CMS Physics Analysis Summary

Contact: cms-pag-conveners-susy@cern.ch 2016/08/04

Search for supersymmetry with multileptons in 13 TeV data

The CMS Collaboration

Abstract

A search for new physics is performed using events with multileptons (≥ 3 electrons or muons) in the final state at the CMS detector. Results are based on a sample of proton-proton collisions collected in 2016 at a center-of-mass energy of 13 TeV at the LHC corresponding to an integrated luminosity of 12.9 fb^{-1}. A similar search has been performed with data collected in 2015. With respect to the previous search, minor updates in the object selection have been made and signal regions have been partially redefined to improve the sensitivity of the analysis with the larger integraded lumiosity available. Search regions are defined by the number of b-tagged jets, missing transverse energy, hadronic transverse energy, and the invariant mass of opposite-sign, same-flavor dilepton pairs in the events. No significant excess above the standard model background expectation is observed.

1 Introduction

The multilepton (three or more leptons) final state is a strongly motivated place to search for new physics. The standard model (SM) processes that produce this final state are characterized by having multiple bosons, which are well-understood both theoretically and experimentally. Many different types of beyond the standard model (BSM) theories can produce multilepton events, with a wide array of unique signatures. This analysis is designed to have broad sensitivity to these possibilties by examining the event yields as a function of several kinematic quantities.

This physics analysis summary describes the methods and results of a search for new physics in final states with three or more leptons in proton-proton collisions at a center-of-mass energy of 13 TeV at the LHC using the CMS detector with 12.9 fb^{-1} of luminosity. A similar search has been performed with data collected in 2015 [1]. With respect to the previous search, minor updates in the object selection have been made and signal regions have been partially redefined to improve the sensitivity of the analysis with the larger integraded lumiosity available. Kinematic regions are defined to optimize the search for events that contain large missing transverse energy and high jet multiplicities. For these final states, the expected irreducible backgrounds come from diboson production ($\text{W}^{\pm}\text{Z}$ and ZZ) or rare SM processes (including $\text{t}\bar{\text{t}}\text{W}^{\pm}$, $\text{t}\bar{\text{t}}\text{Z}$, $\text{t}\bar{\text{t}}\text{H}$, etc.). These backgrounds are modeled using simulation samples that have appropriate corrections applied to match the behavior of reconstructed objects in data. Reducible backgrounds are processes that produce one or more misidentified or nonprompt leptons, i.e. leptons originating from jet or meson decays, which pass all reconstruction, identification, and isolation criteria. Data driven techniques to measure probabilities of observing misidentified or nonprompt leptons are used to predict the reducible background contributions.

Results of the analysis will be interpreted in the context of supersymmetric (SUSY) models that feature gluino pair production with mass spectra that produce final state leptons through the decays of vector bosons. In addition to multiple leptons, these models produce multiple jets and missing transverse energy.

Similar searches have been carried out by CMS using the 8 TeV dataset [2, 3], as well as by ATLAS [4]. With 2.3 fb^{-1} of data collected at 13 TeV with the CMS detector in 2015, gluinos with masses of less than 1125 GeV could be excluded in a model with gluino pair production where the gluino decays to two top quarks and a neutralino [1].

2 Event selection and Monte Carlo simulation

Events used in this analysis are selected by high-level triggers (HLT) that target di- and multilepton events for offline study. One set of triggers requires that the two leptons meet loose isolation criteria and that the leading lepton has $p_{\text{T}} > 23\,\text{GeV}$ and the sub-leading lepton has $p_{\text{T}} > 8(12)\,\text{GeV}$ in the case of muons (electrons). The second set of triggers places no requirements on the isolation, has a lower p_{T} threshold for the two leptons ($p_{\text{T}} > 8\,\text{GeV}$), and also requires that the hadronic activity reconstructed in the HLT be greater than 300 GeV.

The selection requires that there be at least three well-identified leptons in the event and that opposite-sign same-flavor pairs have an invariant mass greater than 12 GeV to reject Drell-Yan and quarkonia processes. The leptons must pass p_{T} thresholds which dependent on the lepton flavor and the hadronic activity in the event. For events with low hadronic activity, the leading lepton must satisfy $p_{\text{T}} > 25\,\text{GeV}$ and subleading muons (electrons) must satisfy $p_{\text{T}} > 10(15)\,\text{GeV}$ respectively. In events with high hadronic activity, the thresholds are relaxed

to 10 (15) GeV for leading and subleading muon (electron). The third lepton must have $p_T >$ 10 GeV all cases. For this offline selection, the trigger efficiency is larger than 97% if at least two leptons have a p_T greater than the p_T required for the leading lepton. A correction for trigger inefficiencies measured in data is applied to simulation if this requirement is not met.

Muon candidates are reconstructed combining the information from both the silicon tracker and the muon spectrometer in a global fit [5]. An identification is performed using the quality of the geometrical matching between the tracker and the muon system measurements. To ensure the candidates are within the fiducial volume of the detector, we require that the candidate pseudorapidity $|\eta| < 2.4$.

Electron candidates are reconstructed using tracking and electromagnetic calorimeter information by combining ECAL superclusters and gaussian sum filter (GSF) tracks [6]. The electron identification is performed using a multivariate discriminant built with shower-shape variables, track-cluster matching variables, and track quality variables. The working point for the selection is chosen to maintain approximately 90% efficiency for accepting electrons produced in the decays of SM bosons and to also efficiently reject candidates originating from jets. To reject electrons originating from photon conversion, electrons are required to have all possible hits in the innermost tracker layers and to not be compatible with any conversion-like secondary vertices. Electrons must have $|\eta| < 2.5$.

Both muon and electron candidates are required to have transverse impact parameter less than 0.5 mm from the event's primary vertex and a longitudinal impact parameter less than 1 mm. In addition, a cut on the 3D impact parameter significance is applied. This variable is the value of impact parameter divided by its uncertainty. This variable should be less than 4 for both electrons and muons. We find that the rejection of nonprompt leptons is much higher using this cut than cuts on the impact parameter that have similar prompt lepton acceptance.

The lepton isolation is constructed using three different variables. The mini isolation, I_{mini}, is the ratio of the amount of energy in a cone with a p_T-dependent radius, $R = \frac{10}{\min(\max(p_T(\ell),50),200)}$. Requiring I_{mini} below a given threshold ensures that the lepton is locally isolated, even in boosted topologies.

The second variable is the ratio of the lepton p_T and p_T of the jet matched to the lepton: $p_T^{\text{ratio}} = \frac{p_T(\ell)}{p_T(\text{jet})}$. This jet is matched geometrically to the lepton and residing within $\Delta R < 0.4$ from it. In most of the cases, this is the jet containing the lepton. If no jet within $\Delta R < 0.4$ is found then lepton itself is used to compute $p_T^{\text{ratio}} = 1$. The use of p_T^{ratio} is a simple way to identify non-prompt low-p_T leptons originating from low-p_T b-quarks which decay with larger aperture than the one used in mini isolation.

The last variable used is the p_T^{rel} variable, which is calculated by subtracting the lepton from the momentum vector of the geometrically matched jet described above and then finding the component of the lepton momentum which is transverse to this new vector. This variable allows us to recover leptons from accidental overlap with jets in boosted topologies.

Using those three variables, a lepton is considered isolated if the following condition is fulfilled:

$$I_{\text{mini}} < I_1 \&\& (p_T^{\text{ratio}} > I_2 || p_T^{\text{rel}} > I_3) \quad (1)$$

The $I_i, i = 1, 2, 3$ values depend of the flavor of the lepton: as the probability to misidentify a jet is higher for the electrons, tighter isolation values are used. The logic behind this isolation is that a lepton should be locally isolated (I_{mini}), carry the major part of the energy of the corresponding jet (p_T^{ratio}), or if not, then this lepton is accepted if its overlap with a jet is accidental

($p_{\mathrm{T}}^{\mathrm{rel}}$). The loose lepton isolation is significantly relaxed: $I_{\mathrm{mini}} < 0.4$ while the other requirements are dropped. For muons (electrons), the tight selection requirements are $I_1 = 0.16(0.12)$, $I_2 = 0.69(0.76)$, and $I_3 = 6.0(7.2)$.

