ATLAS searches for higgsinos with R-parity violating couplings in events with leptons

Otilia Anamaria Ducu, on behalf of the ATLAS Collaboration

Horia Hulubei National Institute of Physics and Nuclear Engineering, IFIN-HH, Bucharest, Romania

E-mail: oducu@cern.ch

Abstract. This document presents two searches for Supersymmetry through the direct production of pairs of higgsinos decaying into final states with leptons and (b-) jets. The analyses are performed using 139 fb^{-1} of the 13 TeV proton-proton collision data collected with the ATLAS detector. The methods used to estimate the Standard Model and detector backgrounds are discussed, as well as their shortcomings. Finally, results in selected signal regions, and some exclusion limits, are presented, illustrating the significant improvement over the previous exclusion limits.

Keywords: RPV SUSY, higgsino, same-charge leptons, multi-leptons

1. Introduction

Supersymmetry [1] (SUSY) remains one of the preferred extensions of the Standard Model (SM) of particle physics, despite the absence of experimental evidence in the LHC [2] Run 1 and Run 2 data. This is mainly because it can solve the SM gauge hierarchy problem without a large fine tuning of fundamental parameters, by predicting for each SM particle a super-partner with a half spin difference. Moreover, weakly interacting particles that are good dark matter candidates are present in the list of SUSY particles: the lightest neutralino ($\tilde{\chi}_1^0$) or the gravitino.

Only the weak production of higgsinos is studied in the searches discussed in this document. Two scenarios with R-parity violated [3] (RPV) are considered:

1) Higgsino bRPV [4]: the R-parity violation is obtained through lepton-number violation. Here, bilinear terms were introduced to the superpotential (bRPV). To ensure higgsino decays to light leptons, thus to suppress decays to tau leptons, $\tan\beta$ parameter is set to 5. The higgsinos are nearly degenerate, with a mass splitting of approximately 2 GeV. The considered production modes are $\tilde{\chi}_1^{\pm}\tilde{\chi}_1^0$, $\tilde{\chi}_1^{\pm}\tilde{\chi}_2^0$ and $\tilde{\chi}_1^0\tilde{\chi}_2^0$. All higgsino possible bRPV decays are allowed in the model, with the dominant decays being: $\tilde{\chi}_1^{\pm} \to W^{\pm}\nu_{\mu}$ and $\tilde{\chi}_{1,2}^0 \to W^{\pm}\ell^{\mp}, W^{\pm}\tau^{\mp}$. Some representative diagrams are shown in Figures 1(a) and 1(b).
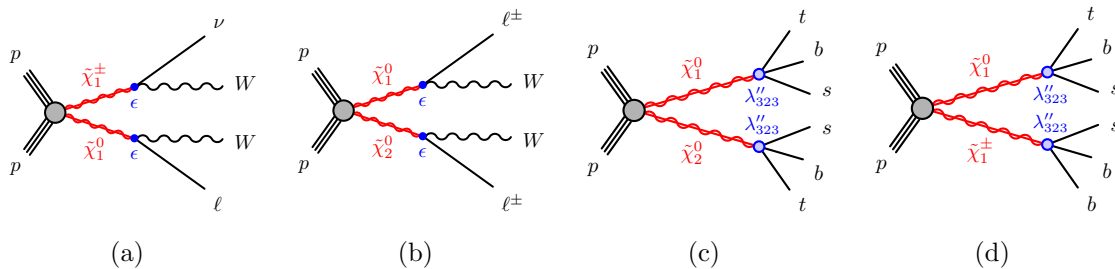
Figure 1. Representative diagrams illustrating the production and subsequent RPV decays of the higgsinos. Reused with permission from [4, 5].

2) Higgsino UDD RPV [4, 5]: the R-parity violation is obtained through baryon number violation, and the UDD-type coupling λ''_{323} is chosen to be non-vanishing. To ensure only prompt decays for $\tilde{\chi}^0_1$ and $\tilde{\chi}^0_2$ with masses > 180 GeV, λ''_{323} is set to have a value of $\mathcal{O}(10^{-3} - 10^{-2})$. The $\tilde{\chi}^0_1$ sparticle is always the LSP. The $\tilde{\chi}^\pm_1$ and $\tilde{\chi}^0_2$ sparticles are assumed to be effectively mass degenerate with the LSP, and all other electroweakinos are assumed to be decoupled and not considered in the model. Both $\tilde{\chi}^0_2$ and $\tilde{\chi}^0_1$ decay in the same way, $\tilde{\chi}^0_{1,2} \to tbs$ with a 100% BR; $\tilde{\chi}^\pm_1$ decays to sbb with a 100% BR. For this simplified model, the considered diagrams are shown in Figures 1(c) and 1(d).

- Signal event samples for these models are generated using MadGraph5 aMC@NLO interfaced to Pythia8 for the modelling of the parton showering, hadronisation and underlying event [4, 5].

As can be seen in Figure 1, the considered RPV models have either top quarks or W bosons in the decay chains, ensuring the possibility of leptonic final states. The searches presented in this document are performed only with events containing one lepton, two same electric charge leptons or three leptons. These, as well as the zero lepton channels, are being extensively considered by both ATLAS and CMS experiments, and the obtained results were published in numerous articles [6, 7].

Even if the signatures including leptons are characterized by small branching ratios, they are well motivated as the level of SM background is quite low. In this document, the focus is mainly on the work done in the channels with two same-charge or three leptons, as the background estimation is more challenging. However, selected results in the one lepton channel are presented for the UDD RPV model illustrated in Figure 1(d).

For these studies, the LHC Run 2 data set of proton-proton collisions at $\sqrt{s} = 13$ TeV recorded by the ATLAS detector [8], and corresponding to an integrated luminosity of 139 fb^{-1} is used. The ATLAS detector is a multipurpose particle detector with a forward–backward symmetric cylindrical geometry and a near 4π coverage in solid angle.† It consists of an inner tracking detector surrounded by a thin superconducting solenoid providing a 2 T axial magnetic field, electromagnetic and hadron calorimeters,

---

† ATLAS uses a right-handed coordinate system. The pseudorapidity is defined in terms of the polar angle θ as $\eta = -\ln\tan(\theta/2)$. Angular distance is measured in units of $\Delta R \equiv \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$.

and a muon spectrometer.

Interesting events are selected by the first-level trigger system implemented in custom hardware, followed by selections made by algorithms implemented in software in the high-level trigger [9]. For the analysis performed only with two same-charge or three leptons, events are selected with the lowest unprescaled di-lepton triggers [10, 11] and missing transverse energy ($E_{\mathrm{T}}^{\mathrm{miss}}$) based triggers [12]. For the analysis including also the one lepton event topology, events are selected with the lowest unprescaled single lepton triggers. The employed trigger selection is ensuring a maximum, and a rather constant wrt. the lepton transverse momentum (p_T) trigger efficiency.

An extensive software suite [13] was used in data simulation, in the reconstruction and analysis of real and simulated data, in detector operations, and in the trigger and data acquisition systems of the experiment.

2. Strategy to look for RPV SUSY

The following strategy is used to search for higgsino signals. At first, dedicated discovery and exclusion signal regions are optimized for each RPV model. This is done via a complete scan of different sets of cuts on the background discriminant kinematic variables representative for the final state. During the optimization, the signal sensitivity is evaluated through its potential discovery significance. Only Monte-Carlo (MC) simulations are used, and to account for uncertainties on the background prediction a 30% systematic uncertainty is considered. To ensure the best sensitivity, more than one signal region per RPV model is defined. The final signal regions can overlap.

Once the final signal regions are selected, the SM and detector background sources are identified and further estimated as best as possible. An extensive validation is also performed, to ensure the correctness of the employed methods. These measurements and checks are done with the help of control and validations regions that are defined for the main background sources.

A simultaneous fit method is used to compute the final signal and background uncertainty estimates, and to statistically test the compatibility between data and background estimation in the signal regions. The method is based on a profile-likelihood ratio test and is implemented in the Histfitter tool [14]. Three different fits are performed to get the final results. A background-only fit [14] is used to estimate the total background in the signal and validation regions, without assumptions on the signal model. These background predictions are independent of the observed number of data events in each signal and validation region. A background-only fit is used to also estimate the normalization factor (and its uncertainty) of the WZ+jets background.

