Available on the CERN CDS information server **CMS PAS SUS-11-028**

CMS Physics Analysis Summary

Contact: cms-pag-conveners-susy@cern.ch 2012/05/04

Search for Supersymmetry in Final States with a Single Lepton, B-jets, and Missing Transverse Energy in Proton-Proton Collisions at $\sqrt{s} = 7$ TeV

The CMS Collaboration

Abstract

Motivated by supersymmetric models with light top and bottom squarks, a search for supersymmetry in final states with a single lepton, b-jets, and missing transverse energy is performed. The analysis is based on data recorded at the CMS experiment during proton-proton collisions at a center of mass energy of $\sqrt{s} = 7$ TeV during 2011, corresponding to an integrated luminosity of 4.96 fb^{-1}. Results are interpreted in the context of the constrained Minimal Supersymmetric Standard Model and different heavy flavor simplified models.

1 Introduction

Searches for new physics in states with 3$^{\text{rd}}$ generation quarks at the Large Hadron Collider (LHC) are motivated by various extensions of the Standard Model (SM). Amongst these, supersymmetric models [1–6] are regarded as attractive, because they resolve the hierarchy problem and may permit the unification of the electroweak and strong interactions.

Supersymmetry (SUSY) predicts that for each particle in the SM there exists a partner particle, often referred to as a sparticle, with identical gauge quantum numbers, but spins differing by 1/2. Assuming R parity[1] conservation, sparticles are produced in pairs and their decay chains terminate with the lightest supersymmetric particle (LSP). In some scenarios the LSP is the lightest neutralino ($\tilde{\chi}_1^0$), a heavy, electrically neutral, weakly interacting particle, which is a viable dark-matter candidate. In this case, SUSY events are characterised by missing transverse energy in the final state. There exist models in which top ($\widetilde{\mathrm{t}}$) or bottom ($\widetilde{\mathrm{b}}$) squarks may be copiously produced at the LHC. This may happen by direct squark production e.g.

$$\tilde{\mathrm{t}}\tilde{\mathrm{t}}^* \rightarrow \mathrm{t}\bar{\mathrm{t}}\tilde{\chi}_1^0\tilde{\chi}_1^0.$$

Another example is a model where the mass of the gluino ($\tilde{\mathrm{g}}$) is larger than the masses of the third generation squarks, but lighter than the squarks of the first two generations. Consequently, the gluino may dominantly decay into the third generation squarks. As the decays of the top and bottom squarks into light quarks are strongly flavor suppressed, this may result in an excess of events with large b-jet multiplicities, e.g.

$$\tilde{\mathrm{g}} \rightarrow \mathrm{t}\tilde{\mathrm{t}}^* \rightarrow \mathrm{t}\bar{\mathrm{t}}\tilde{\chi}_1^0.$$

Hence gluino pair production could even produce events containing four 3$^{\text{rd}}$ generation quarks and isolated leptons may originate from leptonically decaying top quarks and two- or three-body decays of neutralinos and charginos.

In this note a search for SUSY in final states with a single lepton, b-jets, and missing transverse energy ($\not\!\!E_{\mathrm{T}}$) is described. The search is based on the full set of data recorded at the CMS experiment during proton-proton collisions at a center of mass energy of $\sqrt{s} = 7$ TeV during 2011, which corresponds to an integrated luminosity of 4.96 fb^{-1}.

The analysis presented here is not limited to a particular theory. However, due to its simplicity, the constrained Minimal Supersymmetric extension of the Standard Model (cMSSM) [7, 8] is chosen as a benchmark to illustrate the sensitivity of this search for new physics processes. In addition, the impact of the analysis is also tested on different heavy flavor simplified models.

A brief description of the CMS detector is given in Section 2. The datasets and simulated event samples used in this analysis are discussed in Section 3. In Section 4 the selection of physics objects and events is outlined. A definition of the signal region and a description of the data-driven background estimation method are given in Section 5. Systematic uncertainties and selection efficiency corrections applied on simulated events are discussed in Section 6. Results are summarized in Section 7 and interpreted in Section 8.