Jets are reconstructed from particle flow candidates [7], clustered with the anti-k_{T} algorithm and with a distance parameter of $R = 0.4$. Only jets above a transverse momentum $p_{\mathrm{T}} > 30\,\mathrm{GeV}$ and within the tracker acceptance $|\eta| < 2.4$ are considered. Additional criteria are applied to reject events containing noise and mismeasured jets. To avoid double counting due to jets matched geometrically with a lepton, the jet the closest matched to a lepton is not considered as a jet in the event, if the jet is within a cone of $\Delta R < 0.4$. From those selected jets, the key variable H_{T} is defined by $H_{\mathrm{T}} = \sum_{jets} p_{\mathrm{T}}$, where the sum runs over all jets which satisfy the above mentioned criteria.

A combined secondary vertex algorithm [8, 9] is used to assess the likelihood that a jet originates from a bottom quark. Jets in this analysis are considered b-tagged if they pass the algorithm's medium working point, which has a tagging efficiency of 70% and a mistag rate of 1%. With respect to [1] the p_{T} threshold for b-jets has been relaxed from 30 to 25 GeV. B-tagged jets with $p_{\mathrm{T}} < 30\,\mathrm{GeV}$ are not considered for the calculation of H_{T}.

The missing transverse energy $E_{\mathrm{T}}^{\mathrm{miss}}$ is defined as the magnitude of the negative vector sum of all particle flow candidates reconstructed in an event. The jet energy corrections are propagated to the $E_{\mathrm{T}}^{\mathrm{miss}}$ following the procedure described in [10].

In order to estimate the contribution from SM processes yielding prompt leptons in the signal regions and to calculate the efficiency for new physics models, Monte Carlo (MC) simulations are used. The MADGRAPH5 2.2.2 [11] program is used for event generation at leading order (LO) or next-to-leading order (NLO) in perturbative QCD for all SM background processes, except diboson production. For the latter, the POWHEG v2 [12] generator is empoyed. Parton showering and hadronization are simulated using the PYTHIA 8.205 generator [13] with the CUETP8M1 tune [14]. The CMS detector response is determined using a GEANT4-based model [15].

Signal events for interpretation are generated with the MADGRAPH5_AMC@NLO program [11] at LO precision, allowing up to two additional partons in the matrix element calculations. The SUSY particle decays, parton showering, and hadronization are simulated with PYTHIA 8.205 [13]. The detector response for signal events is simulated using a CMS fast-simulation package [16] that is validated against the GEANT4-based model. All simulated events are processed with the same reconstruction procedure as data. Corrections are applied to match the distribution of interactions per bunch crossing in simulation and data. Cross sections for SUSY signal processes, calculated at NLO with next-to-leading-log (NLL) resummation, are taken from the LHC SUSY Cross Sections Working Group [17–21].

3 Search strategy

The goal of this analysis is to search for possible excesses over the expected yields from SM processes in events with three or more leptons. The search is focused on the strongly produced SUSY particles which benefit the most from the rise in the production cross section at 13 TeV. Examples for SUSY processes which can give rise to multilepton final states are shown in Fig. 1. Both models are simplified model spectra (SMS). In these models, SUSY particles that are not directly included in the diagrams are forced to be too heavy to be accessible at the LHC. Therefore, the free parameters in these models are usually the mass of the produced particle – here

gluinos – and the mass of the lightest super partner (LSP) which is a neutralino in the example models.

A typical process within SUSY includes the one known as T1tttt: gluino-pair production where each gluino decays to a $t\bar{t}$ pair and an LSP (Fig. 1a). Another model containing gluino-pair production where each gluino decays to a pair of quarks and a neutralino or chargino and that neutralino or chargino then decays to a W or Z boson and an LSP, depending on the charge, is called T5qqqqVV (Fig. 1b).

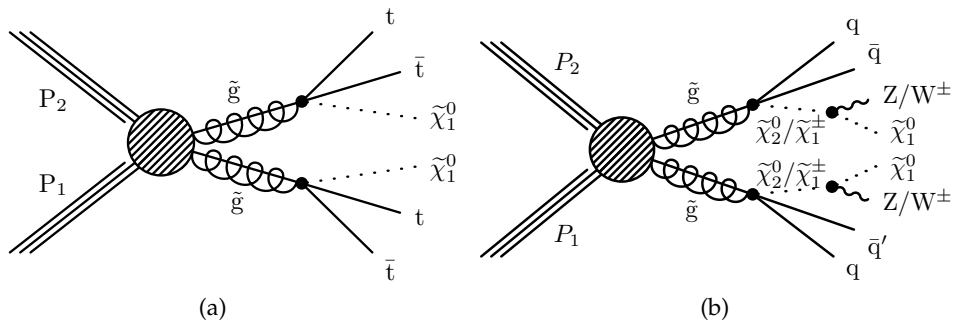


Figure 1: Diagrams for gluino pair production models which can produce multilepton events: T1tttt (a) and T5qqqqVV (b).

For the definition of the signal regions we use several event variables: number of b-jets ($N_{\text{b jets}}$), the hadronic activity (H_{T}), the missing transverse energy ($E_{\text{T}}^{\text{miss}}$), and the invariant mass of opposite-sign, same-flavor pairs in the event.

The separation in b-jet multiplicities ensures the maximization of the signal-to-background ratios for events from different signal models. For example, the T1tttt model features several b-jets, which would be categorized into signal regions which are almost free of WZ background owing to the b-jet requirement. Including the 0 b-tag signal regions keeps the analysis sensitive to signatures without b-jets like the T5qqqqWZ model. Additionally, a categorization in H_{T} and $E_{\text{T}}^{\text{miss}}$ is useful to distinguish between compressed and non-compressed SUSY spectra, i.e. models with small or large mass differences between the SUSY particles in the decay chain.

A baseline selection is applied to the dataset to select events of interest: three or more electrons or muons fulfilling the flavor and H_{T} dependent p_{T} thresholds outlined in Section 2, $m_{\ell\ell} \geq 12\,\text{GeV}$, $N_{\text{jets}} \geq 2$, and $E_{\text{T}}^{\text{miss}} \geq 50\,\text{GeV}$. Two different regions are defined, based on whether an event contains an opposite-sign, same-flavor lepton pair with an invariant mass within a 15 GeV window around the Z mass or not. If such a lepton pair is found the event is categorized as on-Z and else as off-Z. In order to reject Drell-Yan events in on-Z signal regions with low $E_{\text{T}}^{\text{miss}}$, low H_{T}, the minimum $E_{\text{T}}^{\text{miss}}$ requirement is raised from 50 to 70 GeV. Tables 1 and 2 show the definition of the subdivision of the baseline selection into 15 off-Z and 17 on-Z signal regions (SR) respectively. A set of four SR with low or medium H_{T} and low or medium $E_{\text{T}}^{\text{miss}}$ have been defined for each of the b-tag multiplicities 0, 1, and 2. Motivated by the low expected yield of events with 3 or more b-jets, one inclusive SR with $E_{\text{T}}^{\text{miss}} < 300$ and $H_{\text{T}} < 600$ has been defined for high b-tag multiplicities ≥ 3 (SR 13). Two additional SR with ultra high H_{T} (SR 14) and ultra high $E_{\text{T}}^{\text{miss}}$ (SR 15) respectively have been defined since various non-compressed SUSY model can yield events with very high $E_{\text{T}}^{\text{miss}}$ or H_{T}. For the on-Z region, SR 14 and 15 are split in two bins of H_{T} and $E_{\text{T}}^{\text{miss}}$, referred to as SR 14a and 14b and SR 15a and 15b respectively. The split of on-Z SR 14 and SR 15 is motivated by the larger expected background yields in these regions compared to the corresponding off-Z regions. These ultra-high $E_{\text{T}}^{\text{miss}}$ and H_{T} regions

are inclusive in the number of b-jets and priority is given to ultra high $E_{\mathrm{T}}^{\mathrm{miss}}$ region, meaning that every selected event with $E_{\mathrm{T}}^{\mathrm{miss}} \geq 300$ is categorized in this region, while the ultra high H_{T} region 14 is populated with events with $E_{\mathrm{T}}^{\mathrm{miss}} < 300\,\mathrm{GeV}$ and $H_{\mathrm{T}} \geq 600\,\mathrm{GeV}$.