A model-dependent signal fit [14] is used to obtain the exclusion limits. The signal contribution is accounted for, and is given by the number of signal events estimated in the signal region(s) considered in the fit. The fit is performed only in the exclusion signal region(s) [4], or simultaneously in the signal and control regions [5]. It is important to note that more than one signal or control region can be included in this fit only if they

are orthogonal (thus, statistically independent).

A model-independent signal fit [14] is used to obtain the model-independent upper limits on the number of events beyond the expected number of events in each discovery signal region. This information is important especially for theorists, as it can be used to probe any BSM model. As for the model-dependent signal fit, both signal and control regions are used. No assumptions are made for the signal model, and the number of signal events in the signal region is added as a parameter to the fit.

3. Signal region definitions

A detailed description of E_T^{miss}, lepton and (b-) jet objects used to get the results presented in this document is given in [4, 5]. Baseline (signal) leptons collections are defined using loose (tight) lepton identification (plus isolation) criteria, as well as requirements on the transversal and longitudinal impact parameters. The leading and sub-leading (third-leading) leptons are required to have $p_T > 15$ GeV (10 GeV). The selection requirements are a compromise between a reasonably-efficient identification of prompt leptons and a good rejection of background leptons. For example, this efficiency can reach values higher than 90% for leptons with $p_T > 100$ GeV, in a $Z \to \ell^\pm \ell^\mp$ event selection. Jet candidates are reconstructed using the anti-k_t algorithm, with radius parameter $R = 0.4$ and particle-flow objects as inputs [15]. For the b-jet identification, a DL1r b-tagging algorithm [16, 17] is used, with an average identification efficiency of 70% in a $t\bar{t}$ event selection. Only (b-) jets with $p_T > 20$ GeV are used further.

In addition to the counting of the objects discussed above, the discriminating variables presented below are used to define the signal regions.

- E_T^{miss}, that has moderate values in signal events as it comes only from neutrinos.
- The stransverse mass, m_{T2}, an event variable that is correlated to the masses of an unseen pair of particles that are presumed to have decayed semi-invisibly into particles that are seen. It is defined as a function of the momenta of two visible particles and the E_T^{miss} in the event.
- The inclusive effective mass, m_{eff}, computed by summing the signal leptons p_T, jets p_T and E_T^{miss}. It represents the total energy in the transverse plane, and is highly dependent on the initial particle mass.
- The invariant mass of the electron and muon pairs that have different-charges ($m_{e^\pm e^\mp}, m_{\mu^\pm \mu^\mp}, m_{SFOC}$). It helps to reject background events with Z bosons.
- The sum of signal leptons p_T, $\sum p_T(\ell)$, and the sum of jets p_T, $\sum p_T^{jet}$, are variables that help given the high numbers of leptons and jets in the signal final state.
- The ratio of sum of the b-jets p_T, and the sum of all jets p_T, $\sum p_T^{b-jet} / \sum p_T^{jet}$, is a very powerful discriminant when many b-jets are in the final state.
- The minimum angular distance between the leading lepton and the selected jets, $\Delta R(\ell_1, jet)_{min}$. It is a key variable used to separate the UDD RPV signal illustrated in Figure 1(c) from background events coming from e.g. $t\bar{t}$ processes.

Table 1. Exclusion and discovery signal regions defined for the bRPV model. $N_{\mathrm{B}}(\ell)$ ($N_{\mathrm{S}}(\ell)$) stands for the number of baseline (signal) leptons. Reused with permission from [4].

	$\mathrm{SR}_{2\ell}^{\mathrm{bRPV}}$	$\mathrm{SR}_{3\ell}^{\mathrm{bRPV}}$
$N_{\mathrm{B}}(\ell)$	≥ 2	
$N_{\mathrm{S}}(\ell)$	$= 2$	$= 3$
$p_T\,(\ell)$	≥ 20 GeV for (sub)leading leptons	
Charge(ℓ)	same-charge	$-$
m_{T2}	≥ 60 GeV	≥ 80 GeV
$E_{\mathrm{T}}^{\mathrm{miss}}$	≥ 100 GeV	≥ 120 GeV
m_{eff}	$-$	≥ 350 GeV
$n_{b-\mathrm{jets}}$	$= 0$	$-$
$n_{\mathrm{jets}}\ (p_T > 25)$ GeV	$-$	≥ 1
$n_{\mathrm{jets}}\ (p_T > 40)$ GeV	≥ 4	$-$
$m_{e^{\pm}e^{\mp}},\ m_{\mu^{\pm}\mu^{\mp}}$	$-$	$\notin [81, 101]$ GeV

Table 2. Discovery signal regions defined for the UDD RPV model with two same-charge leptons in the final state. Reused with permission from [4].

	$\mathrm{SR}_{2\ell1b}^{\mathrm{RPV}}$		$\mathrm{SR}_{2\ell2b}^{\mathrm{RPV}}$			$\mathrm{SR}_{2\ell3b}^{\mathrm{RPV}}$		
	L	M	L	M	H	L	M	H
$N_{\mathrm{B}}(\ell)$	$= 2$							
$N_{\mathrm{S}}(\ell)$	$= 2$							
$p_T(\ell)$	> 25 GeV							
Charge(ℓ)	same-charge							
$n_{b-\mathrm{jets}}$	$= 1$		$= 2$			≥ 3		
$\sum p_T(\ell)$	≥ 100 GeV		$-$			$-$		
$E_{\mathrm{T}}^{\mathrm{miss}}$	≥ 100 GeV	≥ 50 GeV	≥ 80 GeV			≥ 20 GeV		
$n_{\mathrm{jets}}\ (p_T > 25\ \mathrm{GeV})$	≤ 2	$= 2$ or $= 3$	≤ 3	$= 3$ or $= 4$	≥ 5 and ≤ 6	≤ 3	≤ 3	≤ 6
$\sum p_T^{b-\mathrm{jet}}/\sum p_T^{\mathrm{jet}}$	≥ 0.7	≥ 0.45	≥ 0.9	≥ 0.75	$-$	≥ 0.8	≥ 0.8	≥ 0.5
$\sum p_T^{\mathrm{jet}}$	≥ 120 GeV	≥ 400 GeV	≥ 300 GeV	≥ 420 GeV	≥ 420 GeV	$-$	$-$	≥ 350 GeV
$\Delta R(\ell_1, \mathrm{jet})_{\mathrm{min}}$	≤ 1.2	≤ 1.0	≤ 1.0	≤ 1.0	≤ 1.0	≤ 1.5	$-$	≤ 1.0
$\Delta R(\ell^{\pm}, \ell^{\pm})$	≥ 2.0	≥ 2.5	≥ 2.5	≥ 2.5	≥ 2.0	≥ 2.0	$-$	≥ 2.0

- The angular distance between the two same-charge leptons, $\Delta R(\ell^{\pm}, \ell^{\pm})$. It helps for the UDD RPV model, as the same-charge leptons are typically separated in the detector, for this signal process.

Two non-overlapping signal regions are defined for the bRPV model, one with exactly two same-charge leptons ($\mathrm{SR}_{2\ell}^{\mathrm{bRPV}}$), and one with exactly three leptons ($\mathrm{SR}_{3\ell}^{\mathrm{bRPV}}$). These regions (Table 1) are used both as discovery and exclusion signal regions.

For the UDD RPV model shown in Figure 1(c), three sets of non-overlapping discovery signal regions are selected, as shown in Table 2. Each set has two or three overlapping signal regions that are defined with two same-charge lepton final states, and with exactly one, exactly two and at least three b-jets in the event. These overlapping regions were optimized using a signal generated with $\tilde{\chi}_1^0$ mass set to 180 GeV ($\mathrm{SR}_{2\ell\mathrm{n}b}^{\mathrm{RPV}}$

L), 200 GeV ($\text{SR}_{2\ell nb}^{\text{RPV}}$ M) or 300 GeV ($\text{SR}_{2\ell nb}^{\text{RPV}}$ H).

