex through a cut of the longitudinal impact parameter $|z_0 \sin\theta| < 0.5$ mm. Loose electrons with $p_{\mathrm{T}} > 20$ GeV, called "baseline" electrons, are used to calculate the $E_{\mathrm{T}}^{\mathrm{miss}}$. To be considered as "signal-quality" objects, candidates are required to survive the overlap removal procedure, pass the *MediumLH* likelihood identification criteria [93], pass the *Loose* isolation requirements and have $p_{\mathrm{T}} > 20$ GeV and $|\eta| < 2.47$. Signal-quality electrons are also required to have a transverse impact parameter significance $|d_0|/\sigma(d_0) < 3$.

Muon candidates are reconstructed by matching the ID track with the MS tracks or performing a combined fit of ID tracks with the aligned individual hits found in the MS. After reconstruction, loose muons are required to have $p_{\mathrm{T}} > 6$ GeV, $|\eta| < 2.7$ and to pass the *Medium* identification requirement based on track quality variables [94]. Further rejection of fakes and non-prompt sources is achieved by requiring muon tracks to match the hard-scatter vertex through a cut of the longitudinal impact parameter $|z_0 \sin\theta| < 0.5$ mm. Loose muons with $p_{\mathrm{T}} > 20$ GeV, called "baseline" muons, are used to calculate the $E_{\mathrm{T}}^{\mathrm{miss}}$ and to correct the four-momentum of jets to account for semi-leptonic $b$-hadron decays. This correction consists in the addition of the muon four-momentum to the jet if a muon is found within $\Delta R < 0.4$ of that jet. Signal-quality muons are the subset of loose muons that survive the overlap removal procedure, pass the *TightTrackOnly* (with variable radius) isolation requirements [94] and have $p_{\mathrm{T}} > 20$ GeV and $|\eta| < 2.5$. Signal-quality muons are also required to have a transverse impact parameter significance $|d_0|/\sigma(d_0) < 3$.

To resolve the reconstruction ambiguities between electron, muons and small-radius jets, an overlap removal procedure is applied. First, any baseline electron sharing an ID track with a baseline muon is rejected. Then, if a jet is found to be within the $\Delta R(j, e) < 0.2$ range with respect to a baseline electron, the jet is removed. If a baseline electron is found within $\Delta R(j, e) < \min(0.4, 0.04 + 10\,\mathrm{GeV}/p_{\mathrm{T}}^{e})$ of a remaining jet, where $p_{\mathrm{T}}^{e}$ is the transverse momentum of the electron, the electron is removed. Next, any jet with an associated muon ID track or a baseline muon within $\Delta R(j, \mu) < 0.2$ is removed if the jet has no more than three tracks. Lastly, the muon is removed if $\Delta R(j, \mu) < \min(0.4, 0.04 + 10\,\mathrm{GeV}/p_{\mathrm{T}}^{\mu})$ for any remaining jet, where $p_{\mathrm{T}}^{\mu}$ is the transverse momentum of the muon.

The missing transverse momentum, $p_{\mathrm{T}}^{\mathrm{miss}}$, and its magnitude $E_{\mathrm{T}}^{\mathrm{miss}}$, are built from the negative vector sum of the transverse momenta of all the well-identified and calibrated physics objects in the event, plus an extra term accounting for low energy particles and commonly known as the $E_{\mathrm{T}}^{\mathrm{miss}}$ soft term [95]. The soft term is calculated from ID tracks matched to the primary vertex and not associated to any physics object, and known as the track-based soft term (TST). Baseline identification criteria, which are looser than the signal-quality criteria described above, are applied for muons and electrons in these calculations.

## 5 Event selection and background estimation

Prior to any channel-specific selections, both channels impose data quality requirements to ensure that only events recorded when the entire ATLAS detector was functioning well are used [30]. These selections reject events containing data corruption in the ID and calorimeters, as well as spurious jets caused by non-collision backgrounds [96, 97]. Events containing signal-quality leptons (electrons or muons) are discarded to reduce backgrounds containing leptonically decaying $W$ bosons. Events are also discarded if they contain more than one loose lepton with $p_{\mathrm{T}} > 8$ GeV, where a single loose lepton is allowed to avoid rejecting events containing a semi-leptonic $b$-hadron decay. The loose lepton criteria employed also minimize overlaps with ATLAS analyses targeting the same signal using leptonic final states [14, 16].

### 5.1 High-mass channel

The high-mass channel focuses on detecting final states characterized by high $E_{\mathrm{T}}^{\mathrm{miss}}$, a minimum of three $b$-jets, and no signal-quality leptons. It relies on reconstructing the Higgs bosons resulting from the decay of higgsinos. To estimate the main backgrounds, MC simulations are used, with adjustments made to the normalization of $t\bar{t}$ and $Z$+jets processes derived from data control samples. The contribution from QCD multijet events is estimated through a data-driven approach involving a neural-network-assisted reweighting procedure. A Boosted Decision Tree (BDT) is employed to distinguish between the signal and background events. A set of higgsino mass points are considered, referred to as "mass hypotheses", and for each of them control regions (CRs), validation regions (VRs) and up to four distinct signal regions (SRs) are defined based on the BDT score. The mass hypotheses are given by:

$$m_{\tilde{H}} = \{200, 250, 300, 400, 500, 600, 700, 800, 900, 1000, 1100\}\ \text{GeV}$$

In addition, the three of these signal regions that have the highest significance for the 250 GeV, 500 GeV and 1000 GeV mass hypotheses, denoted 'discovery signal regions', are used to search for and set limits on an excess of beyond Standard Model (BSM) events in this phase space in a more model-independent manner.

#### 5.1.1 Event selection

After the common selections described above and the trigger selections described in Section 3, signal-like events are required to satisfy the following 'standard preselection' requirements:

- between four and seven small-radius jets with $p_{\mathrm{T}} > 25$ GeV are reconstructed in the event, to reduce backgrounds with a large number of additional jets;
- at least three of these jets are $b$-tagged according to the requirements described in Section 4, as expected by the signal topology;
- $E_{\mathrm{T}}^{\mathrm{miss}}$ is greater than 150 GeV, for consistency with the production of invisible particles;
- the minimum difference of azimuthal angle between $E_{\mathrm{T}}^{\mathrm{miss}}$ and any of the four leading jets ($\Delta\phi_{\mathrm{min}}^{4j}$) is greater than 0.4, to reduce backgrounds with spuriously large $E_{\mathrm{T}}^{\mathrm{miss}}$ resulting from mis-measurement of the momentum of a jet.

Scaling factors are applied to MC simulations to correct for discrepancies between simulated- and data-based trigger efficiencies. Such corrections are negligible for events with $E_{\mathrm{T}}^{\mathrm{miss}} \geq 200$ GeV and reach a maximum of about 10% for events with $E_{\mathrm{T}}^{\mathrm{miss}} \simeq 150$ GeV and scalar sum of jet $p_{\mathrm{T}}$ below 250 GeV.

A key element of this analysis is the identification of the Higgs bosons originating from the higgsino decays. This is essential as the masses of the higher- and lower-mass Higgs boson candidates, denoted $m(h_1^{\mathrm{HM}})$ and $m(h_2^{\mathrm{HM}})$ respectively, are used to discriminate between signal and background. In order to obtain these values, the jets originating from the Higgs boson candidates must be identified and then paired together. If there are exactly four $b$-jets in an event, those four are used. If there are more than four $b$-jets, the four with the highest $p_{\mathrm{T}}$ are used. If only three $b$-jets are reconstructed, and one of these jets has a mass greater than 100 GeV, it is considered to be a boosted Higgs boson candidate and no additional jets are considered. Otherwise, the fourth jet is selected as the untagged small-radius jet that minimizes the value of $m(h_1^{\mathrm{HM}})$ that would be obtained by the pairing algorithm discussed in the next paragraph.

The selected jets are paired together to create Higgs boson candidates. If only three jets are selected (in the case where one jet has a mass greater than 100 GeV), the heaviest jet is considered to be a Higgs boson candidate and the remaining two jets are paired to form the second candidate. Otherwise, the quantity $\Delta R^{bb}_{\text{max}}(h_1^{\text{HM}}, h_2^{\text{HM}}) = \max\left(\Delta R(h_1^{\text{HM}}), \Delta R(h_2^{\text{HM}})\right)$ is calculated for each of the three possible pairings of the four jets, where $\Delta R(h)$ is the $\Delta R$ separation of the jets coming from the same Higgs candidate. The pairing that minimizes $\Delta R^{bb}_{\text{max}}$ is used, as pairs of jets in signal events have a more collimated topology compared to the background.

In order to maximize the sensitivity to a broad set of higgsino masses, a BDT is used to discriminate background from signal events and define the regions of the channel. The BDT was trained using the XGBoost algorithm [98] over inclusive background and signal datasets, properly reweighted to account for the different cross sections of the simulated processes. For the classification, the BDT exploited the following inputs:

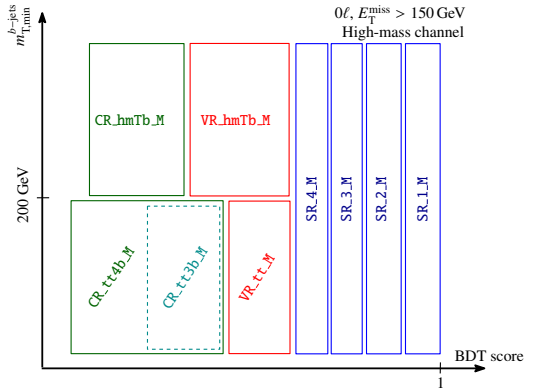
- the number of jets, $N_{\text{jets}}$, and $b$-jets, $N_{\text{b-jets}}$, in the event,
- the scalar sum of the transverse momenta associated to small-radius jets, $H_{\text{T}}$,
- the magnitude, $E_{\text{T}}^{\text{miss}}$, of the missing transverse momentum, and the object-based $E_{\text{T}}^{\text{miss}}$ significance, $\mathcal{S}(E_{\text{T}}^{\text{miss}})$ [99],
- the minimum difference of azimuthal angle between $E_{\text{T}}^{\text{miss}}$ and any of the four highest-$p_{\text{T}}$ jets, $\Delta\phi^{4j}_{\text{min}}$,
- the minimum transverse mass between the $E_{\text{T}}^{\text{miss}}$ and the three leading $b$-jets,

$$m^{b-\text{jets}}_{\text{T,min}} = \min_{i\leq 3}\sqrt{(E_{\text{T}}^{\text{miss}} + p_T^{j_i})^2 - (E_{\text{T}}^{\text{miss}}{}_x + p_x^{j_i})^2 - (E_{\text{T}}^{\text{miss}}{}_y + p_y^{j_i})^2},$$

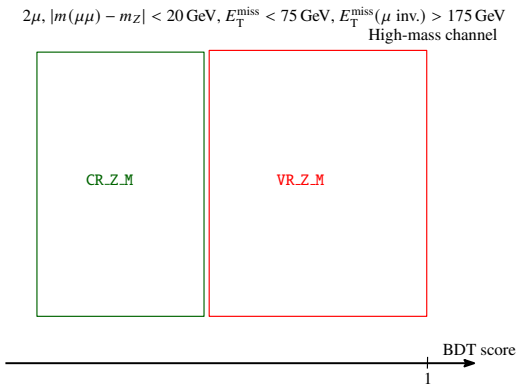
- the minimum angular distance between any 2 $b$-jets, $\Delta R^{bb}_{\text{min}}$,
- the scalar sum of the reclustered large-radius jets masses, $M_J^{\Sigma}$,
- the masses of the reconstructed Higgs boson candidates, $m(h_1^{\text{HM}})$ and $m(h_2^{\text{HM}})$, and the angular distances between the associated jets, $\Delta R(h_1^{\text{HM}})$ and $\Delta R(h_2^{\text{HM}})$.

The BDT was also parametrized with the truth higgsino mass, $m^{\text{sim}}_{\tilde{H}}$, to define signal regions (SRs) each targeting a specific mass hypothesis. The most discriminating variables for low higgsino mass hypotheses are $N_{\text{jets}}$, $N_{\text{b-jets}}$ and $m^{b-\text{jets}}_{\text{T,min}}$, while $E_{\text{T}}^{\text{miss}}$, $\mathcal{S}(E_{\text{T}}^{\text{miss}})$ and $m^{b-\text{jets}}_{\text{T,min}}$ are the most discriminating variables for intermediate to high higgsino mass hypotheses. The optimization of the BDT hyperparameters was performed through a scan to optimize the discovery significance for the benchmark signal $m_{\tilde{H}} = 1000$ GeV. The selected parameters are 500 trees, a learning rate of 0.5, a maximum of 50 bins and a maximum tree depth of six.