[1] R parity is defined as $R = (-1)^{3(B-L)+2S}$, where B is the baryon number, L the lepton number and S the spin.

2 CMS Detector

The CMS detector is a multipurpose apparatus designed to study high-p_T physics processes in proton-proton collisions, as well as a broad range of phenomena in heavy-ion collisions. The central feature of the detector is a superconducting solenoid, 13 m in length and 6 m in diameter, which provides an axial magnetic field of 3.8 T. Within the magnet are (in order of increasing radius from the beam pipe) the high-precision silicon-pixel and silicon-strip detectors for charged particle tracking, covering $0 < \phi < 2\pi$ in azimuth, and $|\eta| < 2.5$. The angle ϕ is measured in a plane perpendicular to the z axis, which points along the direction of the counterclockwise rotating beam, while the pseudorapidity η is defined as $\eta = -\ln[\tan(\theta/2)]$, with θ being the polar angle of the trajectory of a particle with respect to the z axis. The separation between two momentum vectors in η-ϕ space is characterized by the quantity $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$, which is approximately invariant under Lorentz boosts along the z axis.

A lead tungstate crystal electromagnetic calorimeter for measurements of photons, electrons, and the electromagnetic component of jets, and a hadron calorimeter, constructed from scintillating tiles and brass absorbers, for jet-energy measurements, surround the tracking volume and cover the region $|\eta| < 3$. A quartz-steel Cerenkov-radiation-based forward hadron calorimeter extends the coverage to $|\eta| \leq 5$. The detector is nearly hermetic, allowing for energy balance measurements in the plane transverse to the beam directions. Beyond the magnet is the muon system, comprising drift-tube, cathode-strip, and resistive-plate detectors interleaved with steel absorbers. Each detector system comprises subsystems that cover the central (barrel) and forward (endcap) regions. A detailed description of the CMS detector can be found elsewhere [9].

3 Event Samples

The data analyzed here has been recorded using triggers requiring the presence of a muon or electron with a transverse momentum (p_T) above a certain threshold in association with a significant hadronic activity, quantified by H_T^{trigger}, and missing transverse momentum $\not\!\!E_T{}^{\text{trigger}}$. The variable H_T^{trigger} is the scalar sum of the p_T from jets reconstructed in the calorimeter [10] with $p_T > 40$ GeV, whereas $\not\!\!E_T{}^{\text{trigger}}$ is calculated based on the CMS particle-flow algorithm [11, 12], which combines information from different sub-detectors to reconstruct all visible particles. The absolute value of the vectorial sum of these particles is defined as $\not\!\!E_T{}^{\text{trigger}}$.

In order not to exceed the maximal rate of data acquisition and processing, trigger thresholds have been raised with increasing LHC luminosity, resulting in a threshold for muon transverse momentum p_T from $p_T(\mu) > 8$ GeV to $p_T(\mu) > 15$ GeV; and for electrons from 10 to 15 GeV. The requirement on the hadronic activity has been raised from $H_T^{\text{trigger}} > 200$ GeV to $H_T^{\text{trigger}} > 300$ GeV in the muon and $H_T^{\text{trigger}} > 250$ GeV in the electron channel. The requirement on the missing transverse energy, absent in the first data-taking period, has been introduced with $\not\!\!E_T{}^{\text{trigger}} > 20$ GeV and raised towards the end of 2011 to $\not\!\!E_T{}^{\text{trigger}} > 40$ GeV. Only runs where all sub-detectors performed well are included in this analysis.

Simulated event samples were produced using different event generators and the GEANT4 package [13] for the full detector simulation except for the SUSY scans discussed below. The production and the decay of $t\bar{t}$ pairs and vector bosons in association with the production of hard jets are generated using MADGRAPH [14]. The produced parton events are then passed to PYTHIA6 [15] with tune Z2 [16] for simulating parton showers, multiple interactions and fragmentation processes. The decay of τ leptons is simulated using TAUOLA [17]. The pro-

duction and decay of single top and anti-top quarks are simulated with the event generator POWHEG [18, 19] and TAUOLA interfaced to PYTHIA6. QCD-multijet production is simulated using PYTHIA6 directly.