Selected exclusions signal regions for both UDD RPV production modes (Figures 1(c) and 1(d)) are defined with two same-charge leptons, or with at least one lepton.‡ In order to obtain the best sensitivity, events are further categorized into regions based on jet and b-jet multiplicities, and a neural network (NN) discriminant is introduced in some of the jet and b-jet multiplicity regions. These regions, and their background estimation, are discussed in detail in [5]. For illustration, a (quite rare) data event display in a two same-charge signal region with six jets is shown in Figure 2.
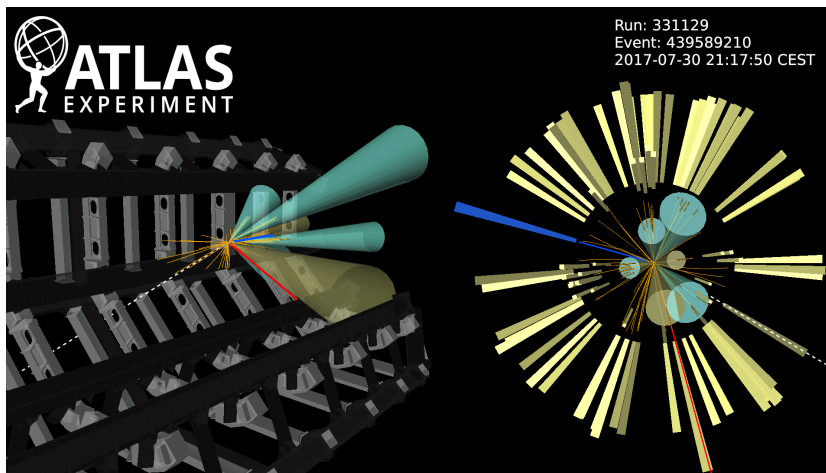

Figure 2. Event display in a UDD RPV signal region, containing a muon and electron with same electric charge, and 6 jets. The signal muon (electron) is indicated by the red (blue) line, and has $p_T = 35$ GeV (61 GeV). The 6 jets from this event have a p_T from 41 GeV to 145 GeV, and among them four are b-jets (shown with a cyan cone). $E_{\text{T}}^{\text{miss}}$ has a value of 31 GeV and is shown with a dotted white line. Figure reused with permission from [18].

To illustrate the statistics obtained after key selections in the analysis, a cut flow is shown in Figure 3. It is obtained for the RPV UDD model illustrated in Figure 1(c), for a signal point generated with $\tilde{\chi}_1^0$ mass set to 250 GeV. The column labeled N_{raw} shows the number of generated events, while N_{events} shows the expected number of events with a luminosity of 139 fb^{-1}. The "Total Eff" column shows the cut flow efficiency with respect to all weighted events.

4. Backgrounds

The background sources can be separated in two main categories. The first category is populated by the SM processes that give one, two same-charge or three prompt lepton final states, like W+jets, WZ or $t\bar{t}W$. In the second category are the detector backgrounds, such as events with electron charge flip or the fake/non-prompt leptons. In this document, only the background strategy used to get the results in the signal

‡ The one lepton regions do not include events entering the two same-charge lepton regions.

$\tilde{\chi}^0_{1,2} \to tbs$, $(m_{\tilde{\chi}^0_{1,2}} = 250$ GeV)	N_{raw}	N_{events}	Total Eff
All Events	269512	14491.03	100%
Lepton trigger	139169	7467.78	51.53%
≥ 1 baseline lepton	124548	6677.48	46.08%
≥ 1 signal lepton	103256	5533.07	38.18%
Leading lep $p_{\text{T}} \geq 27$ GeV	101018	5413.64	37.36%
Signal lepton is leading lepton	99585	5336.54	36.83%

Jet p_{T} threshold	20 GeV	40 GeV	60 GeV	80 GeV	100 GeV
1ℓ category	Total Eff.	Total Eff.	Total Eff.	Total Eff.	Total Eff.
≥ 4 jets	30.8%	19.7%	9.0%	3.8%	1.7%
$== 4$ jets	4.7%	9.2%	5.8%	2.8%	1.3%
$== 5$ jets	7.6%	6.4%	2.4%	0.8%	0.3%
$== 6$ jets	8.2%	2.9%	0.7%	0.1%	0.0%
$== 7$ jets	5.8%	0.9%	0.1%	0.0%	0.0%
$== 8$ jets	2.9%	0.2%	0.0%	0.0%	0.0%
$== 4$ jets, ≥ 4 b-tags	0.05%	0.08%	0.05%	0.02%	0.01%
$== 5$ jets, ≥ 4 b-tags	0.20%	0.17%	0.06%	0.02%	0.01%
$== 6$ jets, ≥ 4 b-tags	0.39%	0.15%	0.03%	0.01%	0.00%
$== 7$ jets, ≥ 4 b-tags	0.42%	0.06%	0.01%	0.01%	0.00%
$== 8$ jets, ≥ 4 b-tags	0.27%	0.02%	0.00%	0.00%	0.00%
$== 4$ jets, ≥ 4 b-tags, NN_{4j} bin 4	0.02%	–	–	–	–
$== 5$ jets, ≥ 4 b-tags, NN_{5j} bin 4	0.06%	–	–	–	–
$== 6$ jets, ≥ 4 b-tags, NN_{6j} bin 4	0.12%	–	–	–	–
$== 7$ jets, ≥ 4 b-tags, NN_{7j} bin 4	0.13%	–	–	–	–
$== 8$ jets, ≥ 4 b-tags, NN_{8j} bin 4	0.10%	–	–	–	–
$2\ell^{\text{sc}}$ category $(m^{\ell j} < 155)$ GeV					
≥ 4 jets	1.2%	0.5%	0.2%	0.1%	0.0%
$== 4$ jets	0.4%	0.3%	0.1%	0.0%	0.0%
$== 5$ jets	0.4%	0.1%	0.0%	0.0%	0.0%
$== 6$ jets	0.3%	0.0%	0.0%	0.0%	0.0%
$== 7$ jets	0.1%	0.0%	0.0%	0.0%	–
$== 8$ jets	0.0%	0.0%	0.0%	–	–
$== 4$ jets, ≥ 3 b-tags	0.1%	0.1%	0.1%	0.0%	0.0%
$== 5$ jets, ≥ 3 b-tags	0.1%	0.1%	0.0%	0.0%	0.0%
$== 6$ jets, ≥ 3 b-tags	0.1%	0.0%	0.0%	0.0%	0.0%
$== 7$ jets, ≥ 3 b-tags	0.0%	0.0%	0.0%	0.0%	–
$== 8$ jets, ≥ 3 b-tags	0.0%	0.0%	–	–	–

Figure 3. Cut flow for the RPV UDD model illustrated in Figure 1(c). $2\ell^{\text{SC}}$ stands for two same-charge leptons, and $m^{\ell j}$ is a variable that is very powerful at rejecting backgrounds that do not have lepton decays from two top quarks (as the signal) [5]. Selections that have not been evaluated in the analysis or are not applicable are denoted with a dash (–). Reused with permission from [18].

regions shown in Tables 1 and 2 is discussed. The background studies done for the UDD RPV exclusion signal regions are presented in detail in [18].

4.1. WZ+jets background

The WZ+jets process is a dominant background in the 0 b-jet bRPV signal regions, and to correct a shape missmodeling seen in the jet multiplicity [19], a dedicated control region is defined (Table 3). Using a background only fit, a normalization factor of 0.88 ± 0.30 is obtained, and further used to scale the WZ+jet MC simulations. The accuracy of this approach is checked in two validation regions, $\text{VR}WZ_{4j}^{WZ,\text{(b)RPV}}$ and $\text{VR}WZ_{5j}^{WZ,\text{(b)RPV}}$ (Table 3). The other SM backgrounds are estimated using only MC simulations. For the $t\bar{t} + Z$ and $t\bar{t} + W$ processes, one common validation region, $\text{VR}t\bar{t} + V^{WZ,\text{(b)RPV}}$, is defined (Table 3). This approach is motivated, as the signal

Table 3. The control and validation regions defined for the bRPV, and UDD RPV discovery signal regions. Reused with permission from [4].