Table 1: Multilepton off-Z signal region definition

N_{jets}	$N_{\mathrm{b\,jets}}$	$E_{\mathrm{T}}^{\mathrm{miss}}$ (GeV)	$60\,\mathrm{GeV} \leq H_{\mathrm{T}} < 400\,\mathrm{GeV}$	$400\,\mathrm{GeV} \leq H_{\mathrm{T}} < 600\,\mathrm{GeV}$	$H_{\mathrm{T}} \geq 600\,\mathrm{GeV}$
≥ 2	0	$50-150$	SR1	SR3	SR14
		$150-300$	SR2	SR4	
	1	$50-150$	SR5	SR7	
		$150-300$	SR6	SR8	
	2	$50-150$	SR9	SR11	
		$150-300$	SR10	SR12	
	≥ 3	$50-300$	SR13		
	inclusive	≥ 300	SR15		

Table 2: Multilepton on-Z signal region definition. $E_{\mathrm{T}}^{\mathrm{miss}}$ cuts labeled with * refer to the low H_{T} regions only. SR14a requires $50\,\mathrm{GeV} < E_{\mathrm{T}}^{\mathrm{miss}} < 150\,\mathrm{GeV}$ and SR14b $150\,\mathrm{GeV} < E_{\mathrm{T}}^{\mathrm{miss}} < 300\,\mathrm{GeV}$.

N_{jets}	$N_{\mathrm{b\,jets}}$	$E_{\mathrm{T}}^{\mathrm{miss}}$ (GeV)	$60\,\mathrm{GeV} \leq H_{\mathrm{T}} < 400\,\mathrm{GeV}$	$400\,\mathrm{GeV} \leq H_{\mathrm{T}} < 600\,\mathrm{GeV}$	$H_{\mathrm{T}} \geq 600\,\mathrm{GeV}$	
≥ 2	0	$50(70^*)-150$	SR1	SR3	SR14a	SR14b
		$150-300$	SR2	SR4		
	1	$50(70^*)-150$	SR5	SR7		
		$150-300$	SR6	SR8		
	2	$50-150$	SR9	SR11		
		$150-300$	SR10	SR12		
	≥ 3	$50-300$	SR13			
	inclusive	≥ 300	SR15a		SR15b	

4 Background Estimation

Backgrounds for the multi-lepton final state can be divided in three categories:

- **Nonprompt or misidentified leptons** are leptons from heavy-flavor decays, misidentified hadrons, muons from light-meson decays in flight, or electrons from unidentified photon conversions. For this analysis $\mathrm{t\bar{t}}$ events can enter the signal regions if nonprompt leptons are present in addition to the prompt leptons from the W decays. $\mathrm{t\bar{t}}$ events are characterized by low H_{T} and low $E_{\mathrm{T}}^{\mathrm{miss}}$ and therefore predominately populate signal regions 1 and 5, with 0 and 1 b-tagged jet respectively. Apart from $\mathrm{t\bar{t}}$, Drell-Yan events can enter the baseline selection, however they are largely supressed by the $E_{\mathrm{T}}^{\mathrm{miss}} > 50\,\mathrm{GeV}$ cut and additional rejection is achieved by increasing the $E_{\mathrm{T}}^{\mathrm{miss}}$ cut to 70 GeV for low H_{T}, low $E_{\mathrm{T}}^{\mathrm{miss}}$ on-Z regions. Processes which yield only one prompt lepton in addition to non-prompt ones like W+jets and various single top channels are effectively suppressed by the three lepton requirement because of the low probablity that two non-prompt leptons pass the tight identification and isolation requirements.
- **Diboson production** can yield multilepton final states with up to three prompt leptons for WZ and up to four prompt leptons for ZZ production, redering these processes irreducible background for this analysis. Especially in signal regions without b-tagged jets, WZ production has a sizable contribution. To estimate this background, its yield as obtained from simulation is scaled using a scale factor measured in a dedicated control region enriched in WZ events.

- **Other rare SM processes** that can yield three or more leptons are $t\bar{t}W$, $t\bar{t}Z$, and triboson production VVV where V= W, Z. We also include the contribution from the SM Higgs boson produced in association with a vector boson or a pair of top quarks in this category of backgrounds, as well as processes that produce additional leptons from internal conversions, which are events that contain a virtual photon that decays to leptons. The internal conversion background components, X+γ, are heavily suppressed by the $E_{\mathrm{T}}^{\mathrm{miss}} > 50\,\mathrm{GeV}$ and $N_{\mathrm{jets}} \geq 2$ cuts. Those rare backgrounds are obtained from simulation and appropriate systematic uncertainties are assigned.

The background contribution from nonprompt and misidentified leptons is estimated using the tight-to-loose ratio method. In this method, the yield in an application region, populated by events that contain at least one lepton which fails the full set of tight identification and isolation requirements but passes the loose requirements, is weighted by $f/(1-f)$, where the tight-to-loose ratio f is the probability that a loosely identified lepton also passes the full set of requirements. This ratio is measured as a function of lepton p_{T} and η in a control sample of QCD multijet events that is enriched in nonprompt leptons (measurement region). Exactly one lepton passing the loose object selection is required in the event. Additionally, one recoiling jet with $\Delta R(\mathrm{jet}, \ell) > 1.0$ and $p_{\mathrm{T}} > 30\,\mathrm{GeV}$ and low $E_{\mathrm{T}}^{\mathrm{miss}}$ and M_{T}, both $< 20\,\mathrm{GeV}$ are required to suppress events with leptons from W and Z decays, where M_{T} is the transverse mass of the lepton and the $E_{\mathrm{T}}^{\mathrm{miss}}$ vector. The remaining contribution from these electroweak processes within the measurement region is subtracted using estimates from MC simulations.

The main advancement of the method with respect to run 1 searches is the reduction of the dependence of the tight-to-loose ratio on the flavor composition of the jets from which the nonprompt leptons originate. This has been achieved by parameterizing the ratio as a function of a variable that correlates stronger with the mother parton p_{T} than with the lepton p_{T}. This variable is calculated by correcting the lepton p_{T} as a function of the energy in the isolation cone around it. This definition leaves the p_{T} of the leptons passing the isolation cut unchanged and modifies the p_{T} of those failing the cut so that it is a better proxy for the mother parton p_{T} and results in a flatter ratio as a function of the mother parton p_{T}. The cone correction significantly improves the results of the method when applying it in simulation. The flavor dependence, which is much more important for the case of electrons, is also reduced by adjusting the loose object selection to obtain similar ratios for nonprompt electrons that originate from both light- and heavy-flavor jets.

The tight-to-loose ratio method for estimating the nonprompt background is validated both in a Monte Carlo closure test and in a data control region exclusive to our baseline selection with minimal signal contamination. This region is defined by having three leptons pass nominal identification, isolation cuts and p_{T} requirement, one or two jets, $20 < E_{\mathrm{T}}^{\mathrm{miss}} < 50$ GeV, and no on-Z dilepton pair. With these cuts a purity in $t\bar{t}$ of 76% can be achieved. We find agreement of the order of 20 – 30% between the predicted and observed yields in this control region in the main event observables.