The SRs are defined using the BDT output scores of the preselected events. For each mass hypothesis, the SRs are built with an iterative procedure that begins from the highest values of the BDT score. The iterative process determines BDT score thresholds for the SRs by requiring at least 0.5 events of background and maximizing the statistical significance $Z$ calculated from the `BinomialExpZ` function of RooStats [100]. For mass hypotheses that were excluded by the previous ATLAS search using 24.3 $-$36.1 fb$^{-1}$ of Run 2 data [12], the cross sections of the signals are scaled down to the previously excluded values for the purposes of this calculation. After an SR is formed, subsequent SRs are created by repeating the procedure on events with BDT scores below the values used in the previous SR. If this would result in a significance $Z$ of less than one, no SR is created and no further SRs are made for that mass hypothesis. If this results in




(a) Definitions of SRs, CRs and VRs targeting $t\bar{t}$ and high $m_{T,min}^{b-jets}$ events




(b) Definitions of CR and VR for $Z$+jets background

Figure 2: General scheme followed in the definition of SRs, VRs and CRs for each mass hypothesis. The left plot shows the SRs, $t\bar{t}$ CRs and VR and high $m_{T,min}^{b-jets}$ CR and VR. The SRs are constructed with all the events with BDT output scores in a specific range that maximizes the expected significance in the SR. Up to four SRs are defined that way. After finding all SRs, the lower threshold in BDT score of a SR for that specific mass hypothesis is used as the upper cut of the VRs, and the procedure is repeated separately in regions of low and high $m_{T,min}^{b-jets}$ to define VRs targeting the $t\bar{t}$ background and high $m_{T,min}^{b-jets}$ selections, as shown in the figure. Once the VRs are found, the process is repeated to define CRs targeting $t\bar{t}$ and high $m_{T,min}^{b-jets}$ events, with a further splitting of the $t\bar{t}$ CRs in $b$-jet multiplicity. The right plot shows the $Z$+jets CR and VR. These are defined using the same procedure as for the $t\bar{t}$ and high $m_{T,min}^{b-jets}$ regions, but with no SRs and a different preselection designed to capture $Z(\rightarrow \mu\mu)$+jets events. $E_T^{miss}(\mu$ inv.) is used in place of $E_T^{miss}$ for the BDT input variables for the right plot to emulate the behavior of $Z(\rightarrow \nu\nu)$+jets. The lower BDT score thresholds for the CRs are not shown in either plot as they depend on the mass hypothesis and the type of CR.

more than four SRs for a given mass hypothesis, the SRs with the lowest BDT scores are merged until there are only four SRs. The SRs are named as `SR_i_M`, where `M` corresponds to the signal mass hypothesis and `i` is an integer between one and four that labels the SR, with lower values of `i` corresponding to more signal-like SRs. Using different SRs for different signal mass hypotheses improves the sensitivity to low higgsino masses by approximately 20% with respect to using the same SRs for all mass hypotheses. The SRs corresponding to different signal mass hypotheses are not required to be orthogonal. These requirements are shown in Table 3, with a diagram of the SRs in Figure 2(a). The signal regions `SR_1_250`, `SR_1_500` and `SR_1_1000` are additionally used to search for excesses with minimal model dependence and are called discovery regions when used in this context.

### 5.1.2 Background estimation strategy

The background estimation of the high-mass channel relies almost completely on MC simulation with data-driven normalization corrections. After preselection, the main background is $t\bar{t}$, followed by QCD multijet processes, primarily at low $E_T^{miss}$, as well as single top and $Z$+jets, which contribute at high $E_T^{miss}$. Smaller contributions arise from $t\bar{t}$+$X$ and diboson production. Distributions of the data and simulated backgrounds normalized to their theory cross sections after preselection are shown in Figure 3. The data and background predictions agree within 10% after taking the statistical uncertainties into account. The normalizations of the dominant $t\bar{t}$ and $Z$+jets processes are measured through a combined maximum

Table 3: Summary of criteria applied to construct the CRs, VRs and SRs of the high-mass channel. The considered variables are the statistical significance $Z$, the number of background events, $n_{\text{bkg}}$, the ratio of signal and background events, $S/B$, the number of $b$-jets, $N_{\text{b-jets}}$, and $m_{\text{T,min}}^{b-\text{jets}}$. The $Z$+jets preselection is discussed in Section 5.1.2. Additional selections based on the BDT score are included in the definitions of the CRs, VRs and SRs, as shown in Figure 2; these are mass-dependent and are therefore omitted from the table.

| Region name | Fixed Requirements | | | Boundary Conditions | | |
|---|---|---|---|---|---|---|
| | Preselection | $m_{\text{T,min}}^{b-\text{jets}}$ | $N_{\text{b-jets}}$ | $Z$ | $n_{\text{bkg}}$ | $S/B$ |
| SR_i_M | Standard | – | – | max. | $\geq 0.5$ | – |
| VR_tt_M | Standard | $< 200$ GeV | – | – | $\geq 25$ | $< 0.2$ |
| VR_hmTb_M | Standard | $> 200$ GeV | – | – | $\geq 25$ | $< 0.2$ |
| VR_Z_M | $Z$+jets | – | – | – | $\geq 25$ | $< 0.2$ |
| CR_tt3b_M | Standard | $< 200$ GeV | $= 3$ | – | $\geq 100$ | $< 0.1$ |
| CR_tt4b_M | Standard | $< 200$ GeV | $\geq 4$ | – | $\geq 100$ | $< 0.1$ |
| CR_hmTb_M | Standard | $> 200$ GeV | – | – | $\geq 100$ | $< 0.1$ |
| CR_Z_M | $Z$+jets | – | – | – | $\geq 100$ | $< 0.1$ |

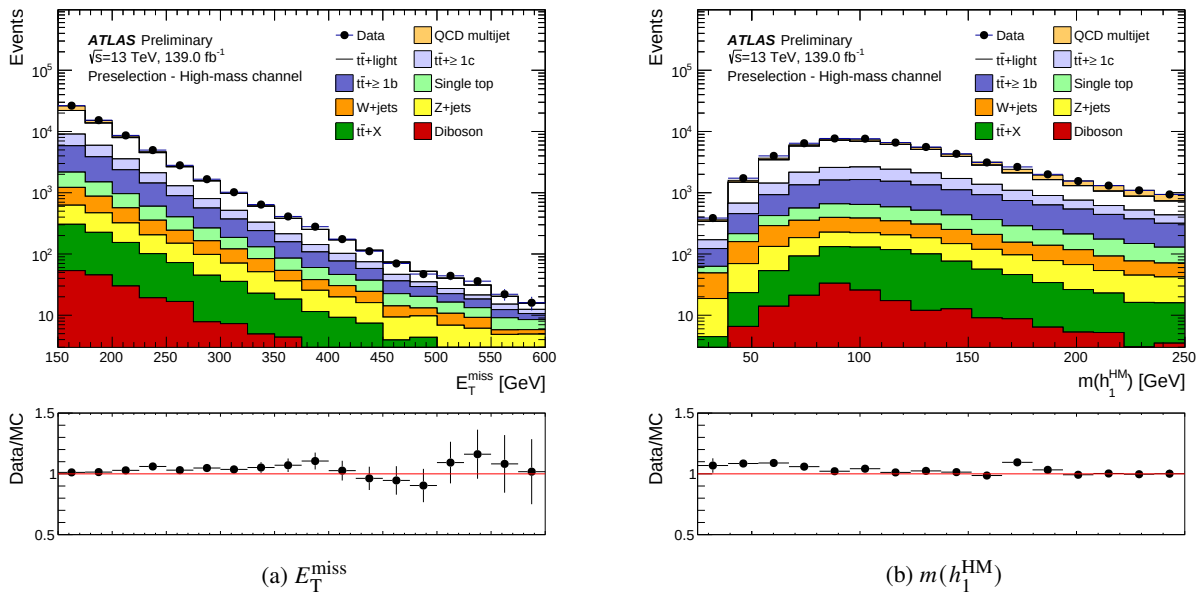
likelihood fit which includes the SRs of the channel, as well as a dedicated set of control regions (CRs) with an enhanced purity of each background component. Three additional parameters are included in the fit to constrain the normalizations of the single top background, which has large but sub-leading contributions to the SRs, and the $t\bar{t}$+$\geq 1b$ and $t\bar{t}$+$\geq 1c$ components of the $t\bar{t}$ background, to improve the modeling in heavy-flavor dominated regions. The agreement between the adjusted background prediction and observed data is checked through dedicated validation regions (VRs). These are named as CR_PROC_M and VR_PROC_M, where PROC labels the physical process and M denotes the signal mass hypothesis. The $t\bar{t}$ and high $m_{\text{T,min}}^{b-\text{jets}}$ CRs and VRs are defined using the same selections as the SRs except for the BDT score requirements. An additional requirement of $m_{\text{T,min}}^{b-\text{jets}} < 200$ GeV ($m_{\text{T,min}}^{b-\text{jets}} > 200$ GeV) is used to enhance the purity of the $t\bar{t}$ (high $m_{\text{T,min}}^{b-\text{jets}}$) CRs and VRs. The $t\bar{t}$ CRs are further split into $3b$ and $4b$ CRs to constrain the normalizations of the $t\bar{t}$+$\geq 1b$ and $t\bar{t}$+$\geq 1c$ backgrounds separately from the overall $t\bar{t}$ background.

The $Z$+jets background is estimated using a separate data sample of muon-enriched events. The dominant component of the $Z$+jets background in the SRs is $Z(\rightarrow \nu\nu)$+jets. The '$Z$+jets preselection' has the same requirements as the standard preselection described above, with the exception of the requirement of two opposite-sign muons satisfying $|m(\mu\mu) - m_Z| < 20$ GeV, where $m_Z$ is the mass of the $Z$ boson, and $E_{\text{T}}^{\text{miss}} < 75$ GeV. The $E_{\text{T}}^{\text{miss}}$ requirement is imposed to suppress the contamination of $t\bar{t}$ background and increase the purity of $Z$+jets events. In addition to the $E_{\text{T}}^{\text{miss}}$ requirement, the cut $E_{\text{T}}^{\text{miss}}(\mu \text{ inv.}) > 175$ GeV is also applied, where $E_{\text{T}}^{\text{miss}}(\mu \text{ inv.})$ is computed considering muons as invisible particles for the emulation of a boosted $Z \rightarrow \nu\nu$ decay. The $E_{\text{T}}^{\text{miss}}(\mu \text{ inv.})$ requirement also allows events to be selected by the $E_{\text{T}}^{\text{miss}}$ trigger, which treats muons as invisible [26]. Once these events are selected, the same BDT used for the signal selection is used to create the CRs and VRs of the $Z$+jets process. All input variables to the BDT using $E_{\text{T}}^{\text{miss}}$ are adjusted to use $E_{\text{T}}^{\text{miss}}(\mu \text{ inv.})$ instead.

For each signal mass hypothesis, the VR for a given background is defined to contain all events with BDT scores below the SR values and greater than a threshold value, selected such that each VR has at least a specific number of background events, denoted as $n_{bkg}$ in Table 3, and a signal contamination less than a threshold value, denoted as $S/B$ in Table 3. The CR is then defined using BDT scores below the values for the VR, with a minimum BDT score selected such that the CR has at least a specific number of background

events and a signal contamination less than a threshold value, as indicated in Table 3. The strategy is illustrated graphically in Figure 2. For the $Z$+jets background, the $S/B$ is estimated to be zero and the signal contamination condition is therefore always satisfied. Due to their different preselection, the $Z$+jets VRs are not required to have BDT scores below the SR values.