	$\mathrm{CR}WZ_{2j}^{WZ,(b)\mathrm{RPV}}$	$\mathrm{VR}WZ_{4j}^{WZ,(b)\mathrm{RPV}}$	$\mathrm{VR}WZ_{5j}^{WZ,(b)\mathrm{RPV}}$	$\mathrm{VR}t\bar{t} + V^{WZ,(b)\mathrm{RPV}}$
$N_{\mathrm{B}}(\ell)$	$= 3$			≥ 2
$N_{\mathrm{S}}(\ell)$	$= 3$			≥ 2
(ℓ)	$p_T > 20$ GeV for (sub)leading leptons			$p_T > 30$ GeV for same-charge leptons
Charge(ℓ)	$-$			same-charge
$n_{b-\mathrm{jets}}$	$= 0$	$= 0$	$= 0$	≥ 1
n_{jets} ($p_T \geq 25$ GeV)	≥ 2	≥ 4	≥ 5	≥ 3 with $p_T > 40$ GeV
Additional requirements	$50 < E_{\mathrm{T}}^{\mathrm{miss}} < 150$ GeV $m_{\mathrm{eff}} < 1$ TeV $81 < m_{\mathrm{SFOC}} < 101$ GeV	$50 < E_{\mathrm{T}}^{\mathrm{miss}} < 250$ GeV $m_{\mathrm{eff}} < 1.5$ TeV $81 < m_{\mathrm{SFOC}} < 101$ GeV $-$ $-$ $-$		$-$ $-$ $-$ $\Delta R(\ell_1, jet)_{\min} > 1.1$ $\sum p_T^{b-\mathrm{jet}} / \sum p_T^{\mathrm{jet}} > 0.4$ $E_{\mathrm{T}}^{\mathrm{miss}}/m_{\mathrm{eff}} > 0.1$
	To ensure negligible signal contamination veto any events entering the bRPV signal regions $n_{b-\mathrm{jets}} \geq 3$ $n_{b-\mathrm{jets}} \geq 1,\ n_{\mathrm{jets}} \geq 4$ ($p_T > 50$ GeV), $E_{\mathrm{T}}^{\mathrm{miss}} > 130$ GeV $n_{b-\mathrm{jets}} = 0,\ n_{\mathrm{jets}} \geq 3$ ($p_T > 50$ GeV), $E_{\mathrm{T}}^{\mathrm{miss}} > 130$ GeV $n_{b-\mathrm{jets}} = 0,\ n_{\mathrm{jets}} \geq 5$ ($p_T > 50$ GeV)			
Purity	85%	84%	77%	62%

regions themselves have a mixture of $t\bar{t} + Z$ and $t\bar{t} + W$ processes, as well as other top processes that contribute with prompt leptons. Ideally, separate validation (and even control) regions should be defined for $t\bar{t}+W$ events, but is difficult given the large amount of fake/non-prompt leptons and the low statistics in data when applying selections closer to the signal regions.

4.2. Electron charge flip background

Backgrounds with electron charge flip, or with a wrong charge measurement, are relevant only for the same-charge lepton final states. When a high p_T electron interacts with the detector material, it can radiate a hard photon that converts into an electron-positron pair. During the reconstruction, if the energy of the radiated electron is too small, then the energy deposited in the calorimeter can be matched with the radiated positron track. If this happens, the charge of the initial electron is wrongly measured. Of course, this can be initiated also by positron candidates. § For muons, such processes are very rare, and the muon charge flip is found to be negligible.

Studies based on MC simulations show that in the signal regions, this type of background originates mainly from $Z \to e^{\pm}e^{\mp}$ (in $\mathrm{SR}_{2\ell}^{\mathrm{bRPV}}$) and dileptonic $t\bar{t}$ processes. As illustrated in Figure 4(a), this type of background is highly suppressed with a boosted decision tree discriminant (BDT) exploiting additional tracks in the vicinity of the electron and track-to-cluster matching variables [20]. Thus, requirements on the BDT output are applied for both baseline and signal electrons.

§ For simplicity, the generic term electron is further used both for electrons and positrons.
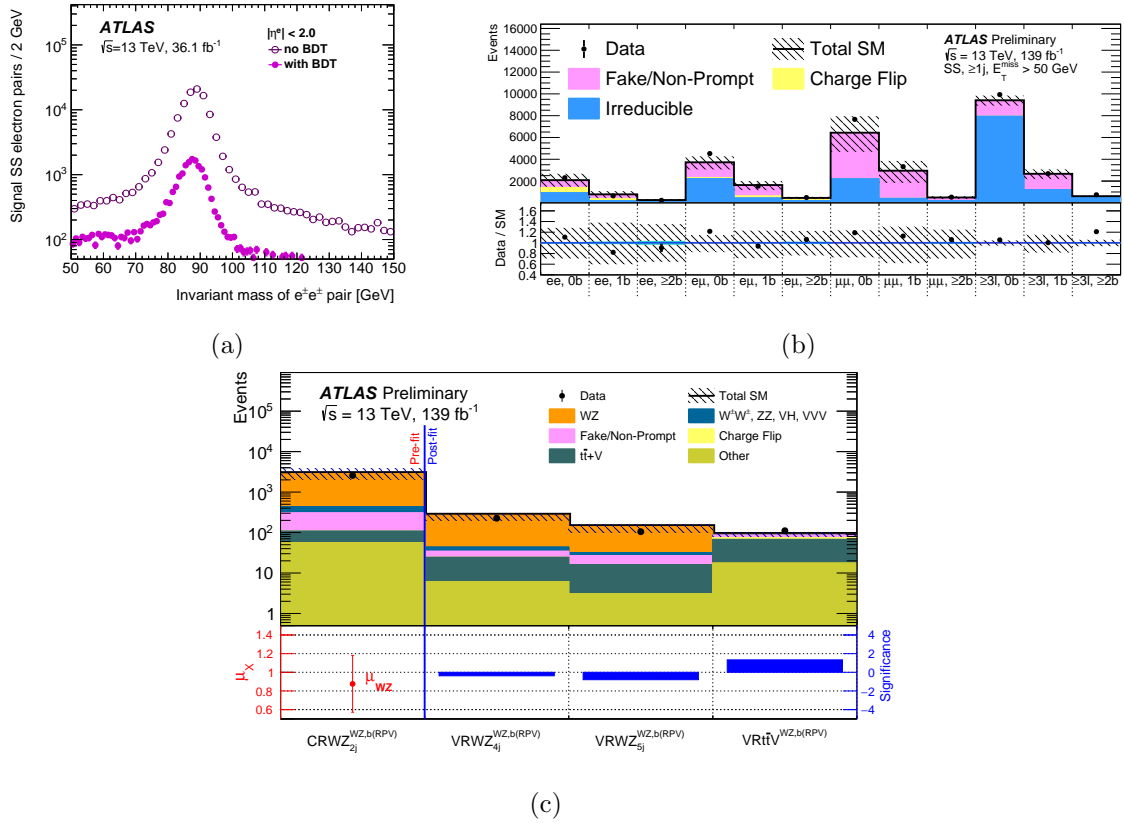
(a)

(b)

(c)

Figure 4. (a) Distribution of the invariant mass of signal $e^{\pm}e^{\pm}$ pairs with (full markers) and without (open markers) electron charge flip BDT selection applied. Reused with permission from [21]. (b) Observed data compared with the background estimation after a loose preselection. The event yields are classified as a function of the lepton flavour and multiplicity, as well as the number of b-jets. The error bars include the statistical uncertainty and the full uncertainties for the data-based background estimates. Reused with permission from [4]. (c) Results in control and validation regions. All uncertainties are included. Reused with permission from [4].
SS stands for same-charge lepton pairs.