The WZ process is one of the main backgrounds in the regions with 0 b-tags. The estimates for this process are taken from MC simulation, but the normalization is obtained from a control region that is highly enriched in this process: three leptons passing the nominal identification and isolation cuts, two leptons form an opposite sign, same flavor pair with $|m_{\ell\ell} - m_{\mathrm{Z}}| < 15$ GeV, the number of jets is zero or one, the number of b-tagged jets is zero, $30\,\mathrm{GeV} < E_{\mathrm{T}}^{\mathrm{miss}} < 100\,\mathrm{GeV}$, and the transverse mass of the third lepton (not in the pair forming the Z) is required to be at least 50 GeV to suppress contamination from Drell-Yan process. The purity of WZ in the selected region is 80%. In this control region, we measure a scale factor for the WZ simulation

needed to match the background prediction and the observation in data. For 12.9 fb^{-1} very good agreement between data and simulation has been found and a scale factor of 0.98 ± 0.13 has been measured. The uncertainty on the scale factor measurement is used as additional uncertainties on the WZ prediction in the SRs.

5 Systematic Uncertainties

The different uncertainties are categorized as experimental, as those related to the jet-energy scale or the b-tagging efficiency; theoretical, such as the uncertainties on the considered cross sections; statistical, due to the limited size of the Monte Carlo samples; and uncertainties on the applied data-driven methods. These uncertainties and their effect on the predicted yields are described below and summarized in Table 3.

One of the major experimental sources of uncertainty is the knowledge of the jet energy scale (JES), a correction aplied to match jet energies measured in data and simulation. This uncertainty affects all simulated background and signal events. For the dataset used in this analysis, the uncertainties on the jet energy scale vary from 1% to 8%, depending of the transverse momentum and pseudorapidity of the jet. The impact of this uncertainties is assessed by shifting the jet energy correction factors for each jet up and down by $\pm 1\sigma$ and recalculating all kinematic quantities. The JES uncertainties are propagated to the missing transverse energy and all variables derived from jets (number of jets, H_{T}, number of b-jets) used in this analysis. The propagation of the variation of the JES results in a variation of 1–10% in the event yields.

A similar approach is used for the uncertainties associated with the corrections for the b-tagging efficiencies for light and bottom flavor jets, which are parametrized as a function of p_{T} and η. The variation of the scale factor correcting for differences between data and simulation is at maximum of the order of 10% per jet, and leads to an overall effect in the range of 1–20% depending on the signal region and on the topology of the event.

Lepton identification and isolation scale factors have been measured as function of lepton p_{T}

source	effect on yield
luminosity	6.2%
jet energy scale	1 – 10%
b-tag efficiency	1 – 20%
pileup	1 – 5%
lepton efficiencies	9%
HLT efficiencies	3%
lepton eff. FastSim	6%
appl. region stat.	15 – 100%
non-prompt extrapol.	30%
EWK subtraction	5%
WZ CR normalization	15%
WZ CR extrapolation	10 – 30%
MonteCarlo stat.	1 – 100%
QCD scales cross-section (ttV)	11 – 13%
QCD scales acceptance (ttV)	3 - 18%
PDFs (ttV)	2 – 3%
NLO/LO (ttV)	1 – 70%
other rare bkgs.	50%

Table 3: Systematic uncertainties and their effect on the event yields.

and η. They are applied to correct for residual differences in lepton selection efficiencies between data and simulation. The corresponding uncertainties are estimated to be about 3% per lepton for both flavors. Assuming correlation between the corrections of the different lepton a flat uncertainty of 9% is taken into account. The uncertainty related to the HLT trigger efficiency correction for the simulated backgrounds amounts to $\pm$3%.

All these uncertainties, related to corrections of the simulation (JES corrections, b-tagging efficiency scale-factors, lepton identification scale-factors) have been estimated also for the fast-simulation used for the signal samples scans and are propagated to the signal event yields following the same procedures as decribed above.

Theoretical uncertainties include the uncertainty on the QCD renormalization (μ_R) and factorization scales (μ_F), and on the knowledge of the parton density functions (PDF). The uncertainties are considered for electroweak processes, namely $t\bar{t}H$, $t\bar{t}Z$, and $t\bar{t}H$, which are dominant backgrounds in some tail signal regions. Both the changes in acceptance and cross sections related to those effects are taken into account.

For the study of the renormalization and factorization uncertainties, fluctuations up and down by a factor of two with respect to the nominal values of μ_F and μ_R are considered. The maximum difference in the yields with respect to the nominal case is observed when both scales are varied up and down simultaneously. The effect on the overall cross section is found to be $\sim$13% for $t\bar{t}W$ and $\sim$11% for $t\bar{t}Z$. An additional, uncorrelated uncertainty on the acceptance corresponding to different signal regions is included. This is found to be between 3% and 18% depending on the search region and process.

The uncertainty related to the PDF is estimated from the the 100 NNPDF 3.0 replicas, computing the deviation with respect to the nominal yields for each of them, and for each signal region (the cross section and acceptance effect are considered together) [22]. The root mean square of the variations is taken as the value of the systematic uncertainty. Since no significant differences between signal regions have been found, a flat uncertainty of 3% (2%) is considered for $t\bar{t}W$ ($t\bar{t}Z$). This value also includes the effect on $\alpha_S(M_Z)$, which is added in quadrature. An additional uncertainty to take into account differences between ttV samples simulated with next-to-leading order accurancy and the leading order samples used in this analysis has been added. Depending on the H_T of the signal region and the background process this uncertainty ranges from 1–70%.

For the $t\bar{t}H$ process, the same Q^2 and PDF related uncertainties as estimated for $t\bar{t}Z$ are considered. A conservative uncertaintiy of 50% is assigned to the remaining rare processes.

For signal samples additional uncertainties for initial state radiation and E_T^{miss} are taken into account, following the recommendation of the CMS SUSY physics analysis group.

The limited size of the generated Monte Carlo samples represents an additional source of uncertainty. For the backgrounds that are estimated from simulation, like $t\bar{t}W$, $t\bar{t}Z$ and $t\bar{t}H$, as well as for all the signal processes, this uncertainty is computed from the number of Monte-Carlo events entering the signal regions.

For the nonprompt and misidentified lepton background, several systematic uncertainties are considered. The statistical uncertainty from the application region which is used to estimate this background contribution ranges from 15% to 100%. The regions where these uncertainties are large are generally regions where the overall contribution of this background is small. In the case where no events are observed in the application region, the upper limit of the background expectation is found by applying the most probable value of the tight-to-loose ratio to a Poisson

fluctuation.

The systematic uncertainty related to the extrapolation from the control regions to the signal regions for the nonprompt lepton background is estimated to be 30%. This magnitude has been extracted from the level of closure achieved in tests which are performed with Monte Carlo samples yielding non-prompt leptons to validate the data-driven background prediction described in Section 4.

The uncertainty associated to the electroweak subtraction in the tight-to-loose ratio computation is propagated along the full analysis process, by replacing the nominal tight-to-loose ratio with another value obtained when the scale factor applied to the electroweak processes in the measurement region is varied up and down by 100%. The overall effect on the nonprompt background yield is of the order of 5%.

Two sources of uncertainty are considered for the estimation of the WZ background. The uncertainty of the normalization in the control region is estimated to be 15%. An additional, signal region dependent uncertainty, ranging from 10 – 30%, is taken into account for the extrapolation from the control region to the signal regions.

6 Results

A comparison of expected background events and data for 12.9 fb^{-1} in distributions of the three event observables used for signal region categorization – H_T, E_T^miss, and $N_\text{b jets}$ – as well as for the lepton p_T spectra, the lepton flavor composition, and the lepton multiplicity in the event is shown in Fig. 2 (Fig. 3), using all the events satisfying the off-Z (on-Z) search region selection criteria. The hatched band represents the statistical and systematic uncertainty in each bin. Fig. 4 graphically presents a summary of predicted background and observed event yields in the individual signal regions. The same data is also presented in Tables 4 and 5 for the off-Z and on-Z regions, respectively.

The number of events observed in data are found to be consistent with predicted background yields in all 32 signal regions. No significant deviation have been found.