A data-driven technique is used to estimate the QCD multijet background. This method exploits a template created by subtracting non-QCD backgrounds estimated with MC simulations from data in a kinematic regime dominated by QCD multijet processes. These events are obtained by replacing the selection $\Delta\phi^{4j}_{\text{min}} > 0.4$ with $\Delta\phi^{4j}_{\text{min}} < 0.2$. In order to get an estimate of the contribution of this background to the CRs, VRs and SRs, the template needs to be evaluated with the discriminating BDT. For this reason, a fake $\Delta\phi^{4j}_{\text{min}}$ distribution is generated for the events in the template, sampled randomly from the expected $\Delta\phi^{4j}_{\text{min}}$ distribution observed in dijet MC simulated samples. This template, however, does not correctly reproduce the correlations between $\Delta\phi^{4j}_{\text{min}}$ and other kinematic variables that are used in the BDT training. Additionally, given the exponentially falling shape of the $\Delta\phi^{4j}_{\text{min}}$ distribution for QCD multijet events, the template is not expected to accurately describe the normalization of this background in the targeted kinematic regime. These issues are resolved by applying a Neural Network (NN) assisted reweighting to the template. The NN is trained in simulated dijet events to separate $\Delta\phi^{4j}_{\text{min}} < 0.2$ and $\Delta\phi^{4j}_{\text{min}} > 0.4$ events, and an event weight based on the per-event NN output score is applied to the data-driven template to correct the normalization and correlations between kinematic variables. The reweighted template is the final estimate for the QCD multijet background. This is validated in kinematic regimes enriched in QCD multijet and is found to describe the data well with up to 10% differences. A non-closure uncertainty of 50% is applied to the prediction of QCD multijet in the CRs, VRs and SRs, found by comparing the performance of the NN-assisted reweighting to the prediction from dijet MC simulation in the analysis regions.




(a) $E_{\text{T}}^{\text{miss}}$ (b) $m(h_1^{\text{HM}})$

Figure 3: Comparisons of data and MC simulations for the standard preselection of the high-mass channel. The left plot shows the distribution of $E_{\text{T}}^{\text{miss}}$ while the right plot shows the distribution of $m(h_1^{\text{HM}})$. The underlays show the ratio of data to MC simulations. The uncertainties on the ratio include the statistical uncertainties from both data and MC simulations.

## 5.2 Low-mass channel

The low-mass channel focuses on final states characterized by four or more $b$-jets and the absence of signal-quality leptons. Two Higgs boson candidates are reconstructed based on the $b$-jets in the event. Signal, control and validation regions are defined in the phase space of leading and sub-leading Higgs masses. In the low-mass channel, the primary sources of background are QCD multijet events and $t\bar{t}$ processes. To estimate these backgrounds, an ABCD method is employed, which involves training a model to adjust the kinematics across control regions with varying $b$-jet multiplicities. The effectiveness of this adjustment is evaluated in validation regions. Within the signal region, two inclusive regions are defined to set model-independent limits on BSM physics. These regions are denoted 'discovery regions' and are optimized to target higgsino mass hypotheses of 150 GeV and 300 GeV respectively.

### 5.2.1 Event selection

The low-mass channel relies on events collected by the $b$-jet triggers listed in Table 1. Events are required to pass at least one of these triggers. Different offline kinematic requirements are also applied to ensure orthogonality between the different trigger selections, allowing correlations between the different trigger scale factors to be ignored.

Events are required to have at least four $b$-jets in order to reconstruct the Higgs candidates. These events will be referred to as the "4$b$" sample. In events with more than four $b$-jets, only the four $b$-jets with the highest $p_{\mathrm{T}}$ are considered. The four jets are then paired to form Higgs candidates by minimizing the quantity $\Delta R^{bb}_{\mathrm{max}}(h_1^{\mathrm{LM}}, h_2^{\mathrm{LM}})$, where $h_1^{\mathrm{LM}}$ is defined as the Higgs boson candidate with the higher reconstructed $p_{\mathrm{T}}$ and $h_2^{\mathrm{LM}}$ is defined as the Higgs boson candidate with the lower $p_{\mathrm{T}}$.

The largest backgrounds passing these selections are QCD multijet and $t\bar{t}$. The leptonic $t\bar{t}$ background is reduced through the signal and loose lepton vetoes included in the common selections. Events are additionally required to have at most two loose leptons[3]. In order to reduce the hadronic $t\bar{t}$ background, a discriminant based on reconstructing the decay of the top quark is used. Top quark candidates are reconstructed using three jets. One of these jets must be from a Higgs boson candidate and is considered to be the $b$-jet from the top quark decay. The other two jets are considered to form a $W$ boson candidate. As a $W$ boson cannot decay into more than one $b$-jet, at least one of the $W$ boson candidate jets must not be associated to a Higgs candidate. The quantity:

$$X_{\mathrm{Wt}} = \sqrt{\left(\frac{m_{jj} - m_W}{0.1 \cdot m_{jj}}\right)^2 + \left(\frac{m_{jjb} - m_t}{0.1 \cdot m_{jjb}}\right)^2} \tag{1}$$

is then calculated for each possible combination of jets subject to the restrictions described above, where $m_{jj}$ and $m_{jjb}$ are the masses of the reconstructed $W$ boson and top quark candidates and $m_W = 80.4\,\mathrm{GeV}$ and $m_t = 172.5\,\mathrm{GeV}$ are the nominal masses of the $W$ boson and top quark. The factor of 0.1 approximates the fractional mass resolution of the reconstructed particle candidates. Events are vetoed if $X_{\mathrm{Wt}} < 1.8$ for any of these combinations.

The low-mass SRs are defined by the requirement $X_{hh}^{\mathrm{SR}} < 1.6$, where $X_{hh}^{\mathrm{SR}}$ is given by:

[3] This veto uses all loose leptons, while the common loose lepton veto only uses loose leptons with $p_{\mathrm{T}} > 8\,\mathrm{GeV}$.

$$X_{hh}^{\mathrm{SR}} = \sqrt{\left(\frac{m(h_1^{\mathrm{LM}}) - 120\,\mathrm{GeV}}{0.1 \cdot m(h_1^{\mathrm{LM}})}\right)^2 + \left(\frac{m(h_2^{\mathrm{LM}}) - 110\,\mathrm{GeV}}{0.1 \cdot m(h_2^{\mathrm{LM}})}\right)^2}, \tag{2}$$

where the denominators are the approximate mass resolutions for the Higgs boson candidates. A separate SR is created for each of the 2016, 2017 and 2018 data-taking periods to account for differences in the triggers and the background estimation procedure. The central values of 120 GeV and 110 GeV are offset from the true Higgs mass due to inefficiencies in reconstruction such as neutrinos produced in $b$-hadron semileptonic decays. The shape of the SR requirement is shown as the innermost contour in Figure 4.

The SRs are binned in $E_{\mathrm{T}}^{\mathrm{miss}}$ and $m_{\mathrm{eff}}$, where $m_{\mathrm{eff}}$ is defined as the $E_{\mathrm{T}}^{\mathrm{miss}}$ plus the scalar sum of the $p_{\mathrm{T}}$ values of the jets associated to Higgs candidates. A fit is performed over the two-dimensional distribution of these variables, with lower bin edges given by:

$$E_{\mathrm{T}}^{\mathrm{miss}} = \{0, 20, 40, 60, 80, 100, 120, 140, 160, 180, 200\}\ \mathrm{GeV}$$
$$m_{\mathrm{eff}} = \{160, 200, 260, 340, 440, 560, 700, 860\}\ \mathrm{GeV}$$

where the last bin is inclusive. In addition, two discovery regions are defined to provide limits on model-independent scenarios. These regions are optimized for the $m_{\tilde{H}} = 150$ GeV (`SR_LM_150`) and $m_{\tilde{H}} = 300$ GeV (`SR_LM_300`) mass points, though they are also sensitive to nearby $m_{\tilde{H}}$ values. The region definitions are shown in Table 4.

Table 4: Discovery region definitions for the low-mass channel.

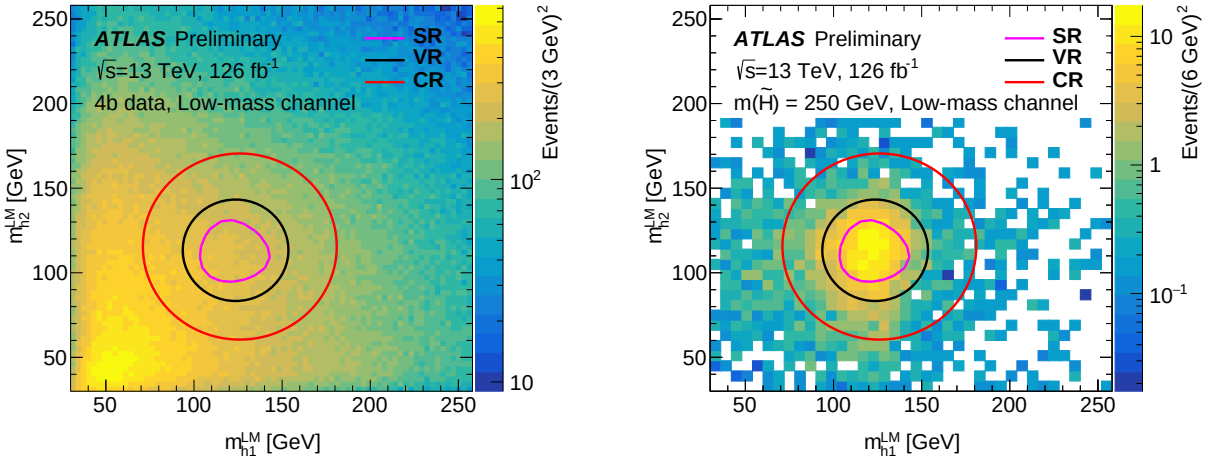
| Region | $E_{\mathrm{T}}^{\mathrm{miss}}$ | $m_{\mathrm{eff}}$ |
|---|---|---|
| `SR_LM_150` | $> 20$ GeV | $> 560$ GeV |
| `SR_LM_300` | $> 150$ GeV | $> 340$ GeV |

### 5.2.2 Background estimation strategy

The dominant background in the low-$E_{\mathrm{T}}^{\mathrm{miss}}$ regime of the low-mass channel comes from QCD multijet processes. Because these processes are difficult to model sufficiently well in simulation, a purely data-driven approach is used to estimate the background from all sources. This estimate makes use of an alternative set of regions with exactly two $b$-jets and two or more non-$b$-tagged jets instead of 4 or more $b$-jets. All other analysis selections are the same as those previously described in Section 5.2.1. These alternative regions, called the "2$b$" sample, have higher statistics and significantly reduced signal contamination with respect to the nominal 4$b$ sample.

The Higgs boson candidates for the 2$b$ sample are reconstructed using the two $b$-jets plus two jets selected randomly from the event's non-$b$-tagged jets. The additional jets are required to have $|\eta| < 2.5$ to match the $\eta$ requirements for $b$-tagging. The masses of the reconstructed Higgs bosons are used to create the VRs and CRs. The VRs are defined for each data-taking period by $R_{hh}^{\mathrm{VR}} < 30$ GeV and $X_{hh}^{\mathrm{SR}} \geq 1.6$, where $R_{hh}^{\mathrm{VR}}$ is given by:

$$R_{hh}^{\mathrm{VR}} = \sqrt{(m(h_1^{\mathrm{LM}}) - 123.6\,\mathrm{GeV})^2 + (m(h_2^{\mathrm{LM}}) - 113.3\,\mathrm{GeV})^2}. \tag{3}$$




(a) $4b$ data

(b) $m_{\tilde{H}} = 250\,\text{GeV}$

Figure 4: Mass of the reconstructed Higgs bosons for the low-mass channel. The left plot shows the $4b$ data while the right plot shows the $4b$ signal MC for a higgsino mass of 250 GeV. The red, black and pink contours correspond to the outer boundary of the control, validation and signal regions respectively.

The CRs are defined by $R_{hh}^{\text{CR}} < 55\,\text{GeV}$ and $R_{hh}^{\text{VR}} \geq 30\,\text{GeV}$, where $R_{hh}^{\text{CR}}$ is given by:

$$R_{hh}^{\text{CR}} = \sqrt{(m(h_1^{\text{LM}}) - 126.0\,\text{GeV})^2 + (m(h_2^{\text{LM}}) - 115.5\,\text{GeV})^2}. \tag{4}$$

The VRs and CRs form ellipses around the SRs with the central values shifted by factors of 1.03 and 1.05 respectively. These shifts avoid over-representing low $(m(h_1^{\text{LM}}),m(h_2^{\text{LM}}))$ values, which have a higher density of events. The SRs, VRs and CRs are each defined for both the $2b$ and the $4b$ samples. The $2b$ CRs, $4b$ CRs and $2b$ SRs are used to create a background model for the $4b$ SRs where the search is performed. Figure 4 shows the $m(h_1^{\text{LM}})$ versus $m(h_2^{\text{LM}})$ distributions for the $4b$ data and the 250 GeV signal simulation with the region definitions overlaid.