To estimate the electron charge flip background in a given region, data events with different-charge leptons, but otherwise passing an identical selection, are taken and weighted with the charge flip probability, ξ. This weight can be written as $w_{\text{flip}} = \xi_1(1 - \xi_2) + (1 - \xi_1)\xi_2$, with ξ_i set to zero if the lepton $i(=1 \text{ or } 2)$ is a muon. ξ is computed with a likelihood based method [20], using $Z \to e^{\pm}e^{\mp}$ events, in bins of electron p_T and η. Checks are done using also $t\bar{t}$ MC simulations. The electron charge flip probabilities are found to be $\mathcal{O}(10^{-6})$ in the low p_T - η regions, increasing to $\mathcal{O}(1\%)$ in high p_T - η regions. Main sources of uncertainties on the ξ parameter come from the low statistics in the measurement data sample, from the final-state radiation multijet production background estimation, and from the tight requirements on the di-electron invariant mass imposed to ensure a cleaner $Z \to e^{\pm}e^{\mp}$ data selection. All the statistical and systematic uncertainties estimated for ξ are propagated to the final electron charge flip estimation, leading to a 5% to 45% uncertainty in the predicted signal region yields

for this background source.

4.3. Fake/non-prompt lepton background

The fake/non-prompt (not isolated) leptons stem from weak hadron decays, photon conversions in detector material or trident electrons. Most common properties shared by these objects are an incorrect response to the lepton identification requirements, non-zero impact parameters, and are not well isolated. These characteristics are used to discriminate the fake/non-prompt leptons against the prompt and isolated (real) leptons, and ultimately to estimate this source of background. According to MC simulations, in the signal regions, this type of background originates mainly from semi-leptonic or di-leptonic $t\bar{t}$ processes. The cases with a prompt leading lepton and a fake/non-prompt sub-leading lepton are dominant, and events with two fake/non-prompt leptons are negligible.

This type of background is evaluated with the data-based matrix method [22, 23], and cross-checked with a MC based method (MC-Template). The matrix method relies on loose and tight lepton categories: the loose (tight) leptons are baseline lepton candidates that fail (pass) the signal lepton requirements. It relates the number of events containing prompt or fake/non-prompt leptons to the number of observed events with tight or loose leptons, using the probability for loose prompt or fake/non-prompt leptons to satisfy the tight lepton criteria.

The real lepton efficiency, or the probability for prompt leptons to pass the signal lepton requirements, is measured in $t\bar{t}$ MC simulations using generator level information, as a function of p_T and η. These "true" efficiencies are corrected with dedicated scale factors that account for the data – MC simulation differences in the reconstructed energy, or in the efficiency of identification and isolation of leptons. For a p_T around 15 GeV they are found to be around 50–60% (70%) for electrons (muons), increasing up to 98% (99%) for leptons with $p_T > 100$ GeV (60 GeV). Main uncertainties are from systematic sources associated to the correction scale factors, and come mainly from the lepton identification and isolation efficiency measurements. The total uncertainties vary between 0.3–7% (0.1–3%) for electrons (muons), depending on p_T and η.

The fake/non-prompt lepton efficiency, or the probability for fake/non-prompt leptons to pass the signal lepton requirements, is the most challenging to measure. The control regions enriched in this type of background should have a similar composition as the signal regions, similar kinematics, and be definable also in data. This is because not all sources of fake/non-prompt leptons are well modeled in MC simulations, thus an efficiency measurement in data is mandatory. Six semileptonic or dileptonic $t\bar{t}$ enriched control regions are defined with $e^{\pm}e^{\pm}$, $e^{\pm}\mu^{\pm}$, $\mu^{\pm}e^{\pm}$, $\mu^{\pm}\mu^{\pm}$, $\ell^{\pm}\ell'^{\mp}e^{\mp}$ and $\ell^{\pm}\ell'^{\mp}\mu^{\mp}$ lepton configurations, at least (one) two (b-) jets and $E_T^{\text{miss}} > 30$ GeV. The minimum E_T^{miss} requirement helps to remove most of the QCD events with two fake/non-prompt leptons. Upper cuts on E_T^{miss} and m_{eff} are placed to minimize the signal contamination. The fake/non-prompt efficiency is measured per lepton flavor using the Tag & Probe [20]

Table 4. The systematic uncertainties associated to the fake/non-prompt lepton efficiency, measured in $t\bar{t}$ MC simulation [4, 23].

Source of uncertainty	Electrons	Muons
Extrapolation to higher p_T	0% (covered by measurement uncertainties and/or next item)	
Underlying jet kinematics / event topology	$\pm 30\%$ for $p_T < 100$ GeV	$\pm 30\%$ for $p_T < 30$ GeV
	$\pm 50\%$ for $p_T > 100$ GeV	$^{+30\%}_{-50\%}$ for $30 < p_T < 50$ GeV
		$^{+30\%}_{-80\%}$ for $p_T > 50$ GeV

method, in bins of p_T and η. The tag lepton must pass the signal lepton requirements and tighter identification and isolation operating points, to ensure it is the prompt lepton originating from one of the leptonically-decaying top quarks. In the measurement, the tag is taken to be the leading lepton from the same-charge pair. The other lepton from the pair is the probe, and assumed to be the fake/non-prompt one. The efficiencies in data are then simply obtained as the fraction of probe leptons satisfying the signal lepton requirements, after subtracting the expected contributions from SM processes with two or three prompt leptons and, when relevant, the electron charge flip. They are measured independently in each control region, and a weighted combination gives the final efficiency. The fake/non-prompt lepton efficiency has values of $\approx$10–20% for both electrons and muons up to p_T of around 45 GeV, and increases to 30–40% for $p_T > 60$ GeV. When two b-jets are present in the event, the fake/non-prompt lepton efficiencies are much higher, with an increase of up to a factor two.

As shown in Table 4, various sources of systematic uncertainties are considered, to account for all variations in composition and event kinematics [4, 23]. The statistical uncertainty, the uncertainties associated to the electron charge flip subtraction, as well as a 30% uncertainty on the SM background subtraction are also considered in the total uncertainty associated to the fake/non-prompt lepton efficiency. All these uncertainties are propagated through the matrix method to the fake/non-prompt background estimate, leading to a 10% to 40% uncertainty in the predicted signal regions yields for this background source.

The data and estimated background agree well, within the assigned uncertainties, as shown in Figure 4(b). The very good agreement in control and validation regions defined in Table 3 is illustrated in Figure 4(c): here the electron charge flip background has a contribution only in the ttZ/W validation region, as the WZ+jets regions are defined with three leptons.

The MC-Template method uses MC simulations to extrapolate the detector background predictions from control regions defined with low jet multiplicities and low m_{eff} or E_T^{miss}, to the signal regions. The main assumption is that the MC simulations describe the kinematic distributions correctly and predict accurately e.g. the rate of fake/non-prompt leptons up to a global factor independent of the event kinematics and the process type. This makes the method a suitable cross-check for the matrix method that assumes that the lepton fake/non-prompt efficiencies are the same in the control and
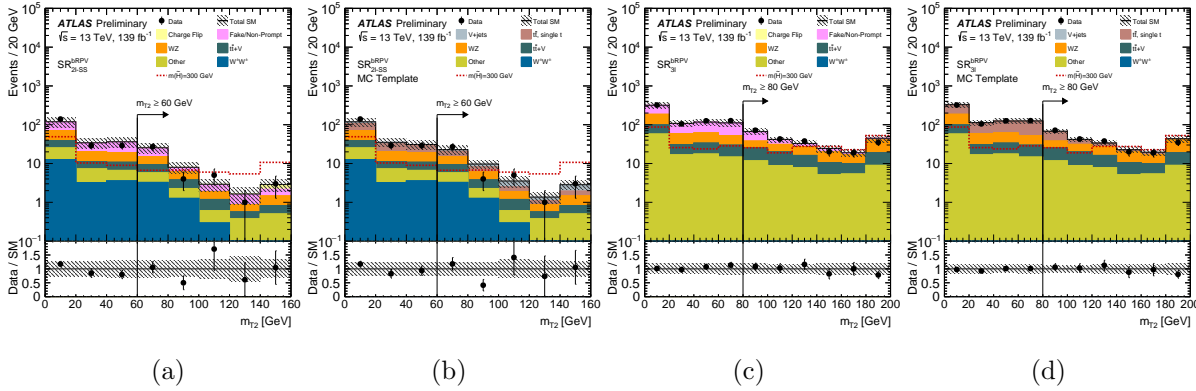
Figure 5. Observed data versus the estimated background close to the selected signal regions. The bRPV signal regions requirements are all applied except for that on m_{T2}. Expected distributions of representative signal mass points are overlaid, and shown with interrupted red lines. In Figures a and c (b and d) the detector background is estimated with the data-based (MC-Template) methods described in the text. The vertical black lines and the corresponding arrows indicate the corresponding signal region. All uncertainties are included. In each figure the bottom panel shows the ratio of data to the estimated background. Reused with permission from [4].

signal regions, and independent of the selection requirements. The second assumption is that the fake/non-prompt fractions are uncorrelated in events with multiple fake leptons.