Observing consistent numbers for data and background prediction, the search is interpreted by setting limits on gluino and neutralino masses using the T1tttt and the T5qqqqWZ simplified models. For each mass point, the observation, background predictions, and expected signal yields from all on-Z and off-Z search regions are combined to extract a cross section that can be excluded at a 95% confidence level (CL) using the LHC-type CL_s method. Log-normal nuisance parameters are used to describe the systematic uncertainties listed in Section 5. The results of the limit setting procedure are shown in Fig. 5a for the T1tttt model and in Fig. 5b for the T5qqqqWZ model. For the latter model the WW and ZZ final states have been filtered out and the gluino-gluino cross section has been scaled down accordingly. For both exclusion plots the gluino pair-production cross section is calculated at NLO-NLL (next-to-leading-logarithm) accuracy and assumes that other SUSY particles are decoupled (i.e. very massive).

For both signal models the most sensitive signal regions have been identified for each one compressed and one non-compressed mass point close to the exclusion limit. For the T1tttt model, Table 6 shows the expected number of background and signal events, the observed yields and the expected and observed exclusion limits obtained with only one signal region for an uncompressed mass spectrum with a gluino mass of 1200 GeV and an LSP mass of 100 GeV and for a compressed spectrum with the same gluino mass but a higher LSP mass of 700 GeV. For both mass points, the three most sensitive signal regions are listed, ordered by

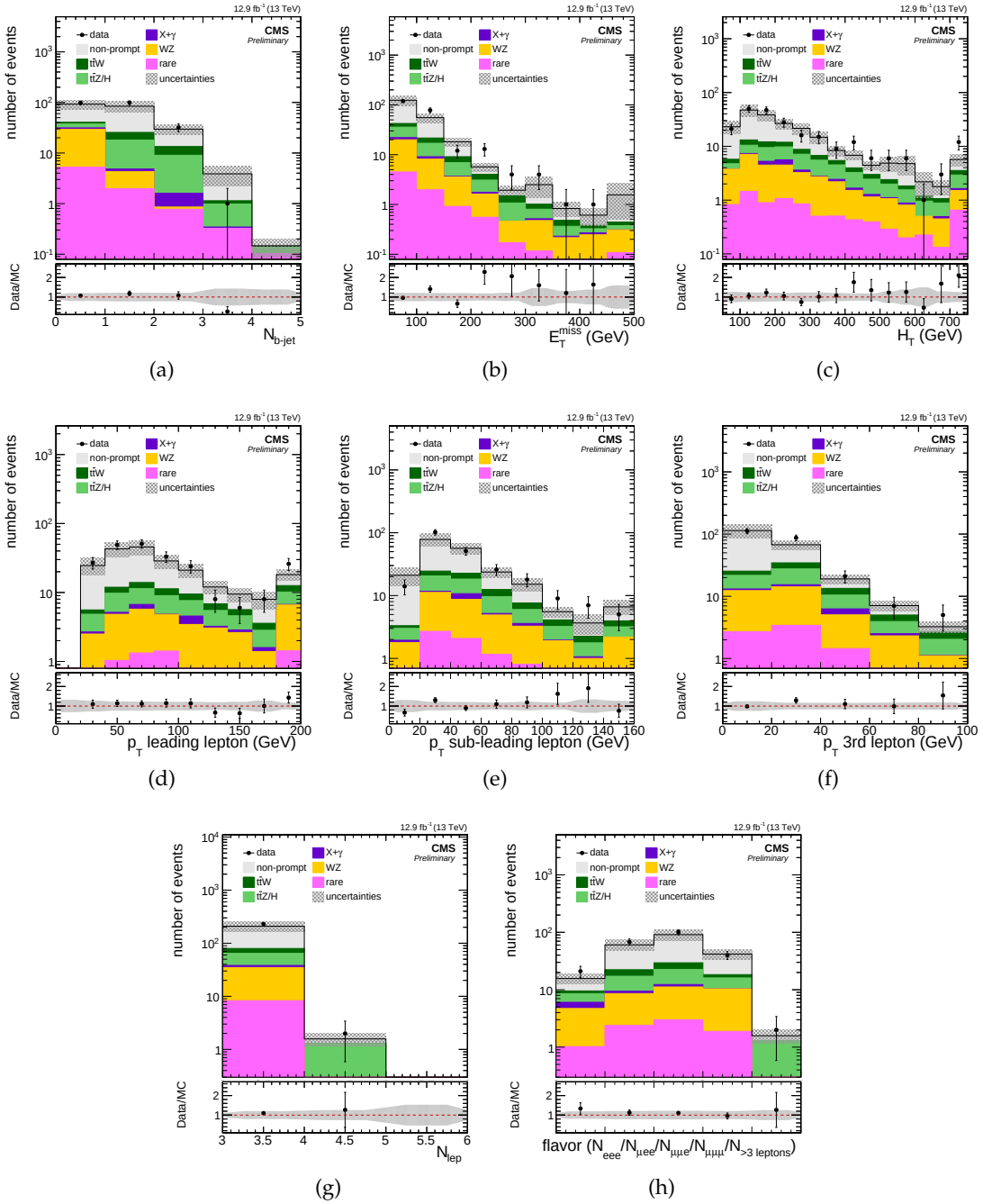


Figure 2: Background prediction and observation in key observables of the off-Z baseline selection for 12.9 fb^{-1}. Good agreement between data and expected background yields is observed in the event variables used for signal region categorization – the b-jet multiplicity, $E_{\mathrm{T}}^{\mathrm{miss}}$, and H_{T}. Additionally, the distibutions of the lepton p_{T} spectra, the flavor composition of the leptons, and the lepton multiplicity in the event are shown. The hatched area represents the statistical and systemtic uncertainties on the prediction.