The background model is created by normalizing and reweighting the $2b$ SR to estimate the behavior of the background in the $4b$ SR. The normalization factor is determined from the CRs as:

$$\mu_{\text{CR}} = \frac{N_{\text{CR}}^{4b}}{N_{\text{CR}}^{2b}},$$

where $N_{\text{CR}}^{4b}$ ($N_{\text{CR}}^{2b}$) is the number of observed data events in the $4b$ ($2b$) CR. The parameter $\mu_{\text{CR}}$ is measured separately for the 2016, 2017 and 2018 data-taking periods due to the differences in the triggers used to record the data.

The $2b$ regions are reweighted using BDTs to correct any kinematic differences between the $2b$ and $4b$ regions. The BDTs are trained on data events in the CRs separately for the 2016, 2017 and 2018 data-taking periods. The hyperparameters used in common for each year are the maximum number of layers (5), the minimum number of events per node (250), the learning rate (0.3) and the sampling fraction (0.4). The

number of trees was set to 50 for 2016, 75 for 2017 and 100 for 2018. A greater number of trees are used for later years due to the increased size of the data sample. A large set of 51 input variables is provided, as the BDT training can select which variables to weight most strongly. These include the mass, energy, $p_{\mathrm{T}}$, $\eta$ and $\phi$ of each Higgs boson candidate and Higgs boson candidate jet; the mass and $p_{\mathrm{T}}$ of the di-Higgs system; the number of jets; the $E_{\mathrm{T}}^{\mathrm{miss}}$; $X_{Wt}$ without the light jet requirement; the number of track-jets associated with each Higgs candidate and 14 angular variables.

For each tree of the BDT, events are split into two nodes using the input variable and the cut value maximizing the $\chi^2$ difference between the nodes. This process repeats on each resulting node until either the maximum number of layers or minimum number of events per node is reached.

Each of the final nodes (or leaves) is given a weight

$$w = e^{\lambda N^{4b}_{\mathrm{CR,leaf}}/N^{2b}_{\mathrm{CR,leaf}}},$$

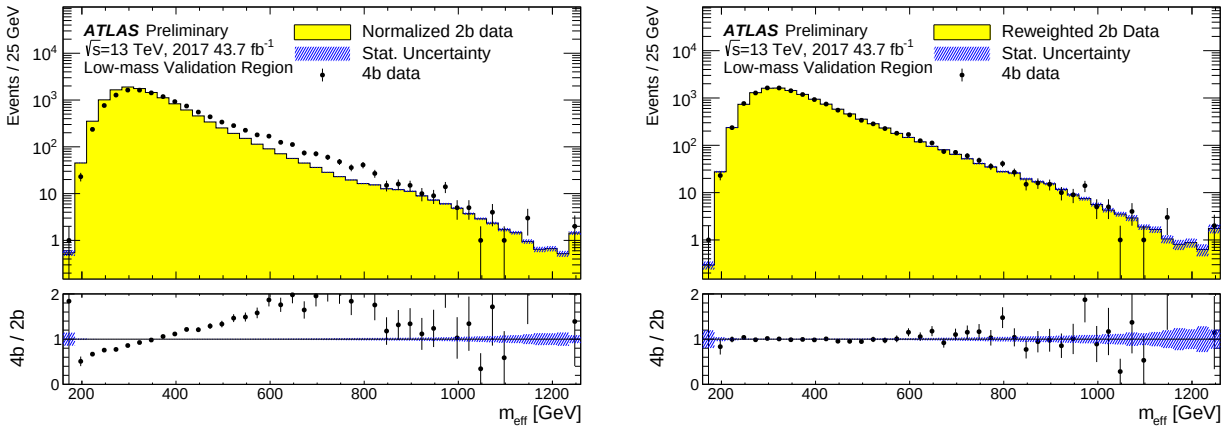
where $\lambda$ is the learning rate and $N^{4b}_{\mathrm{CR,leaf}}$ ($N^{2b}_{\mathrm{CR,leaf}}$) is the number of events in the leaf for the $4b$ ($2b$) CR. The results are then input into a new decision tree. This process iterates until the desired number of trees is reached. The final weight for each event is the product of the event's weight in each tree, normalized such that the number of events in the reweighted $2b$ CR matches that of the $4b$ CR. After training on the CR, the BDT is used to calculate weights for all $2b$ events. The reweighted $2b$ SR then forms the background estimate for the $4b$ SR.

A bootstrapping procedure is used to increase the stability and estimate the statistical uncertainty of the BDT reweighting. In addition to the nominal BDT, an ensemble of 100 BDTs were trained for each year. These BDTs were trained using the same input events but with random Poisson weights with $\mu = 1$ to model the effect of statistical variations on the training data. The background estimate in any region of phase space is then set to the value obtained from the median of the 100 + 1 variations. Half of the difference between the 84th percentile and 16th percentile of the 100 + 1 bootstrap variations is set as the statistical uncertainty due to the reweighting. This is combined in quadrature with the Poisson uncertainties from the $2b$ samples.

Figure 5 shows the $m_{\mathrm{eff}}$ distribution in the VR for the 2017 dataset before and after the reweighting is applied. Excellent agreement is observed between the reweighted $m_{\mathrm{eff}}$ for the $2b$ and $4b$ data. The agreement was also checked for other years and input variables and found to be good in all cases.

# 6 Systematic uncertainties

Systematic uncertainties from various sources are evaluated for this analysis. Uncertainties are divided into three types: experimental uncertainties, theoretical uncertainties and uncertainties on the data-driven background estimate. Experimental uncertainties quantify systematic effects due to the ATLAS detector and LHC conditions. Theoretical uncertainties account for possible mismodeling in the simulations of various physical processes. Experimental and theoretical uncertainties apply to the signal models used by both channels as well as the MC background estimates for the high-mass channel. Uncertainties on the data-driven background estimate are applied to the low-mass channel to account for potential mismodeling from the normalization factors and the BDTs used to estimate the background.




(a) Before reweighting

(b) After reweighting

Figure 5: Comparison of the $2b$ and $4b$ $m_{\text{eff}}$ distributions in the 2017 VR of the low-mass channel. The left plot shows the $m_{\text{eff}}$ distribution before BDT reweighting. The right plot illustrates the same distribution, with the $2b$ events having the reweighting applied to match the $4b$ data. The final bin contains overflow events. The $2b$ distribution is normalized to the $4b$ integral to show the shape differences. The blue shading shows the uncertainty due to the $2b$ sample statistics for the left plot and both the reweighting and the $2b$ sample statistics for the right plot.

## 6.1 Experimental uncertainties

Systematic uncertainties corresponding to the jet energy scale and resolution [85], jet mass scale [101], flavor-tagging efficiencies [89], trigger efficiencies, soft $E_{\text{T}}^{\text{miss}}$ terms [102], pileup conditions [103, 104] and luminosity are included in this analysis. The uncertainty on the combined 2015–2018 integrated luminosity is 1.7% [105], obtained using the LUCID-2 detector [106] for the primary luminosity measurements. These uncertainties are assessed separately for each signal and background process. For the high-mass channel, experimental uncertainties are correlated across all of the SRs, VRs and CRs corresponding to the higgsino mass hypothesis under test, while the low-mass channel only applies these systematics to the $4b$ SRs as no other regions are used in the fit. For the high-mass channel, the jet energy scale, jet energy resolution and jet mass scale uncertainties are derived using a single merged CR, VR and SR for each mass point. These regions are created for each mass point by merging each of the mass point's CRs, VRs and SRs respectively and are used to avoid instabilities in the fit due to large statistical fluctuations. For the low-mass channel, the trigger and offline flavor-tagging uncertainties are calculated jointly to properly account for correlations. The jet energy resolution uncertainties for the low-mass channel are smoothed with a Gaussian kernel to mitigate the effect of statistical fluctuations.

## 6.2 Theoretical uncertainties

The signal models used by both channels are also subject to theoretical uncertainties. Truth samples are run on a representative set of mass points for variations in the factorization and renormalization scales, merging scale, parton shower tuning and radiation uncertainties. The differences between variations are set as the systematic uncertainties. For the low-mass channel, these differences are smoothed using a Gaussian kernel.

Theoretical uncertainties are also applied to the background samples for the high-mass channel. Uncertainties quantifying the systematic effects due to possible mismodeling associated the with MC generator and with parton showering models, as well as with the modeling of initial- and final-state radiation, are applied to the $t\bar{t}$ and single top backgrounds. For the $tW$ process, the uncertainty associated to interference with $t\bar{t}$ production was estimated by comparison with an alternative sample generated using the diagram subtraction scheme [35, 107]. Uncertainties related to the renormalization, factorization, resummation and matching scales are applied to the $W$+jets and $Z$+jets backgrounds. 50% uncertainties are applied to the $t\bar{t}+X$ and diboson backgrounds. As the $t\bar{t}+X$ and diboson backgrounds make up a small fraction of the overall background, these uncertainties have no significant impact on the analysis.

Figure 6 illustrates the total post-fit uncertainties for the high-mass channel, as well as the dominant component of its representative SRs. Statistical uncertainties dominate. At masses below 1000 GeV, $t\bar{t}$ theoretical uncertainties associated with the process generation and parton showering represent one of the most important systematic uncertainty components due to the dominance of this background process. In contrast, signal regions with $m_{\tilde{H}} \geq 1000$ GeV are more affected by uncertainies on $Z$+jets processes due to the larger importance of $Z \rightarrow \nu\nu$ background at higher $E_{\mathrm{T}}^{\mathrm{miss}}$.

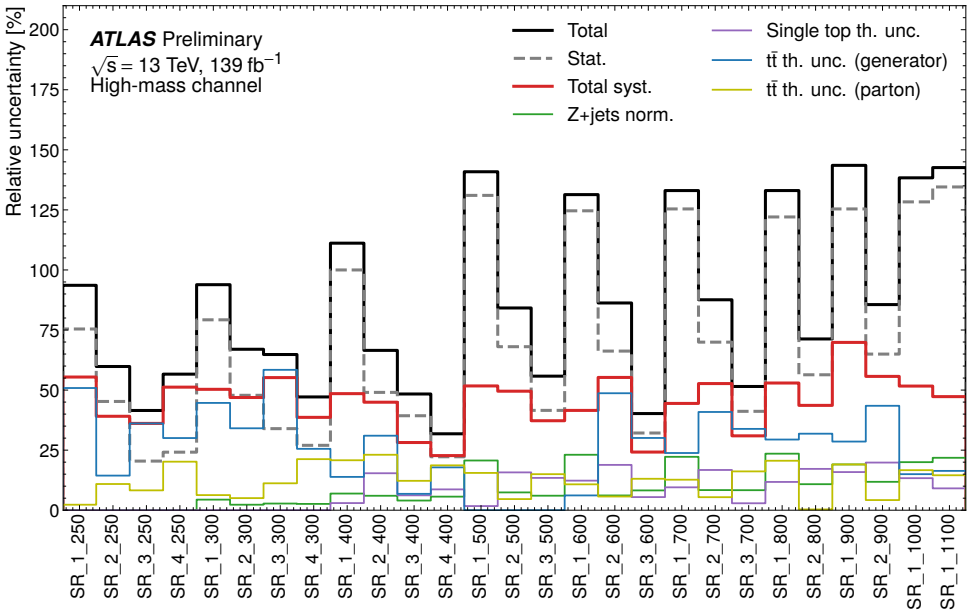



Figure 6: Breakdown of the dominant post-fit systematic uncertainties in the SRs of the high-mass channel. The total systematic uncertainty may be greater than the sum of its components due to correlations between systematics. The total uncertainty shown is the quadrature sum of the statistical and total systematic uncertainties.

### 6.3 Data-driven background uncertainties

The low-mass channel has three sources of systematic uncertainties on the data-driven background estimate. These are each derived separately for the 2016, 2017 and 2018 data-taking periods.

A non-closure uncertainty is assigned to account for imperfections in the reweighting procedure. This is estimated by evaluating the bin-by-bin differences between the $4b$ and reweighted $2b$ samples in the CRs. A nuisance parameter corresponding to the discrepancy is assigned to each bin where the difference is

greater than the combined statistical uncertainty of the $2b$ and $4b$ samples. No nuisance parameters are assigned to bins where the discrepancy is less than this uncertainty. The non-closure uncertainties are treated as uncorrelated in the statistical analysis.