The MC-Template fake/non-prompt lepton control regions are similar to the ones used to measure the fake/non-prompt lepton efficiency needed by the matrix method. As the $t\bar{t} + W$ processes are not perfectly modeled at low m_{eff} in MC simulations, a dedicated control region is defined. The $\mathrm{CR}WZ_{2j}^{WZ,(\mathrm{b})\mathrm{RPV}}$ region is also considered. With these control regions six correction factors are measured, separately for $t\bar{t} + W$, WZ, electron charge flip, fake electron, non-prompt electron, and fake/non-prompt muon contributions, via a simultaneous (1D or 2D) fit ‖. The fit uses a likelihood function defined as the product of the Poisson probabilities describing the observed events in the binned distributions from the expected number of events rescaled by the six correction factors which are left free to float in the fit. Among the discriminant variables used are the leading and sub-leading lepton p_T and η, m_{eff} or the jet multiplicity.

Several sources of uncertainties are considered. The statistical uncertainty is given by the fit, and corresponds to how much the correction factors need to be varied for one standard deviation change in the likelihood function. Systematic uncertainties are estimated by looking at the differences in the correction factors between various discriminant variables and/or control regions used in the fit.

4.4. Background validation

The good observed data – estimated background agreement close to the selected signal regions is illustrated in Figures 5–8. Here one can see also the good agreement between

‖ Note that the $t\bar{t} + W$ correction factor obtained with the MC-Template method is applied to $t\bar{t} + W$ MC simulations when the fake/non-prompt lepton efficiency is measured.
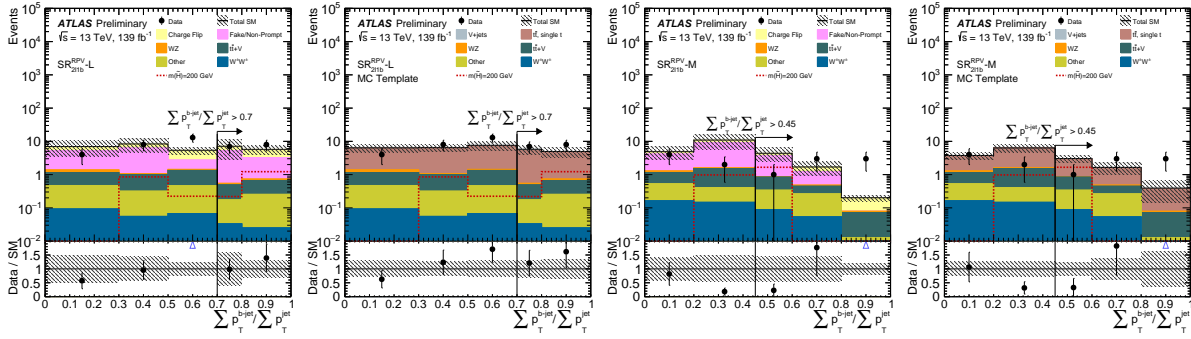
Figure 6. Similar to Figure 5, but for the $\text{SR}_{2\ell1b}^{\text{RPV}}$ signal regions [4]. All the signal region requirements except the one on $\sum p_T^{b-\text{jet}} / \sum p_T^{\text{jet}}$ are applied. Reused with permission from [4].
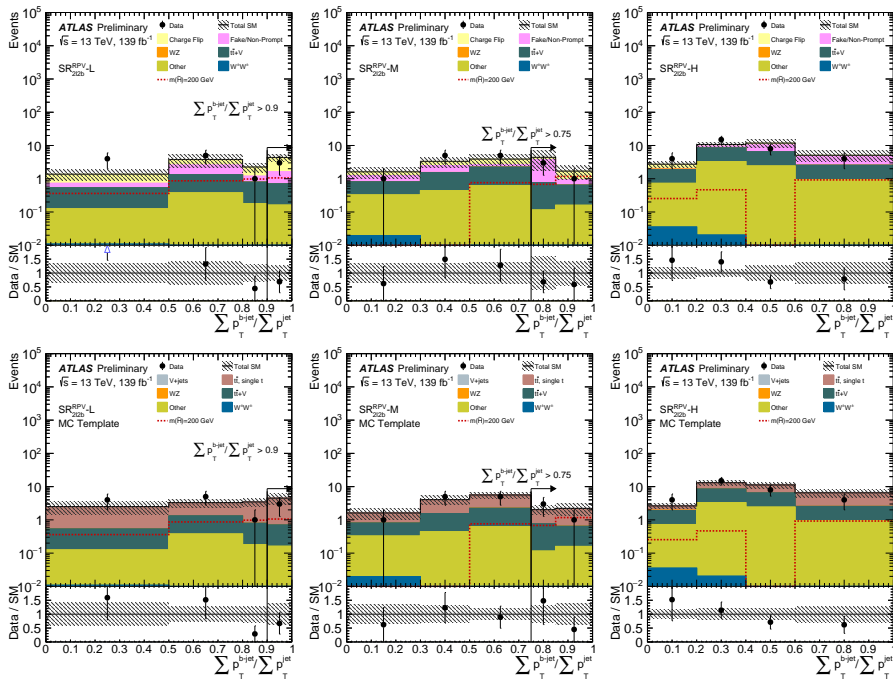

Figure 7. Similar to Figure 6, but for the $\text{SR}_{2\ell2b}^{\text{RPV}}$ signal regions. Top (bottom) the detector background is estimated with the data-based (MC-Template) methods described in the text. Reused with permission from [4].