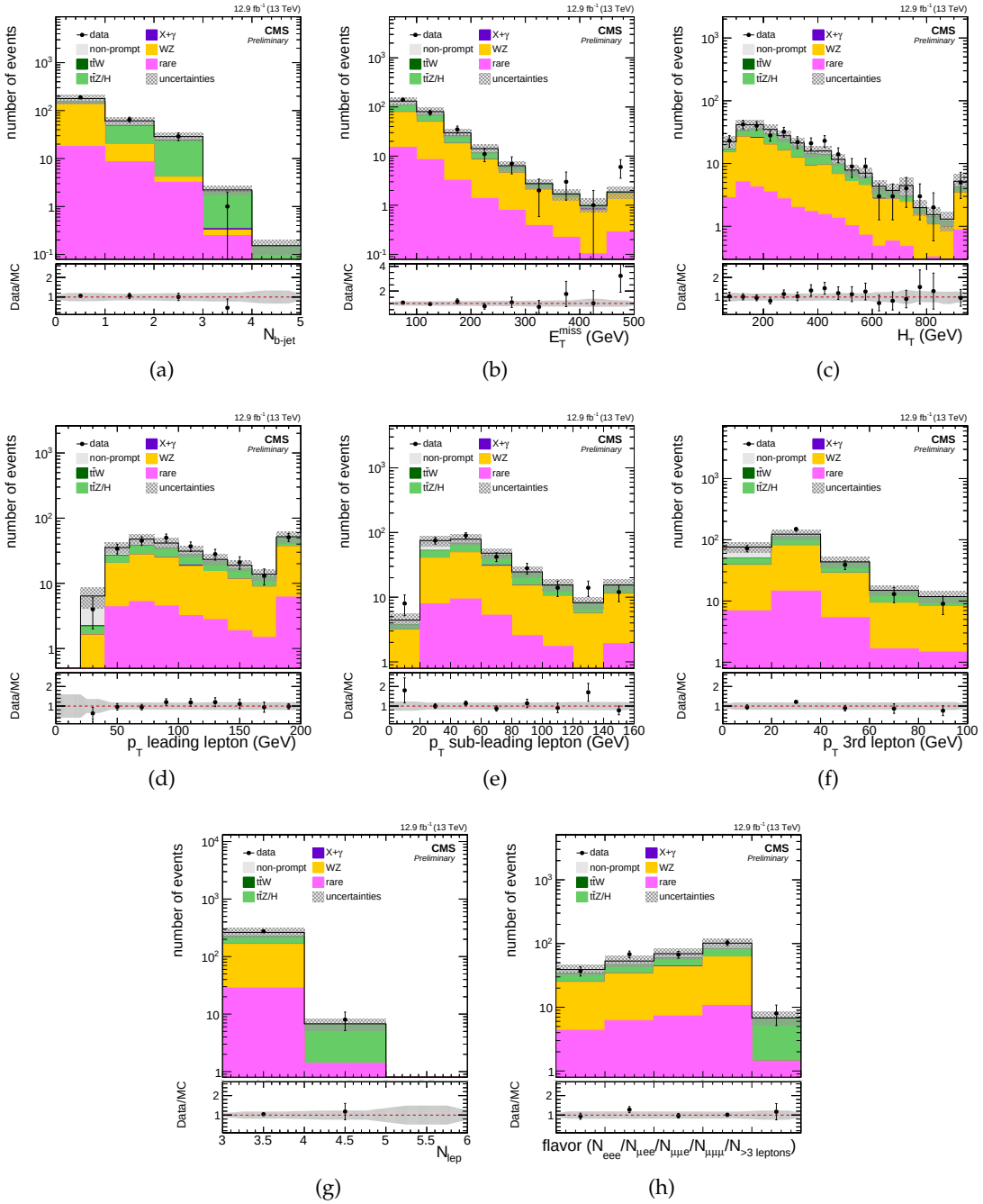


Figure 3: Background prediction and observation in key observables of the on-Z baseline selection for 12.9 fb^{-1}. Good agreement between data and expected background yields is observed in the event variables used for signal region categorization – the b-jet multiplicity, $E_{\mathrm{T}}^{\mathrm{miss}}$, and H_{T}. Additionally, the distibutions of the lepton p_{T} spectra, the flavor composition of the leptons, and the lepton multiplicity in the event are shown. The hatched area represents the statistical and systemtic uncertainties on the prediction.

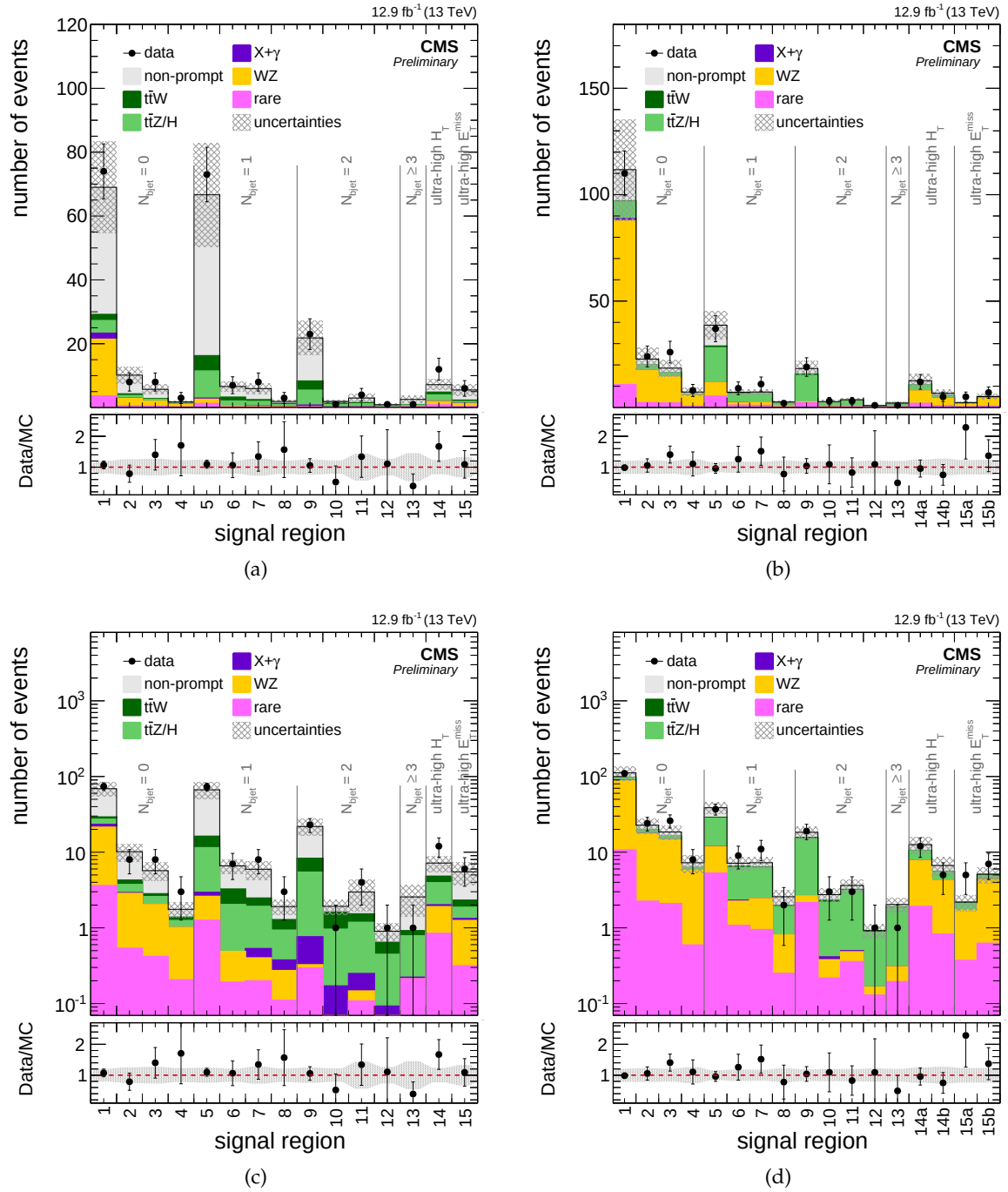


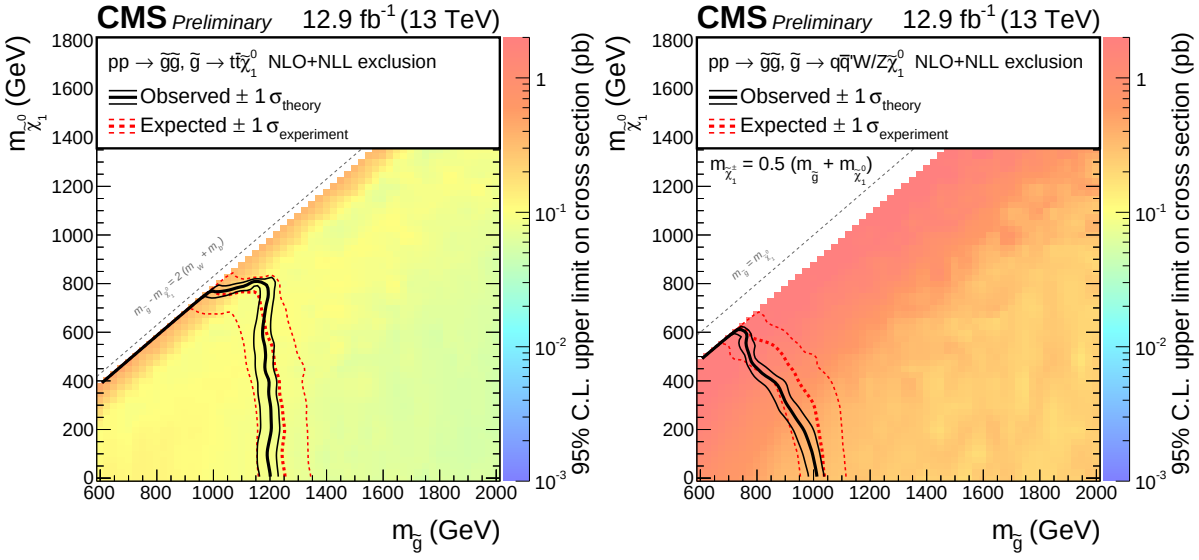
Figure 4: Background prediction and observation in the 15 off-Z signal regions (a,c) and in the 17 on-Z signal regions (b,d) with linear and logarithmic y-axis. No significant deviation between data and prediction can be found in 12.9 fb^{-1} of data.