A transfer uncertainty is assigned to take into account the validity of extrapolating results from the CRs to the SRs. This transfer uncertainty is divided into a normalization and a shape component. To estimate the transfer shape uncertainty, the BDTs were retrained using data from the VRs instead of the CRs. All other aspects of the training, including the bootstrapping procedure, were unchanged. This creates alternative background estimates with weights derived closer to the SRs. Each VR-derived estimate is then normalized to the corresponding nominal (CR-derived) estimate in the SR to separate this uncertainty from the normalization uncertainty. The difference between the two estimates is calculated for each bin and set as an uncertainty if the difference is greater than the combined statistical uncertainty of the estimates. As with the non-closure uncertainty, the transfer shape uncertainty is treated as uncorrelated across bins and is not set for bins where the difference is less than the combined statistical uncertainty of the estimates.

The third source is the transfer normalization uncertainty, which is computed by two methods. In the first method, the uncertainty is set as the fractional difference between the predicted number of SR events using the nominal and VR-derived background estimates. In the second method, the uncertainty for each data-taking period is set as the fractional difference between the predicted and observed number of events in the VR. The total normalization uncertainty is set to the maximum of the two estimates.

A slightly different treatment is used to assign transfer uncertainties for the discovery regions. As the discovery regions consist of single bins, the transfer normalization and shape uncertainties are calculated jointly as the difference between the number of predicted SR events using the nominal and VR-derived background estimates. In addition, the VR non-closure is calculated as the difference between the number of predicted and observed events in the VR. The transfer uncertainties for the discovery regions are then set to the maximum of the transfer and VR non-closure uncertainties. The boostrap and non-closure uncertainties are assigned in the same way as in the exclusion regions, except that the non-closure uncertainty is set even if it is less than the statistical uncertainty.

The different components of the background uncertainty for the low-mass channel are shown in Figure 7. Considering all the sources of systematic uncertainties, the non-closure, jet energy resolution and jet flavor composition (a component of the jet energy scale) uncertainties tend to have the greatest effect.

A 50% uncertainty is applied to the QCD multijet prediction for the high-mass channel to account for potential mismodeling from the NN-assisted reweighting.

# 7 Results

## 7.1 High-mass channel

The post-fit agreement of data and background prediction is shown for each VR of the high-mass channel in Figure 8. All data observations are found to be in agreement with the background prediction within the analysis uncertainties. The $t\bar{t}$ and $Z$+jets normalization factors are measured to be between 0.9 and 1.3 for all mass hypotheses.

The SRs observations are shown in Figure 9, together with the post-fit background prediction. No statistically significant excess is found, and the largest excess is observed in `SR_1_1000` with a local

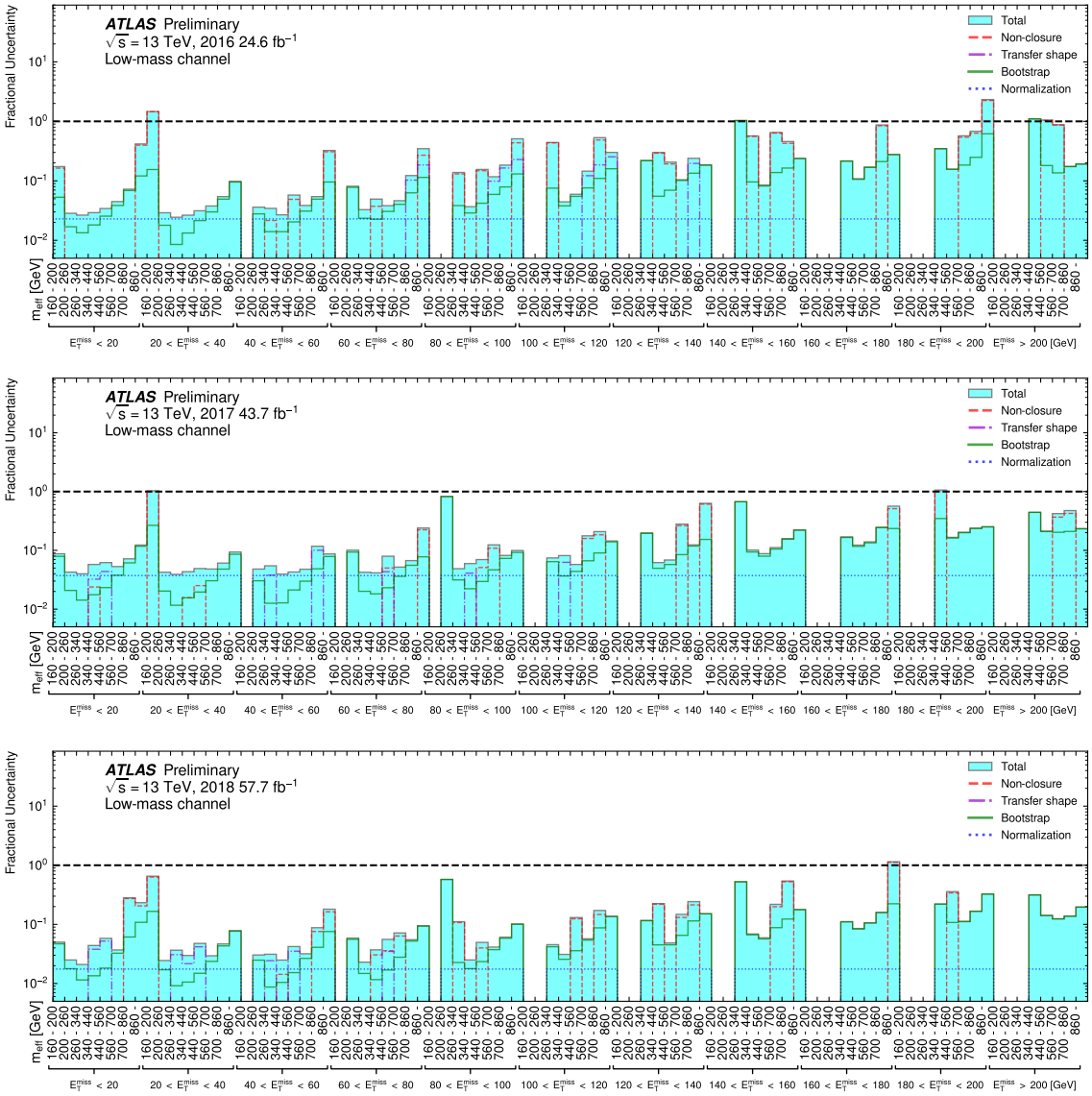



Figure 7: Background uncertainties as a fraction of background yields for the 2016, 2017 and 2018 SRs. Pre-fit values are shown. The bootstrap component also includes the Poisson component of the statistical uncertainty of the $2b$ sample. No uncertainties are shown for bins with 0 predicted background events. The total uncertainty shown is the quadrature sum of the individual uncertainties. All uncertainties are treated as symmetric.

statistical significance of $1.9\sigma$. SRs targeting $m_{\tilde{H}}$ with values included between 900 GeV and 1100 GeV show similar excesses due to their highly correlated SR definitions. All observed data in these regions are found to be the same.

## 7.2 Low-mass channel

The observations and normalized background predictions are shown for each $E_{\mathrm{T}}^{\mathrm{miss}}$ and $m_{\mathrm{eff}}$ bin of the VRs in Figure 10. The lower panel shows the significance of the deviations using the profile likelihood method

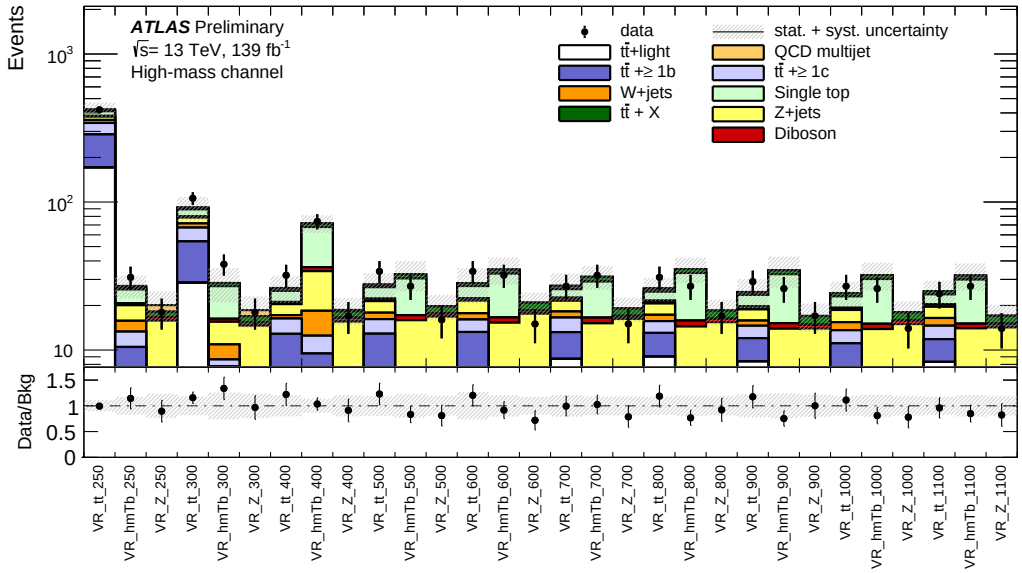



Figure 8: Post-fit VRs yields of the high-mass channel. The upper panel shows the observed number of events, as well the post-fit background predictions in each region. The bottom panel shows the ratio of the observed data and the total background prediction. The shaded areas correspond to the total statistical and systematic uncertainties obtained after the fit and described in Section 6.

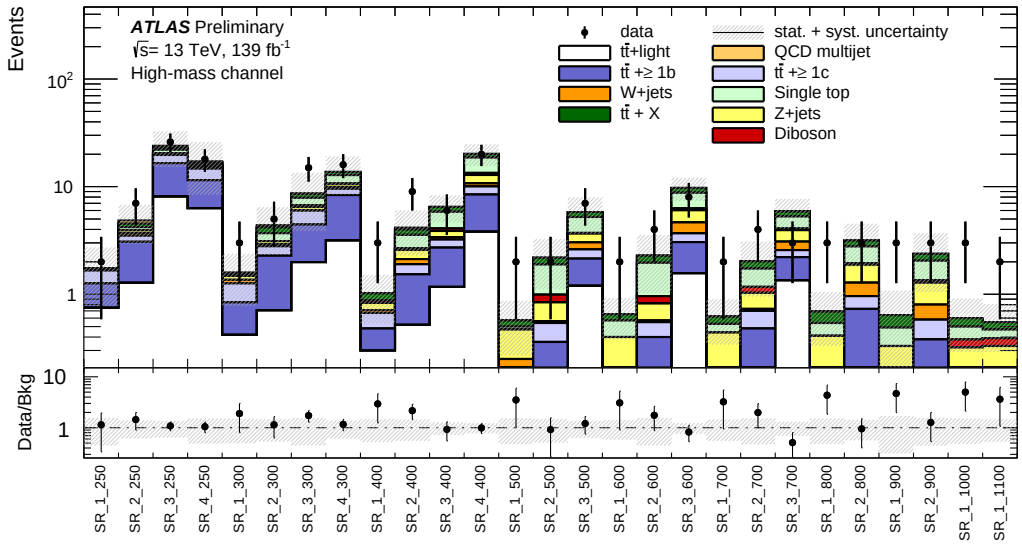



Figure 9: Post-fit SRs yields of the high-mass channel. The upper panel shows the observed number of events, as well the post-fit background predictions in each region. The bottom panel shows the ratio of the observed data and the total background prediction. The shaded areas correspond to the total statistical and systematic uncertainties obtained after the fit and described in Section 6.