the detector background estimations with the data-based and MC-Template methods, proving that all methods used for the background estimation are robust. Note also the results in the signal regions, highlighted with an arrow in the figures. In some regions, the MC-Template estimations have lower uncertainties and this will be studied in more detail in future. One improvement will be to combine the fake/non-prompt lepton background estimations obtained with the matrix and MC-Template methods. This will help not only with the reduction of uncertainties, but also with an improved estimation especially in the two same-charge leptons plus 2 or 3 b-jets regions. The latter is hinted at by the data to background ratio seen in Figures 7 and 8.
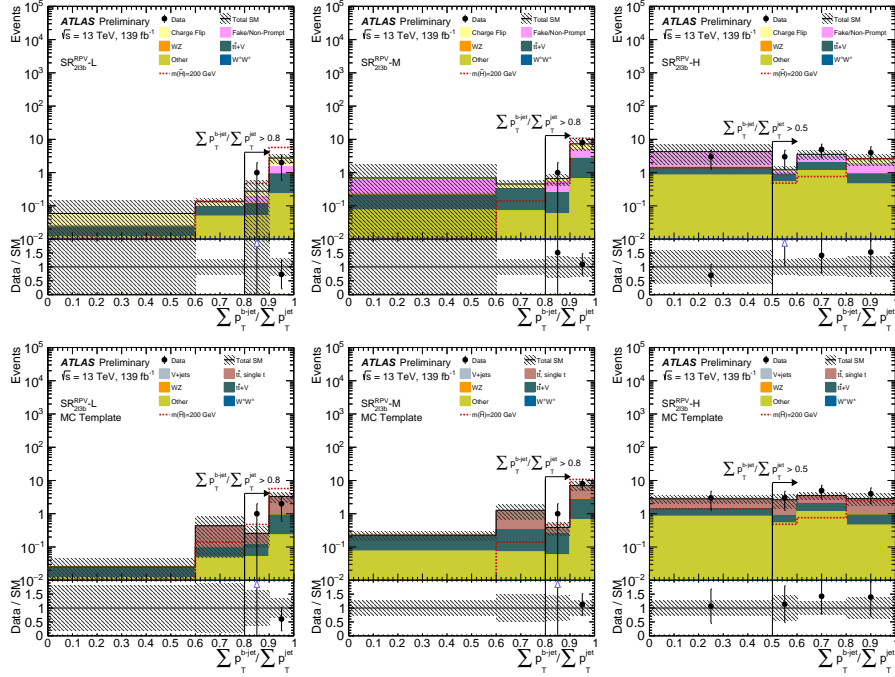
Figure 8. Similar to Figure 7, but for the $SR_{2\ell 3b}^{RPV}$ signal regions. Reused with permission from [4].
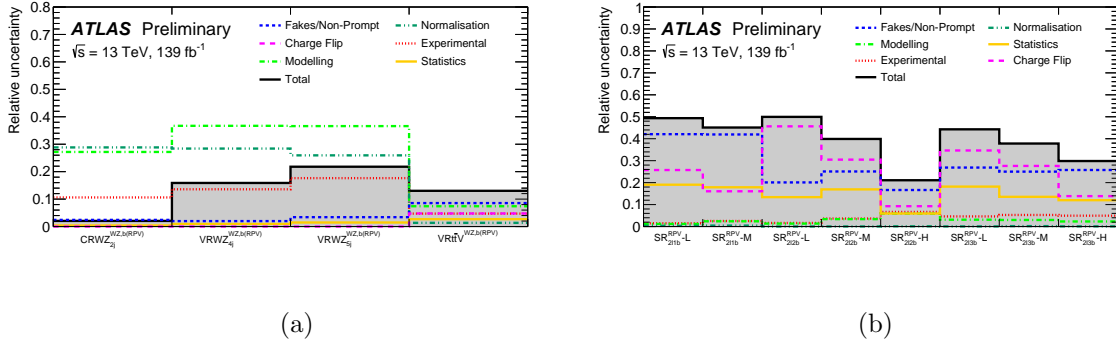

| (a) | (b) |

Figure 9. Contributions of different categories of uncertainties relative to the expected background yields in the (a) control and validation regions, and (b) in the UDD RPV discovery signal regions. The total uncertainty takes into account also the correlation between different nuisance parameters. Reused with permission from [4].

5. Systematic uncertainties

Beside the uncertainties on the detector background estimates, all the experimental and theoretical sources of uncertainties are considered as detailed in [4, 5]. The breakdown of various uncertainty sources in the background prediction in the WZ+jets control region, in the WZ+jets and ttZ/W validation regions, and in the UDD RPV discovery signal regions is shown in Figure 9. In the "Fakes/non-prompt" ("Charge-flip") category, all the systematic uncertainties associated to the fake/non-prompt lepton (electron charge flip) background are considered. In the "Normalisation"
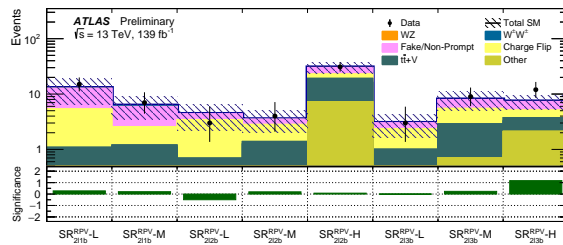
Figure 10. Results in the UDD RPV discovery signal regions. Reused with permission from [4].

category, all the systematic uncertainties associated to WZ+jets normalisation are considered, and in the "Modelling" the systematic uncertainties associated to theoretical modelling for all SM backgrounds are also taken into account, respectively. As the name suggests, the "Statistics" ("Experimental") category has all the statistical (experimental) uncertainties.

In the WZ+jets validation regions, the "Normalisation" and "Modelling" sources have a significant contribution (Figure 9(a)). This is expected, as there is no fit in the validation regions, and the anti-correlations between the theory uncertainties are not accounted for. The analysis sensitivity is not affected too much, as most of the discussed discovery signal regions do not have as a dominant component the WZ+jets background. Nevertheless, in future it would be better to use WZ+jets control regions with simultaneous fits, as done to get the results in e.g the one lepton signal regions discussed in [5] (even if it is more time consuming).

In the selected UDD RPV discovery signal regions, the main uncertainties are the ones associated to the detector backgrounds – not very surprising, as these have the electron charge flip and the fake/non-prompt leptons as the main background. With more data available (possible with the LHC Run 3), these uncertainties can be reduced as the various efficiencies needed as input by the data-based methods can be measured more precisely. Also the matrix method could be expanded to use as input efficiencies measured per fake/non-prompt lepton source.

In the bRPV signal regions, the total uncertainty is around 26% in $SR_{2\ell}^{bRPV}$ with "Fakes/non-prompt" being the main contribution, while in $SR_{3\ell}^{bRPV}$ it is around 16% with the "Fakes/non-prompt" and "Modelling" categories contributing equally.

6. Results

As hinted by Figures 7 and 8 and summarized in Figure 10, there is no significant excess in any of the selected discovery signal regions (nor in the exclusion UDD RPV signal regions [5]). The highest excess is only around 1σ in $SR_{2\ell3b}^{RPV}$, and if the MC-Template method would be used instead of the data-based methods for the detector background, this excess would be even lower. Using these yields, 95% CL model independent upper limits on the number of observed (obs) and expected (exp) BSM events (S^{95}) that may

Table 5. 95% CL upper limits on the visible cross section times efficiency ($\langle\epsilon\sigma\rangle^{95}_{\mathrm{obs}}$). The upper limits on the observed signal events (S^{95}_{obs}), as well on the signal events given the expected number of background events (S^{95}_{exp}) and $\pm1\sigma$ variations of the expected number) are also shown. Reused with permission from [4].

Signal region	$\langle\epsilon\sigma\rangle^{95}_{\mathrm{obs}}$ [fb]	S^{95}_{obs}	S^{95}_{exp}
$\mathrm{SR}^{\mathrm{RPV}}_{2\ell 1b}$-L	0.13	17.5	$15.1^{+4.8}_{-3.7}$
$\mathrm{SR}^{\mathrm{RPV}}_{2\ell 1b}$-M	0.07	10.1	$8.9^{+3.1}_{-1.7}$
$\mathrm{SR}^{\mathrm{RPV}}_{2\ell 2b}$-L	0.04	6.1	$6.2^{+2.4}_{-1.1}$
$\mathrm{SR}^{\mathrm{RPV}}_{2\ell 2b}$-M	0.05	6.8	$6.0^{+2.3}_{-1.2}$
$\mathrm{SR}^{\mathrm{RPV}}_{2\ell 2b}$-H	0.15	20.7	$18.6^{+6.0}_{-4.3}$
$\mathrm{SR}^{\mathrm{RPV}}_{2\ell 3b}$-L	0.04	6.1	$5.7^{+1.9}_{-1.0}$
$\mathrm{SR}^{\mathrm{RPV}}_{2\ell 3b}$-M	0.08	11.5	$9.7^{+3.2}_{-1.8}$
$\mathrm{SR}^{\mathrm{RPV}}_{2\ell 3b}$-H	0.10	13.5	$8.6^{+3.2}_{-2.5}$

contribute to the discovery signal regions are set. Normalizing these to the luminosity of the data sample (139 fb^{-1}), upper limits on the visible BSM cross-section are also obtained: $\sigma = \sigma_{\mathrm{prod}} \times A \times \epsilon = \mathrm{S}^{95}/139$ fb^{-1}. A and ϵ are the corresponding fiducial acceptance and selection efficiency of a BSM signal, in the considered signal region. Selected results are presented in Table 5.