Table 4: The total expected yields in the off-Z search regions with the 12.9 fb^{-1} of data. Uncertanties are given as $\pm$ stat $\pm$ syst.

b-tags	H_T (GeV)	E_T^{miss} (GeV)	expected	observed	SR
0 b-tags	60-400	50-150	$69.0 \pm 4.9 ^{+13.5}_{-13.4}$	74	SR1
		150-300	$10.2 \pm 1.7 ^{+2.0}_{-2.0}$	8	SR2
	400-600	50-150	$5.8 \pm 1.1 ^{+1.1}_{-1.1}$	8	SR3
		150-300	$1.8 \pm 0.3 ^{+0.3}_{-0.3}$	3	SR4
1 b-tags	60-400	50-150	$66.7 \pm 5.0 ^{+15.4}_{-15.4}$	73	SR5
		150-300	$6.6 \pm 0.8 ^{+1.2}_{-1.2}$	7	SR6
	400-600	50-150	$6.0 \pm 1.2 ^{+1.1}_{-1.1}$	8	SR7
		150-300	$1.9 \pm 0.3 ^{+0.3}_{-0.3}$	3	SR8
2 b-tags	60-400	50-150	$21.8 \pm 3.2 ^{+4.3}_{-4.3}$	23	SR9
		150-300	$2.0 \pm 0.3 ^{+0.3}_{-0.3}$	1	SR10
	400-600	50-150	$3.0 \pm 1.2 ^{+0.5}_{-0.5}$	4	SR11
		150-300	$0.9 \pm 0.2 ^{+0.1}_{-0.1}$	1	SR12
$\geq$ 3 b-tags	60-600	50-300	$2.6 \pm 1.0 ^{+0.5}_{-0.5}$	1	SR13
inclusive	$>$ 600	50-300	$7.2 \pm 1.1 ^{+1.3}_{-1.2}$	12	SR14
inclusive	inclusive	$\geq$ 300	$5.5 \pm 1.5 ^{+1.0}_{-1.1}$	6	SR15

Table 5: The total expected yields in the on-Z search regions with the 12.9 fb^{-1} of data. Uncertanties are given as $\pm$ stat $\pm$ syst.

b-tags	H_T (GeV)	E_T^{miss} (GeV)	expected	observed	SR
0 b-tags	60-400	70-150	$111.8 \pm 3.4 ^{+23.4}_{-23.8}$	110	SR1
		150-300	$22.7 \pm 1.7 ^{+4.8}_{-4.7}$	24	SR2
	400-600	50-150	$18.5 \pm 0.9 ^{+3.8}_{-3.8}$	26	SR3
		150-300	$7.3 \pm 1.1 ^{+1.5}_{-1.5}$	8	SR4
1 b-tags	60-400	70-150	$38.7 \pm 1.7 ^{+6.2}_{-6.2}$	37	SR5
		150-300	$7.2 \pm 0.4 ^{+1.3}_{-1.3}$	9	SR6
	400-600	50-150	$7.3 \pm 0.5 ^{+1.2}_{-1.2}$	11	SR7
		150-300	$2.6 \pm 0.4 ^{+0.4}_{-0.4}$	2	SR8
2 b-tags	60-400	50-150	$18.4 \pm 1.2 ^{+3.3}_{-3.3}$	19	SR9
		150-300	$2.8 \pm 0.3 ^{+0.5}_{-0.5}$	3	SR10
	400-600	50-150	$3.6 \pm 0.2 ^{+0.7}_{-0.7}$	3	SR11
		150-300	$0.9 \pm 0.0 ^{+0.2}_{-0.2}$	1	SR12
$\geq$ 3 b-tags	60-600	50-300	$2.0 \pm 0.1 ^{+0.4}_{-0.4}$	1	SR13
inclusive	$>$ 600	50-150	$12.6 \pm 1.4 ^{+2.8}_{-2.8}$	12	SR14a
inclusive	$>$ 600	150-300	$6.7 \pm 1.1 ^{+1.5}_{-1.5}$	5	SR14b
inclusive	60-400	$\geq$ 300	$2.2 \pm 0.2 ^{+0.5}_{-0.5}$	5	SR15a
inclusive	$>$ 400	$\geq$ 300	$5.1 \pm 0.4 ^{+1.2}_{-1.2}$	7	SR15b



(a) (b)

Figure 5: Excluded region at 95% confidence in the $m(\tilde{\chi}^0)$ versus $m(\tilde{g})$ plane for the T1tttt (a) and for the T5qqqqWZ (b) simplified model. The color scale indicates the excluded cross section at a given point in the mass plane. The excluded regions are to the left and below the observed and expected limit curves.

Table 6: Most sensitive signal regions for an uncompressed ($m_{\tilde{g}}$=1200 GeV, $m_{\tilde{\chi}^0_1}$=100 GeV) and a compressed ($m_{\tilde{g}}$=1200 GeV, $m_{\tilde{\chi}^0_1}$=700 GeV) mass scenario of the T1tttt simplified model. For the three most sensitive signal regions the number of expected background and signal events, the observed yield, and the expected and observed exclusion limit in terms of signal strength modifier are shown.

$m_{\tilde{g}}$ / $m_{\tilde{\chi}^0_1}$ (GeV)	SR	$N^{\text{exp}}_{\text{bkg}}$	N_{sig}	N_{obs}	exp. limit	obs. limit
1200 / 100	15 off-Z	$5.5 \pm 1.5\,^{+1.0}_{-1.1}$	$9.0 \pm 0.4\,^{+3.4}_{-2.8}$	6	0.9	1.0
	14 off-Z	$7.2 \pm 1.1\,^{+1.3}_{-1.2}$	$4.2 \pm 0.3\,^{+1.5}_{-1.3}$	12	2.2	3.8
	15b on-Z	$5.1 \pm 0.4\,^{+1.2}_{-1.2}$	$1.6 \pm 0.2\,^{+0.6}_{-0.5}$	7	4.7	6.2
1200 / 700	13 off-Z	$2.6 \pm 1.0\,^{+0.5}_{-0.5}$	$2.6 \pm 0.2\,^{+1.0}_{-0.8}$	1	1.5	2.0
	15 off-Z	$5.5 \pm 1.5\,^{+1.0}_{-1.1}$	$3.6 \pm 0.2\,^{+1.4}_{-1.1}$	6	2.5	2.3
	14 off-Z	$7.2 \pm 1.1\,^{+1.3}_{-1.2}$	$2.3 \pm 0.2\,^{+0.9}_{-0.7}$	12	4.0	7.0

sensitivity. It can be seen that off-Z signal region 15 ($E_{\text{T}}^{\text{miss}} > 300$ GeV) is the most sensitive one for uncompressed spectra while off-Z SR 13 ($N_{\text{b jets}} \geq 3$, $E_{\text{T}}^{\text{miss}} < 300$ GeV) drives the limit in the absence of large $E_{\text{T}}^{\text{miss}}$ and H_{T} for compressed spectra. The observed downward fluctuation in off-Z SR 13 and the upward fluctuation in off-Z SR 15 explain the relative position of the observed limit with respect to the expected limit in Figure 5a. Similarly, Table 7 shows that the on-Z signal region 15b ($E_{\text{T}}^{\text{miss}} > 300$ GeV, $H_{\text{T}} > 600$ GeV) is the most sensitive SR for an uncompressed scenario of the T5qqqqWZ model with a gluino mass of 1000 GeV and an LSP mass of 100 GeV. For a more compressed spectrum with a gluino mass of 800 GeV and an LSP mass of 500 GeV, on-Z signal region 2 ($N_{\text{b jets}} = 0$, 60 GeV $< H_{\text{T}} <$ 400 GeV, 150 GeV $< E_{\text{T}}^{\text{miss}} <$ 300 GeV) shows the leading sensitivity.