Table 5: Results of the high-mass and low-mass channel discovery regions. The columns include, from left to right: the number of observed events $N_{\text{obs}}$, number of predicted events $N_{\text{pred}}$ and 95% CL upper limits on the visible cross section ($\langle\epsilon\sigma\rangle^{95}_{\text{obs}}$) and on the number of signal events ($S^{95}_{\text{obs}}$ ). The $S^{95}_{\text{exp}}$ column shows the 95% CL upper limit on the number of signal events, given the expected number (and $\pm 1\sigma$ excursions on the expectation) of background events. The last column indicates the discovery $p$-value ($p(s = 0)$) and its corresponding significance. The $p$-values are capped at 0.5.

| Signal channel | $N_{\text{obs}}$ | $N_{\text{pred}}$ | $\langle\epsilon\sigma\rangle^{95}_{\text{obs}}$ [fb] | $S^{95}_{\text{obs}}$ | $S^{95}_{\text{exp}}$ | $p(s = 0)$ |
|---|---|---|---|---|---|---|
| SR_1_250 | 2 | $1.8 \pm 1.0$ | 0.04 | 6.2 | $5.9^{+1.7}_{-0.9}$ | 0.48 (0.05) |
| SR_1_500 | 2 | $0.58 \pm 0.30$ | 0.04 | 5.5 | $4.0^{+1.7}_{-0.6}$ | 0.18 (0.92) |
| SR_1_1000 | 3 | $0.60 \pm 0.31$ | 0.05 | 6.7 | $4.3^{+0.9}_{-0.9}$ | 0.03 (1.9) |
| SR_LM_150 | 1790 | $1860 \pm 50$ | 0.73 | 92 | $127^{+48}_{-34}$ | 0.5 (0.00) |
| SR_LM_300 | 97 | $77.0 \pm 5.3$ | 0.31 | 39 | $22^{+9}_{-6}$ | 0.03 (1.8) |

from Ref. [108]. No significant mismodeling is observed.

The results for the SRs are shown for each $E_{\text{T}}^{\text{miss}}$ and $m_{\text{eff}}$ bin in Figure 11. No significant excess above the Standard Model is observed. The largest upward fluctuation for the 213 kinematically allowed bins is observed in the $E_{\text{T}}^{\text{miss}} > 200\,\text{GeV}$, $m_{\text{eff}} > 860\,\text{GeV}$ bin for the 2017 data-taking period. This bin has 6 observed events and $1.51 \pm 0.35$ predicted background events, corresponding to a local significance of $2.6\sigma$. This excess occurs for events with high $E_{\text{T}}^{\text{miss}}$, where the high-mass channel has greater sensitivity.

# 8 Statistical interpretation

Following the Neyman–Pearson lemma, upper limits are set using test statistics [109] based on the profile likelihood ratio. The $p$-values of the statistical tests are obtained following the $\text{CL}_{\text{s}}$ prescription [110] and, unless stated otherwise, using the asymptotic approximation described in Ref. [109]. Results obtained using the asymptotic approximation were confirmed using pseudo-experiments. Two different types of upper limits are provided for each channel: model-independent limits and model-dependent limits for the various simulated higgsino masses $m_{\tilde{H}}$.

## 8.1 Model-independent limits

Model-independent limits on the number of beyond-the-SM physics events are set for each discovery SR of the analysis. The upper limit values are obtained by assuming no signal contamination in the analysis CRs, and by adding to each SR a number of signal events corresponding to the BSM signal strength $\mu$ of the fit. The different background and model-independent signals are then adjusted to assess the 95% CL (confidence level) upper limit values. Results are calculated using 5000 pseudo-experiments for the high-mass discovery regions SR_1_250, SR_1_500 and SR_1_1000 and 20000 pseudo-experiments for the low-mass discovery regions SR_LM_150 and SR_LM_300. The upper limit values are summarized in Table 5 for the discovery regions of the high-mass and low-mass channels.

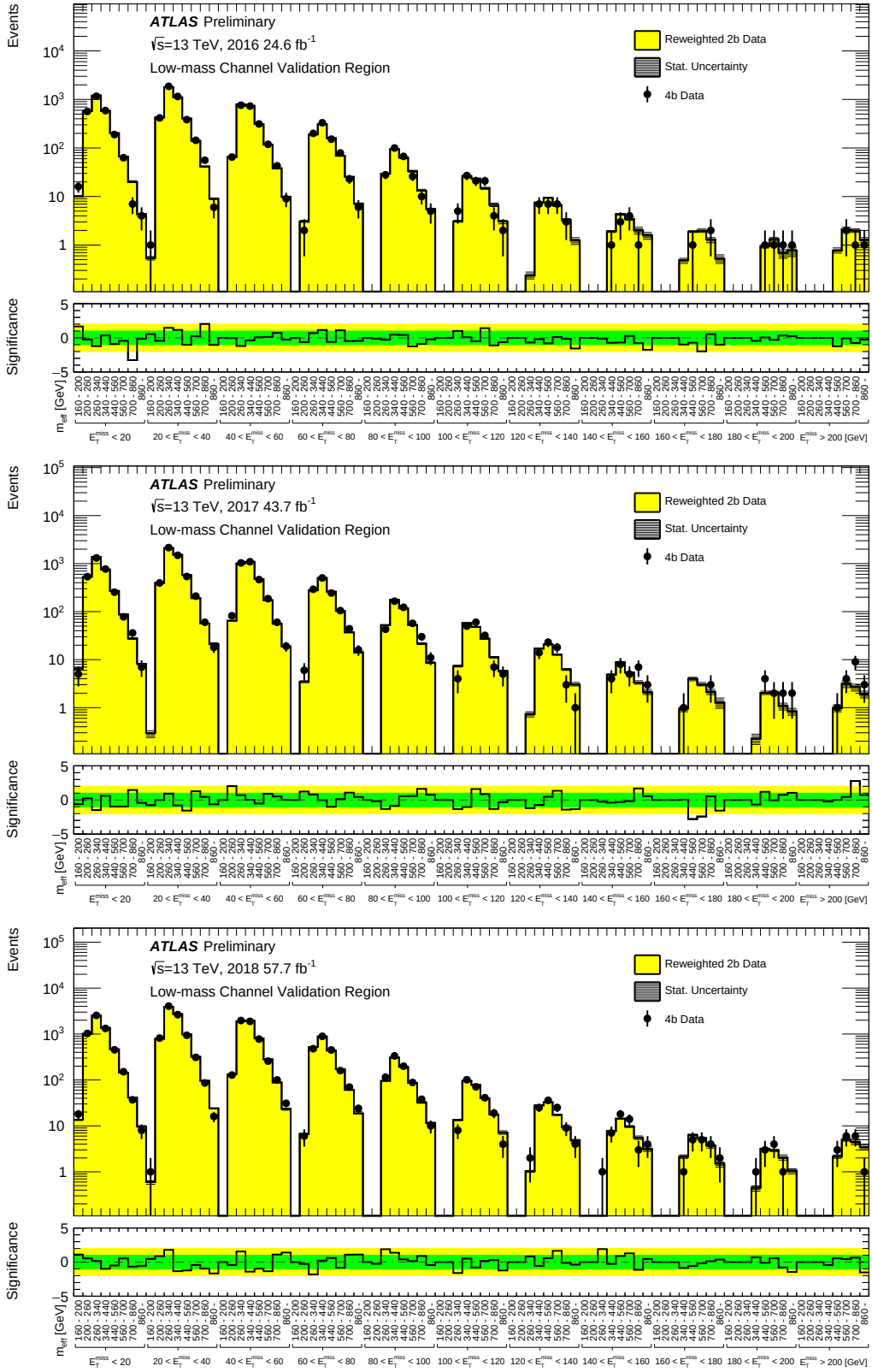



Figure 10: Data and background (reweighted $2b$) predictions for each $4b$ VR $E_{\mathrm{T}}^{\mathrm{miss}}$ and $m_{\mathrm{eff}}$ bin of the low-mass channel for the 2016, 2017 and 2018 data taking periods. The background is normalized to the $4b$ data. The bottom panel shows the significance of any differences between the observed $4b$ data and the background prediction. The $1\sigma$ and $2\sigma$ bands are shown in green and yellow respectively.

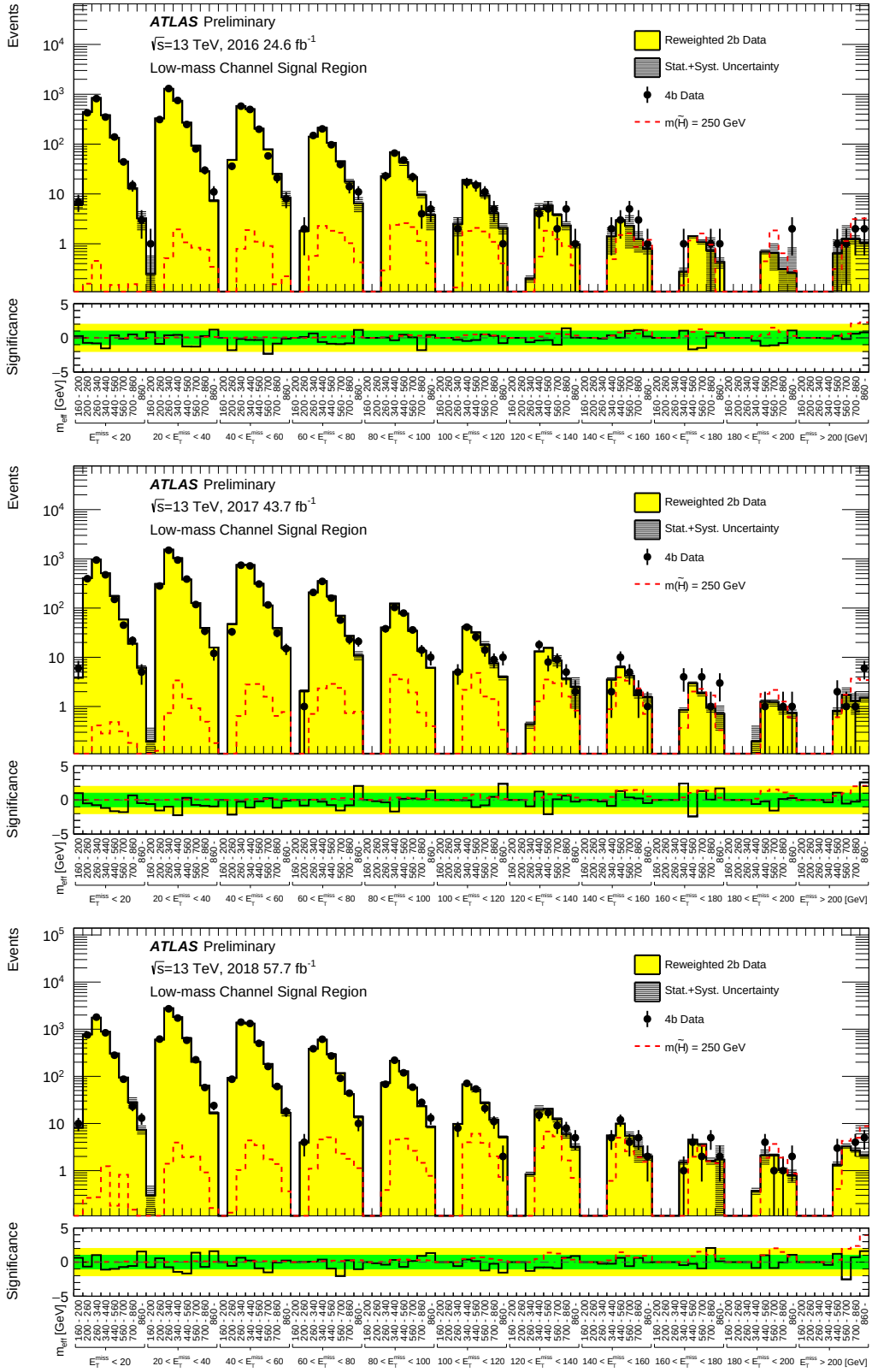
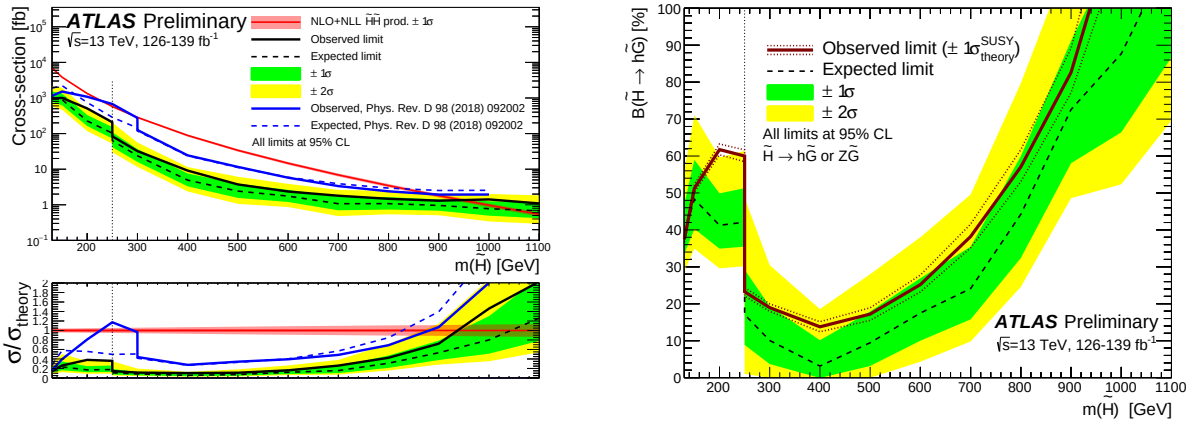



Figure 11: Pre-fit data and background (reweighted $2b$) predictions for each $4b$ SR $E_{\mathrm{T}}^{\mathrm{miss}}$ and $m_{\mathrm{eff}}$ bin of the low-mass channel for the 2016, 2017 and 2018 data taking periods. The bottom panel shows the significance of any differences between the observed $4b$ data and the background prediction. The $1\sigma$ and $2\sigma$ bands are shown in green and yellow respectively.