Model dependent exclusion limits are set on the UDD RPV $pp \to \tilde{\chi}^{\pm}_1\tilde{\chi}^0_{1,2}$, $\tilde{\chi}^0_1\tilde{\chi}^0_2$ production cross-section versus the $\tilde{\chi}^0_1$ (LSP) mass, and shown in [4, 5]. Upper limits on the production cross section range from 6.5 pb to 0.18 pb, when the higgsino LSP mass varies from 180 GeV to 400 GeV. LSP masses between 200 GeV and 320 GeV are excluded, thanks to the usage of a neural network in the signal regions defined with a one lepton selection, and to the $m^{\ell j} < 155$ GeV requirement applied for the two same-charge lepton selection. This was seen when performing the optimization of the signal regions in [5], and confirmed in [4] where the discovery signal regions from Table 2 are used to obtain exclusion limits and only the 200 GeV mass point is excluded. The conclusion is that, for discovery, more inclusive (general) signal regions should still be used, as in nature SUSY will not manifest exactly as in a simplified model, while for exclusion limits the usage of machine learning techniques will greatly help.

Figure 11 shows the signal acceptance and efficiency for the two RPV UUD production modes illustrated in Figures 1(c) and 1(d). The results are shown in the one lepton signal regions defined with requirements on the NN discriminant in the 4–8 jet bins, as well as in the two same-charge lepton signal regions [5]. The acceptance gives the number of events passing the selection cuts at generator level, and the efficiency accounts for reconstruction losses like lepton identification, jet energy resolution, jet tagging efficiency, $E^{\mathrm{miss}}_{\mathrm{T}}$ resolution, etc. For the $\tilde{\chi}^{\pm}_1\tilde{\chi}^0_1$ production mode, values of $A \times \epsilon$ are at maximum 15%, while for the $\tilde{\chi}^0_1\tilde{\chi}^0_2$ production mode they are 25% in the one lepton
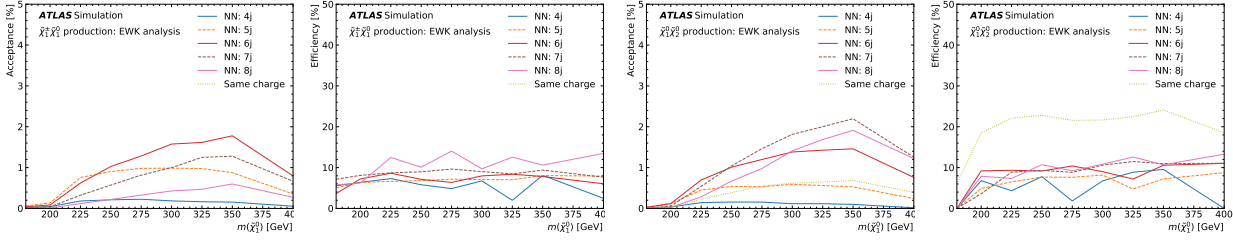
Figure 11. $\tilde{\chi}_1^\pm \tilde{\chi}_1^0$ (left) and $\tilde{\chi}_1^0 \tilde{\chi}_2^0$ (right) UDD RPV signal acceptance and efficiency. Reused with permission from [18].

signal regions, and 15% in the two same-charge lepton signal regions, respectively.

With the bRPV model, the $\tilde{\chi}_1^0$, $\tilde{\chi}_2^0$, $\tilde{\chi}^\pm$ masses are excluded up to 440 GeV, assuming an inclusive higgsino production and allowing all predicted sparticle decay modes. The upper limits on the production cross section range from 0.18 pb to 0.025 pb, when the higgsino masses vary from 200 GeV to 700 GeV. In $\text{SR}_{2\ell}^{\text{bRPV}}$ ($\text{SR}_{3\ell}^{\text{bRPV}}$), values of $A \times \epsilon$ are maximum of 1.5% (5%).

7. Conclusions

Two searches for RPV SUSY through the direct production of pairs of higgsinos have been discussed. The results have been obtained in final states with one, two same-charge or three leptons, using 139 fb^{-1} of LHC data. The analyses have been optimized using two RPV SUSY models, with the R-parity violation obtained though lepton number violation or through baryon number violation. Only prompt decays have been considered. The methods used to estimate the background in two same-charge and three lepton regions have been discussed in detail, as well as their shortcomings. Some ideas for improvements have also been mentioned. Finally, the model dependent and independent upper limits have been presented.

8. Bibliography

[1] S. P. Martin, *A Supersymmetry primer*, Adv. Ser. Direct. High Energy Phys. **18** (1998), 1-98, [arXiv:hep-ph/9709356 [hep-ph]]

[2] L. Evans and P. Bryant, *LHC Machine*, JINST **3** (2008) S08001

[3] H. K. Dreiner, *An Introduction to explicit R-parity violation*, Adv. Ser. Direct. High Energy Phys. **21** (2010), 565-583, [arXiv:hep-ph/9707435 [hep-ph]]

[4] ATLAS Collaboration, *Search for direct production of winos and higgsinos in events with two same-sign or three leptons in pp collision data at 13 TeV with the ATLAS detector*, ATLAS-CONF-2022-057, https://cds.cern.ch/record/2826603

[5] ATLAS Collaboration, *Search for R-parity-violating supersymmetry in a final state containing leptons and many jets with the ATLAS experiment using $\sqrt{s} = 13 TeV$ proton–proton collision data*, Eur. Phys. J. C **81** (2021) no.11, 1023, [arXiv:2106.09609 [hep-ex]]

[6] ATLAS collaboration, twiki.cern.ch/AtlasPublic/SupersymmetryPublicResults

[7] CMS collaboration, twiki.cern.ch/CMSPublic/PhysicsResultsSUSY

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

[9] ATLAS Collaboration, *Performance of the ATLAS Trigger System in 2015*, Eur. Phys. J. C **77** (2017) no.5, 317, [arXiv:1611.09661 [hep-ex]]

[10] ATLAS Collaboration, *Performance of electron and photon triggers in ATLAS during LHC Run 2*, Eur. Phys. J. C **80** (2020) no.1, 47, [arXiv:1909.00761 [hep-ex]]

[11] ATLAS Collaboration, *Performance of the ATLAS muon triggers in Run 2*, JINST **15** (2020) no.09, P09015, [arXiv:2004.13447 [physics.ins-det]]

[12] 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]]

[13] ATLAS Collaboration, *The ATLAS Collaboration Software and Firmware*, ATL-SOFT-PUB-2021-001, https://cds.cern.ch/record/2767187

[14] M. Baak, G. J. Besjes, D. Côte, A. Koutsman, J. Lorenz and D. Short, *HistFitter software framework for statistical data analysis*, Eur. Phys. J. C **75** (2015), 153, [arXiv:1410.1280 [hep-ex]]

[15] ATLAS Collaboration, *Jet reconstruction and performance using particle flow with the ATLAS Detector*, Eur. Phys. J. C **77** (2017) no.7, 466, [arXiv:1703.10485 [hep-ex]]

[16] 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) no.11, 970, [arXiv:1907.05120 [hep-ex]]

[17] ATLAS Collaboration, *Optimisation and performance studies of the ATLAS b-tagging algorithms for the 2017-18 LHC run*, ATL-PHYS-PUB-2017-013, https://cds.cern.ch/record/2273281

[18] ATLAS Collaboration, [5] ATLAS public page, https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PAPERS/SUSY-2019-04/

[19] ATLAS Collaboration, *Measurement of $W^{\pm}Z$ production cross sections and gauge boson polarisation in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, Eur. Phys. J. C **79** (2019) no.6, 535, [arXiv:1902.05759 [hep-ex]]

[20] ATLAS Collaboration, *Electron and photon performance measurements with the ATLAS detector using the 2015–2017 LHC proton-proton collision data*, JINST **14** (2019) no.12, P12006, [arXiv:1908.00005 [hep-ex]]

[21] ATLAS Collaboration, ATLAS public page, https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PAPERS/SUSY-2016-14/

[22] ATLAS Collaboration, *Tools for estimating fake/non-prompt lepton backgrounds with the ATLAS detector at the LHC*, submitted to JINST, [arXiv:2211.16178 [hep-ex]]

[23] ATLAS Collaboration, *Search for squarks and gluinos in final states with same-sign leptons and jets using 139 fb^{-1} of data collected with the ATLAS detector*, JHEP **06** (2020), 046, [arXiv:1909.08457 [hep-ex]]