Table 7: Most sensitive signal regions for an uncompressed ($m_{\tilde{g}}$=1000 GeV, $m_{\tilde{\chi}^{\pm}}$=550 GeV, $m_{\tilde{\chi}_1^0}$=100 GeV) and a compressed ($m_{\tilde{g}}$=800 GeV, $m_{\tilde{\chi}^{\pm}}$=650 GeV, $m_{\tilde{\chi}_1^0}$=500 GeV) mass scenario of the T5qqqqWZ simplified model. For the three most sensitive signal regions the number of expected background and signal events, the observed yield, and the expected and observed exclusion limit in terms of signal strength modifier are shown.

$m_{\tilde{g}}$ / $m_{\tilde{\chi}_1^0}$ (GeV)	SR	N_{bkg}^{exp}	N_{sig}	N_{obs}	exp. limit	obs. limit
1000 / 100	15b on-Z	$5.1 \pm 0.4\,^{+1.2}_{-1.2}$	$8.4 \pm 0.5\,^{+3.1}_{-2.5}$	7	0.9	1.2
	14b on-Z	$6.7 \pm 1.1\,^{+1.5}_{-1.5}$	$3.1 \pm 0.3\,^{+1.1}_{-0.9}$	5	2.8	2.4
	14a on-Z	$12.6 \pm 1.4\,^{+2.8}_{-2.8}$	$1.5 \pm 0.2\,^{+0.6}_{-0.5}$	12	8.6	8.2
800 / 500	2 on-Z	$22.7 \pm 1.7\,^{+4.8}_{-4.7}$	$13.9 \pm 1.5\,^{+5.2}_{-4.3}$	24	1.3	1.4
	15b on-Z	$5.1 \pm 0.4\,^{+1.2}_{-1.2}$	$5.0 \pm 0.8\,^{+2.5}_{-2.1}$	7	1.8	2.5
	15a on-Z	$2.2 \pm 0.2\,^{+0.5}_{-0.5}$	$2.5 \pm 0.6\,^{+1.0}_{-0.8}$	5	2.3	4.3

7 Conclusions

An analysis to search for beyond the standard model physics in final states with $\geq$ 3 leptons, electrons or muons, using 12.9 fb^{-1} of data collected with the CMS detector in 2016 at $\sqrt{s} = 13$ TeV has been presented. The analysis makes use of data-driven techniques to estimate reducible backgrounds and validates simulation for use in estimating irreducible background processes. To maximize sensitivity to a broad range of possible signal models, 32 exclusive signal regions have been investigated. No significant deviation from the expected standard model background has been observed.

In the absence of any observed excesses in the data, the result has been interpreted using a simplified gluino-pair production model that features cascade decays producing four top quarks in the final state. In this model, we exclude gluinos with a mass of up to $\sim$1250 GeV in the case of a massless LSP. The maximum excluded LSP mass is $\sim$750 GeV for gluino masses up to 1150 GeV. In both masses, this represents an improvement of the order of 100 GeV with respect to the exclusion limit set in a similar search based on 2.3 fb^{-1} collected with the CMS detector in 2015 [1]. For the simplified model with gluino-gluino production and light jets and each one W and Z boson in the final state, gluino masses up to $\sim$1025 GeV and neutralino masses up to $\sim$600 GeV can be excluded. The limit on gluino mass for a light neutralino extends the corresponding limit from the previous analysis by about 200 GeV, while the limit on the LSP mass improves by about 100 GeV.

References

[1] CMS Collaboration, "Search for SUSY with multileptons in 13 TeV data", *CMS Physics Analysis Summary* **SUS-16-003** (2016).

[2] CMS Collaboration, "Search for anomalous production of events with three or more leptons in pp collisions at $\sqrt{s} = 8$ TeV", *Phys. Rev. D* **90** (2014) 032006, doi:10.1103/PhysRevD.90.032006, arXiv:1404.5801.

[3] CMS Collaboration, "Searches for supersymmetry based on events with b jets and four W bosons in pp collisions at 8 TeV", *Phys. Lett. B* **745** (2015) 5, doi:10.1016/j.physletb.2015.04.002, arXiv:1412.4109.

[4] ATLAS Collaboration, "Search for supersymmetry at $\sqrt{s}$=8 TeV in final states with jets and two same-sign leptons or three leptons with the ATLAS detector", *JHEP* **06** (2014) 035, doi:10.1007/JHEP06(2014)035, arXiv:1404.2500.

[5] CMS Collaboration, "Performance of CMS muon reconstruction in pp collision events at sqrt(s) = 7 TeV", *JINST* **7** (2012) P10002, doi:10.1088/1748-0221/7/10/P10002, arXiv:1206.4071.

[6] CMS Collaboration, "Performance of Electron Reconstruction and Selection with the CMS Detector in Proton-Proton Collisions at vs = 8 TeV", *JINST* **10** (2015) P06005, doi:10.1088/1748-0221/10/06/P06005, arXiv:1502.02701.

[7] CMS Collaboration, "Commissioning of the Particle-Flow Reconstruction in Minimum-Bias and Jet Events from pp Collisions at 7 TeV", CMS Physics Analysis Summary CMS-PAS-PFT-10-002, 2010.

[8] CMS Collaboration, "Identification of b-quark jets with the CMS experiment", *JINST* **8** (2013) P04013, doi:10.1088/1748-0221/8/04/P04013, arXiv:1211.4462.

[9] CMS Collaboration, "Performance of b tagging at sqrt(s)=8 TeV in multijet, $t\bar{t}$ and boosted topology events", CMS Physics Analysis Summary CMS-PAS-BTV-13-001, 2013.

[10] CMS Collaboration, "Performance of the missing transverse energy reconstruction by the CMS experiment in $\sqrt{s}$ = 8 TeV pp data", (2014). arXiv:1411.0511. Submitted to *JINST*.

[11] 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, doi:10.1007/JHEP07(2014)079, arXiv:1405.0301.

[12] T. Melia, P. Nason, R. Rontsch, and G. Zanderighi, "W+W-, WZ and ZZ production in the POWHEG BOX", *JHEP* **11** (2011) 078, doi:10.1007/JHEP11(2011)078, arXiv:1107.5051.

[13] T. Sjostrand, S. Mrenna, and P. Z. Skands, "A Brief Introduction to PYTHIA 8.1", *Comput. Phys. Commun.* **178** (2008) 852, doi:10.1016/j.cpc.2008.01.036, arXiv:0710.3820.

[14] CMS Collaboration, "Event generator tunes obtained from underlying event and multiparton scattering measurements", *Eur. Phys. J. C* **76** (2016) 155, doi:10.1140/epjc/s10052-016-3988-x, arXiv:1512.00815.

[15] GEANT4 Collaboration, "GEANT4—a simulation toolkit", *Nucl. Instrum. Meth. A* **506** (2003) 250, doi:10.1016/S0168-9002(03)01368-8.

[16] S. Abdullin et al., "The fast simulation of the CMS detector at LHC", *J. Phys. Conf. Ser.* **331** (2011) 032049, doi:10.1088/1742-6596/331/3/032049.

[17] W. Beenakker et al., "Production of charginos, neutralinos, and sleptons at hadron colliders", *Phys. Rev. Lett.* **83** (1999) 3780, doi:10.1103/PhysRevLett.83.3780, arXiv:hep-ph/9906298.

[18] A. Kulesza and L. Motyka, "Threshold resummation for squark-antisquark and gluino-pair production at the LHC", *Phys. Rev. Lett.* **102** (2009) 111802, doi:10.1103/PhysRevLett.102.111802, arXiv:0807.2405.

[19] 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, doi:10.1103/PhysRevD.80.095004, arXiv:0905.4749.

[20] W. Beenakker et al., "Soft-gluon resummation for squark and gluino hadroproduction", *JHEP* **12** (2009) 041, doi:10.1088/1126-6708/2009/12/041, arXiv:0909.4418.

[21] W. Beenakker et al., "Squark and gluino hadroproduction", *Int. J. Mod. Phys. A* **26** (2011) 2637, doi:10.1142/S0217751X11053560, arXiv:1105.1110.

[22] J. Butterworth et al., "PDF4LHC recommendations for LHC Run II", *J. Phys.* **G43** (2016) 023001, doi:10.1088/0954-3899/43/2/023001, arXiv:1510.03865.