(a) (b)

Figure 12: Exclusion limits of the low-mass and high-mass channels. The low-mass channel is used for $m_{\tilde{H}} < 250\,\text{GeV}$ while the high-mass channel is used for $m_{\tilde{H}} \geq 250\,\text{GeV}$. The first interpretation (a) shows the observed (solid) and expected (dashed) 95% CL upper limits on the cross section of higgsino pair production, assuming a higgsino decay branching ratio of $\text{BR}(\tilde{H} \to h + \tilde{G}) = 100\%$. The theory cross section and its uncertainty are shown in red and results from a previous ATLAS search using 24.3-36.1 $\text{fb}^{-1}$ [12] are shown in blue. The bottom panel shows the ratio of the limits to the theory cross section. The second interpretation (b) shows the 95% CL observed (solid) and expected (dashed) upper limits on $\text{BR}(\tilde{H} \to h + \tilde{G})$, assuming the theory cross section for higgsino pair production. The higgsinos are assumed to decay as $\tilde{H} \to h + \tilde{G}$ or $\tilde{H} \to Z + \tilde{G}$. For both plots, the phase space above the lines are excluded.

### 8.2 Model-dependent exclusion limits

Since no significant excess is observed above the SM prediction, 95% CL upper limits are calculated for the simplified signal models described in Section 1. Each mass point uses the channel with the better expected limit, with the transition in interpretations between the high-mass and low-mass channels occuring at $m_{\tilde{H}} = 250\,\text{GeV}$. Figure 12a shows the limit on the higgsino pair production cross section for a branching ratio of $\text{BR}(\tilde{H} \to h + \tilde{G}) = 100\%$. Higgsino masses below 940 GeV are excluded.

The results are also interpreted for the case where $\text{BR}(\tilde{H} \to h + \tilde{G})$ is allowed to vary, with all other decays assumed to create $Z$ bosons through $\tilde{H} \to Z + \tilde{G}$. The cross section for higgsino pair production is assumed to match the theoretical prediction. Limits for this interpretation are shown in Figure 12b. For the low-mass channel, the $Z$ decays peak outside of the SRs in the $(m(h_1^{\text{LM}}), m(h_2^{\text{LM}}))$ plane and would be included in the data-driven background estimate. The limit on the branching ratio is therefore equal to the square root of the limit on $\sigma/\sigma_{\text{theory}}$ shown in Figure 12a. For the high-mass channel, both the Higgs boson and the $Z$ boson decays are included in the signal model, leading to additional sensitivity in the branching ratio plane. $\tilde{H} \to h + \tilde{G}$ branching ratios as low as 14% are excluded.

## 9 Conclusion

A search for pair-produced and mass-degenerate higgsinos decaying to gravitinos and Higgs bosons was performed using the $\geq 3$ $b$-jets+$E_{\text{T}}^{\text{miss}}$ final state. The search includes separate channels optimized for

the low-mass and high-mass higgsino regimes and exploits the 139 fb$^{-1}$ of $\sqrt{s}$ = 13 TeV proton–proton collision data collected by the ATLAS detector during the Run 2 of the LHC. This analysis improves upon previous ATLAS results [12] in several ways, including using a Boosted Decision Tree to provide enhanced background rejection for the high-mass channel, reoptimizing the $b$-tagging efficiency working point and binning structure for the low-mass channel, implementing a new method for reconstructing Higgs boson candidates and using improved jet reconstruction and $b$-tagging algorithms.

No significant excess above the SM prediction is observed. Higgsinos with masses between 130 GeV and 940 GeV are excluded at 95% confidence level for the BR($\tilde{H} \rightarrow h + \tilde{G}$) = 100% hypothesis. Upper limits are set on BR($\tilde{H} \rightarrow h + \tilde{G}$) as low as 14% at 400 GeV. Model-independent limits are also set on the visible cross section for new physics processes.

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

[9] S. Dimopoulos, S. Raby, and F. Wilczek, *Supersymmetry and the scale of unification*,
Phys. Rev. D **24** (1981) 1681 (cit. on p. 2).

[10] L. E. Ibáñez and G. G. Ross, *Low-energy predictions in supersymmetric grand unified theories*,
Phys. Lett. B **105** (1981) 439 (cit. on p. 2).

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

[12] ATLAS Collaboration, *Search for pair production of higgsinos in final states with at least three b-tagged jets in $\sqrt{s} = 13$ TeV pp collisions using the ATLAS detector*,
Phys. Rev. D **98** (2018) 092002, arXiv: 1806.04030 [hep-ex] (cit. on pp. 2, 9, 25, 26).

[13] CMS Collaboration, *Search for higgsinos decaying to two Higgs bosons and missing transverse momentum in proton–proton collisions at $\sqrt{s} = 13$ TeV*, JHEP **05** (2022) 014,
arXiv: 2201.04206 [hep-ex] (cit. on p. 2).

[14] ATLAS Collaboration, *Search for supersymmetry in events with four or more charged leptons in 139 $fb^{-1}$ of $\sqrt{s} = 13$ TeV pp collisions with the ATLAS detector*, JHEP **07** (2021) 167,
arXiv: 2103.11684 [hep-ex] (cit. on pp. 2, 7).

[15] ATLAS Collaboration,
*Search for charginos and neutralinos in final states with two boosted hadronically decaying bosons and missing transverse momentum in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*,
Phys. Rev. D **104** (2021) 112010, arXiv: 2108.07586 [hep-ex] (cit. on p. 2).

[16] ATLAS Collaboration, *Searches for new phenomena in events with two leptons, jets, and missing transverse momentum in 139 $fb^{-1}$ of $\sqrt{s} = 13$ TeV pp collisions with the ATLAS detector*,
Eur. Phys. J. C **83** (2022) 515, arXiv: 2204.13072 [hep-ex] (cit. on pp. 2, 7).

[17] ATLAS Collaboration, *Search for pair-produced Higgsinos decaying via Higgs or Z bosons to final states containing a pair of photons and a pair of b-jets with the ATLAS detector*,
ATLAS-CONF-2023-009, URL: https://cdsweb.cern.ch/record/2854839 (cit. on p. 2).

[18] P. Meade, N. Seiberg, and D. Shih, *General Gauge Mediation*, Prog. Theor. Phys. Suppl. **177** (2009) 143, arXiv: 0801.3278 [hep-ph] (cit. on p. 2).

[19] C. Cheung, A. L. Fitzpatrick, and D. Shih, *(Extra)ordinary gauge mediation*, JHEP **07** (2008) 054, arXiv: 0710.3585 [hep-ph] (cit. on p. 2).

[20] M. Dine and W. Fischler, *A phenomenological model of particle physics based on supersymmetry*, Phys. Lett. B **110** (1982) 227 (cit. on p. 2).

[21] L. Alvarez-Gaumé, M. Claudson, and M. B. Wise, *Low-energy supersymmetry*, Nucl. Phys. B **207** (1982) 96 (cit. on p. 2).

[22] C. R. Nappi and B. A. Ovrut, *Supersymmetric extension of the SU(3) x SU(2) x U(1) model*, Phys. Lett. B **113** (1982) 175 (cit. on p. 2).

[23] S. Dimopoulos, M. Dine, S. Raby, and S. D. Thomas, *Experimental Signatures of Low Energy Gauge-Mediated Supersymmetry Breaking*, Phys. Rev. Lett. **76** (1996) 3494, arXiv: hep-ph/9601367 (cit. on p. 2).

[24] K. T. Matchev and S. D. Thomas, *Higgs and Z-boson signatures of supersymmetry*, Phys. Rev. D **62** (2000) 077702, arXiv: hep-ph/9908482 (cit. on p. 2).

[25] P. Meade, M. Reece, and D. Shih, *Prompt Decays of General Neutralino NLSPs at the Tevatron*, JHEP **05** (2010) 105, arXiv: 0911.4130 [hep-ph] (cit. on p. 2).

[26] 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 pp. 2, 11).

[27] ATLAS Collaboration, *Configuration and performance of the ATLAS b-jet triggers in Run 2*, Eur. Phys. J. C **81** (2021) 1087, arXiv: 2106.03584 [hep-ex] (cit. on p. 2).

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

[29] 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).

[30] 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 pp. 4, 7).

[31] 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. 4).

[32] R. D. Ball et al., *Parton distributions with LHC data*, Nucl. Phys. B **867** (2013) 244, arXiv: 1207.1303 [hep-ph] (cit. on pp. 4, 6).

[33] 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. 4).

[34] D. J. Lange, *The EvtGen particle decay simulation package*, Nucl. Instrum. Meth. A **462** (2001) 152 (cit. on p. 4).

[35] 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 pp. 4, 18).

[36] T. Gleisberg and S. Höche, *Comix, a new matrix element generator*, JHEP **12** (2008) 039, arXiv: 0808.3674 [hep-ph] (cit. on pp. 4, 6).

[37] 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. 4, 6).

[38] 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. 4, 6).

[39] 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. 4, 6).

[40] 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. 4, 6).

[41] 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. 4, 6).

[42] F. Buccioni et al., *OpenLoops 2*, Eur. Phys. J. C **79** (2019) 866, arXiv: 1907.13071 [hep-ph] (cit. on pp. 4, 6).

[43] 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. 4, 6).

[44] 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. 4, 6).

[45] W. Beenakker et al., *Production of Charginos, Neutralinos, and Sleptons at Hadron Colliders*, Phys. Rev. Lett. **83** (1999) 3780, arXiv: hep-ph/9906298 (cit. on p. 4), Erratum: Phys. Rev. Lett. **100** (2008) 029901.

[46] G. Bozzi, B. Fuks, and M. Klasen, *Threshold Resummation for Slepton-Pair Production at Hadron Colliders*, Nucl. Phys. B **777** (2007) 157, arXiv: hep-ph/0701202 [hep-ph] (cit. on p. 4).

[47] B. Fuks, M. Klasen, D. R. Lamprea, and M. Rothering, *Gaugino production in proton-proton collisions at a center-of-mass energy of 8 TeV*, JHEP **10** (2012) 081, arXiv: 1207.2159 [hep-ph] (cit. on p. 4).

[48] B. Fuks, M. Klasen, D. R. Lamprea, and M. Rothering, *Precision predictions for electroweak superpartner production at hadron colliders with RESUMMINO*, Eur. Phys. J. C **73** (2013) 2480, arXiv: 1304.0790 [hep-ph] (cit. on p. 4).

[49] B. Fuks, M. Klasen, D. R. Lamprea, and M. Rothering, *Revisiting slepton pair production at the Large Hadron Collider*, JHEP **01** (2014) 168, arXiv: 1310.2621 [hep-ph] (cit. on p. 4).

[50] J. Fiaschi and M. Klasen, *Slepton pair production at the LHC in NLO+NLL with resummation-improved parton densities*, JHEP **03** (2018) 094, arXiv: 1801.10357 [hep-ph] (cit. on p. 4).

[51] 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 p. 6).

[52] 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. 6).

[53] 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. 6).

[54] W. Beenakker et al., *NNLL resummation for squark and gluino production at the LHC*, JHEP **12** (2014) 023, arXiv: 1404.3134 [hep-ph] (cit. on p. 6).

[55] W. Beenakker et al., *Towards NNLL resummation: hard matching coefficients for squark and gluino hadroproduction*, JHEP **10** (2013) 120, arXiv: 1304.6354 [hep-ph] (cit. on p. 6).

[56] W. Beenakker et al., *NNLL resummation for squark-antisquark pair production at the LHC*, JHEP **01** (2012) 076, arXiv: 1110.2446 [hep-ph] (cit. on p. 6).

[57] W. Beenakker et al., *Soft-gluon resummation for squark and gluino hadroproduction*, JHEP **12** (2009) 041, arXiv: 0909.4418 [hep-ph] (cit. on p. 6).

[58] A. Kulesza and L. Motyka, *Soft gluon resummation for the production of gluino-gluino and squark-antisquark pairs at the LHC*, Phys. Rev. D **80** (2009) 095004, arXiv: 0905.4749 [hep-ph] (cit. on p. 6).

[59] A. Kulesza and L. Motyka, *Threshold Resummation for Squark-Antisquark and Gluino-Pair Production at the LHC*, Phys. Rev. Lett. **102** (2009) 111802, arXiv: 0807.2405 [hep-ph] (cit. on p. 6).

[60] W. Beenakker, R. Höpker, M. Spira, and P. Zerwas, *Squark and gluino production at hadron colliders*, Nucl. Phys. B **492** (1997) 51, arXiv: hep-ph/9610490 (cit. on p. 6).

[61] J. Butterworth et al., *PDF4LHC recommendations for LHC Run II*, J. Phys. G **43** (2016) 023001, arXiv: 1510.03865 [hep-ph] (cit. on p. 6).

[62] T. Sjöstrand et al., *An introduction to PYTHIA 8.2*, Comput. Phys. Commun. **191** (2015) 159, arXiv: 1410.3012 [hep-ph] (cit. on p. 6).

[63] S. Frixione, P. Nason, and G. Ridolfi, *A positive-weight next-to-leading-order Monte Carlo for heavy flavour hadroproduction*, JHEP **09** (2007) 126, arXiv: 0707.3088 [hep-ph] (cit. on p. 6).

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

[65] 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 p. 6).

[66] 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 p. 6).

[67] R. D. Ball et al., *Parton distributions for the LHC run II*, JHEP **04** (2015) 040, arXiv: 1410.8849 [hep-ph] (cit. on p. 6).

[68] 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. 6).

[69] 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. 6).

[70] 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. 6).

[71] 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. 6).

[72] 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. 6).

[73] 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. 6).

[74] 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. 6).

[75] 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. 6).

[76] 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. 6).

[77] N. Kidonakis, "Top Quark Production," *Proceedings, Helmholtz International Summer School on Physics of Heavy Quarks and Hadrons (HQ 2013)* (JINR, Dubna, Russia, July 15–28, 2013) 139, arXiv: 1311.0283 [hep-ph] (cit. on p. 6).

[78] 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. 6).

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

[80] T. Gleisberg et al., *Event generation with SHERPA 1.1*, JHEP **02** (2009) 007, arXiv: 0811.4622 [hep-ph] (cit. on p. 6).

[81] ATLAS Collaboration, *Vertex Reconstruction Performance of the ATLAS Detector at $\sqrt{s}$ = 13 TeV*, ATL-PHYS-PUB-2015-026, 2015, URL: https://cds.cern.ch/record/2037717 (cit. on p. 5).

[82] 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. 5).

[83] 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. 5).

[84] ATLAS Collaboration, *Tagging and suppression of pileup jets with the ATLAS detector*,
ATLAS-CONF-2014-018, 2014, URL: https://cds.cern.ch/record/1700870 (cit. on p. 5).

[85] 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 pp. 5, 17).

[86] ATLAS Collaboration, *Jet reclustering and close-by effects in ATLAS run II*,
ATLAS-CONF-2017-062, 2017, URL: https://cds.cern.ch/record/2275649 (cit. on p. 5).

[87] D. Krohn, J. Thaler, and L.-T. Wang, *Jet trimming*, JHEP **02** (2010) 084,
arXiv: 0912.1342 [hep-ph] (cit. on p. 6).

[88] ATLAS Collaboration, *ATLAS flavour-tagging algorithms for the LHC Run 2 $pp$ collision dataset*, (2022), arXiv: 2211.16345 [physics.data-an] (cit. on p. 6).

[89] 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 pp. 6, 17).

[90] ATLAS Collaboration, *Measurement of the c-jet mistagging efficiency in $t\bar{t}$ events using $pp$ collision data at $\sqrt{s} = 13$ TeV collected with the ATLAS detector*, Eur. Phys. J. C **82** (2021) 95,
arXiv: 2109.10627 [hep-ex] (cit. on p. 6).

[91] ATLAS Collaboration,
*Calibration of the light-flavour jet mistagging efficiency of the b-tagging algorithms with Z+jets events using 139 $fb^{-1}$ of ATLAS proton–proton collision data at $\sqrt{s} = 13$ TeV*, (2023),
arXiv: 2301.06319 [hep-ex] (cit. on p. 6).

[92] 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).

[93] ATLAS Collaboration, *Electron efficiency measurements with the ATLAS detector using 2012 LHC proton–proton collision data*, Eur. Phys. J. C **77** (2017) 195, arXiv: 1612.01456 [hep-ex] (cit. on pp. 6, 7).

[94] 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. 7).

[95] 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).

[96] ATLAS Collaboration, *Characterisation and mitigation of beam-induced backgrounds observed in the ATLAS detector during the 2011 proton–proton run*, JINST **8** (2013) P07004,
arXiv: 1303.0223 [hep-ex] (cit. on p. 7).

[97] ATLAS Collaboration,
*Selection of jets produced in* 13 *TeV proton–proton collisions with the ATLAS detector*,
ATLAS-CONF-2015-029, 2015, URL: https://cds.cern.ch/record/2037702 (cit. on p. 7).

[98] T. Chen and C. Guestrin, *XGBoost: A Scalable Tree Boosting System*, (2016),
arXiv: 1603.02754 [cs.LG] (cit. on p. 9).

[99] 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. 9).

[100] L. Moneta et al., *The RooStats Project*, PoS **ACAT2010** (2010) 057, ed. by T. Speer et al.,
arXiv: 1009.1003 [physics.data-an] (cit. on p. 9).

[101] 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. 17).

[102] ATLAS Collaboration, *Performance of missing transverse momentum reconstruction with the ATLAS detector using proton–proton collisions at* $\sqrt{s}$ = 13 *TeV*, Eur. Phys. J. C **78** (2018) 903,
arXiv: 1802.08168 [hep-ex] (cit. on p. 17).

[103] 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. 17).

[104] ATLAS Collaboration, *Measurement of the Inelastic Proton–Proton Cross Section at* $\sqrt{s}$ = 13 *TeV with the ATLAS Detector at the LHC*, Phys. Rev. Lett. **117** (2016) 182002,
arXiv: 1606.02625 [hep-ex] (cit. on p. 17).

[105] ATLAS Collaboration,
*Luminosity determination in* $pp$ *collisions at* $\sqrt{s}$ = 13 *TeV using the ATLAS detector at the LHC*,
ATLAS-CONF-2019-021, 2019, URL: https://cds.cern.ch/record/2677054 (cit. on p. 17).

[106] G. Avoni et al., *The new LUCID-2 detector for luminosity measurement and monitoring in ATLAS*, JINST **13** (2018) P07017 (cit. on p. 17).

[107] 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. 18).

[108] R. D. Cousins, J. T. Linnemann, and J. Tucker,
*Evaluation of three methods for calculating statistical significance when incorporating a systematic uncertainty into a test of the background-only hypothesis for a Poisson process*,
Nucl. Instrum. Meth. A **595** (2008) 480, arXiv: physics/0702156 [physics.data-an]
(cit. on p. 22).

[109] 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. 22), Erratum: Eur. Phys. J. C **73** (2013) 2501.

[110] A. L. Read, *Presentation of search results: the* $CL_S$ *technique*, J. Phys. G **28** (2002) 2693
(cit. on p. 22).

# Auxiliary material

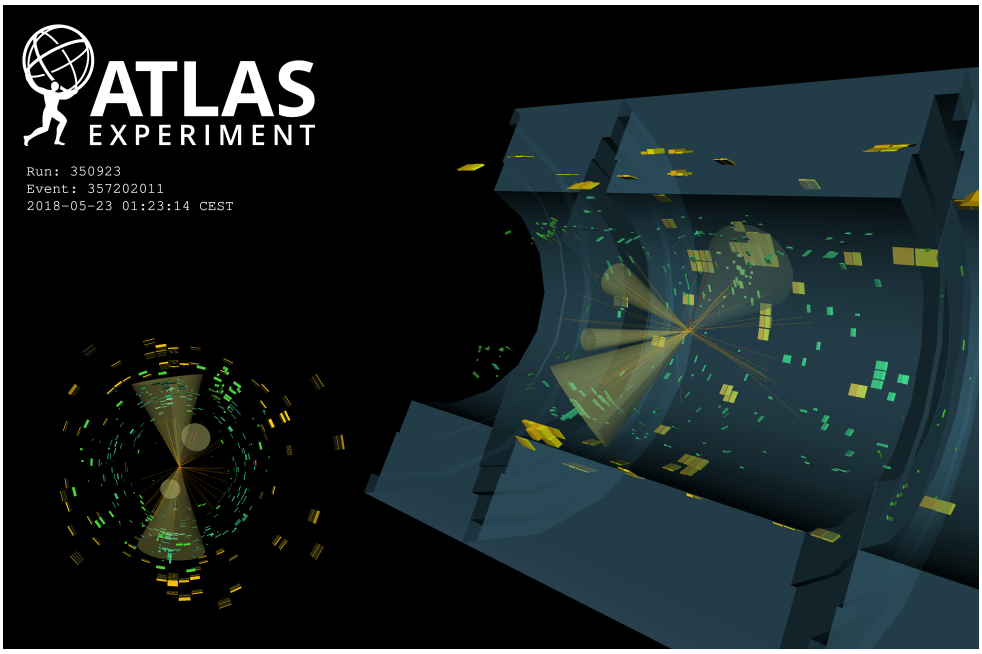



Figure 13: An event display for the low-mass channel. The yellow cones are jets, the dashed red line is $E_{\mathrm{T}}^{\mathrm{miss}}$ and the orange lines are tracks. This event has $E_{\mathrm{T}}^{\mathrm{miss}}$ = 14.0 GeV and $m_{\mathrm{eff}}$ = 283.3 GeV.

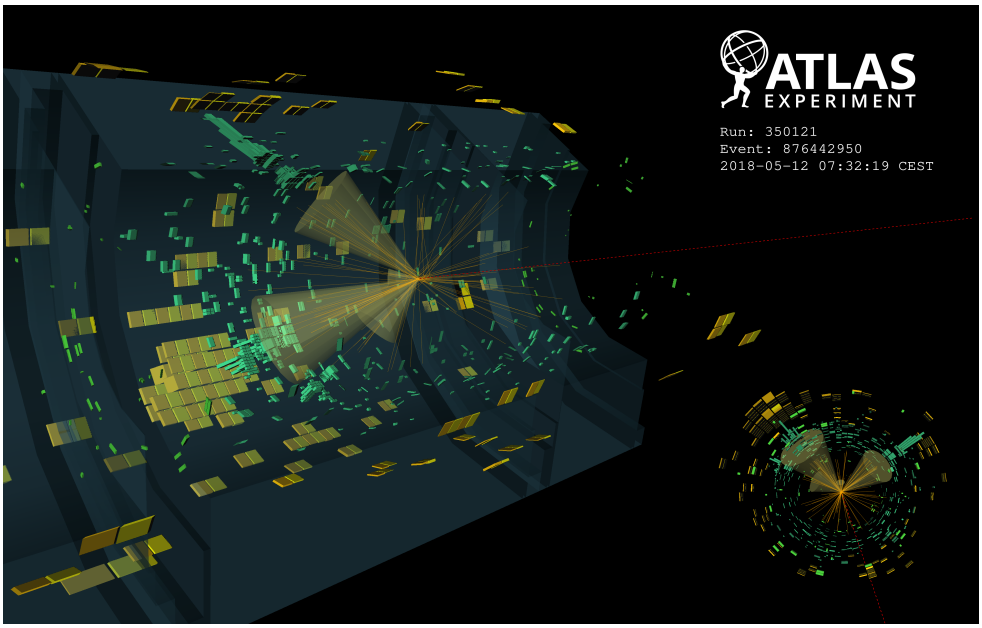



Figure 14: An event display for the high-mass channel. The yellow cones are jets, the dashed red line is $E_{\mathrm{T}}^{\mathrm{miss}}$ and the orange lines are tracks. This event has $E_{\mathrm{T}}^{\mathrm{miss}}$ = 550.3 GeV.