SUSY mass spectra and branching ratios have been calculated at the electroweak scale using the renormalization equations implemented in the SOFTSUSY package [20], interfaced to PYTHIA6. The low mass scenarios LM3, LM8 and LM13 [21] are used as cMSSM benchmark scenarios, as they are consistent with the assumption of one top or bottom squark being lighter than the quarks of the 1st and 2nd generation. In addition, the gluino is lighter than the 1st and 2nd generation squarks in these scenarios, resulting in a large branching ratios for its decays into 3rd generation quarks. The LM3 (LM8, LM13) scenario is defined by $m_0 = 330$ GeV (500 GeV, 270 GeV), $m_{1/2} = 240$ GeV (300 GeV, 218 GeV) $A_0 = 0$ GeV (-300 GeV, -553 GeV), $\tan\beta = 20$ (10, 40), and $\mu > 0$. The leading order (LO) cross section for the LM3 (LM8, LM13) scenario is 3.4 pb (0.73 pb, 6.90 pb). The corresponding next-to leading order (NLO) k-factors, calculated with PROSPINO [22], are 1.40, 1.41 and 1.42, respectively. The cMSSM benchmark points have been generated with the full detector simulation. A scan in the cMSSM parameter space is performed for a fixed set of parameters A_0, $\tan\beta$ and sign μ, where a grid in the m_0 - $m_{1/2}$ plane is defined by varying m_0 and $m_{1/2}$ in steps of a size of 20 GeV. NLO k factors are applied to all contributing subprocesses. For each grid point in the scans 10 000 events are generated using a fast detector simulation.

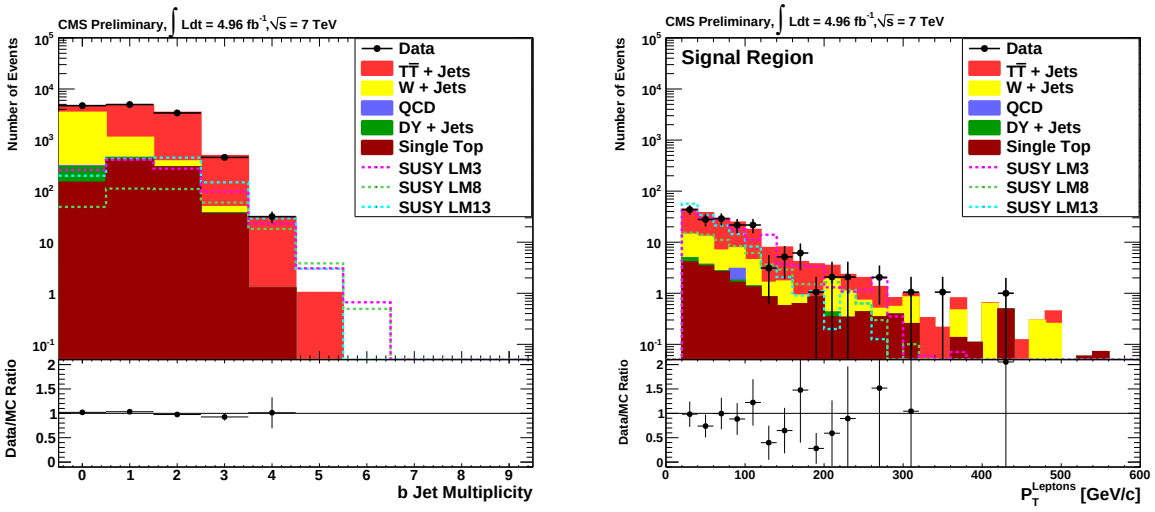
In addition, the results are interpreted in the context of two different simplified models [23]. The model "T1tttt" contains the pair-production of gluinos, which subsequently decay with branching fraction: $\mathcal{B}(\tilde{g} \to t + \bar{t} + \tilde{\chi}_1^0) = 1$. The model "T2ttww" covers the direct production of a $\tilde{b}\tilde{b}^*$ pair decaying each to a top quark and a chargino ($\tilde{\chi}_1^{\pm}$), with the chargino decaying to a W boson and the lightest neutralino ($\tilde{\chi}_1^0$). The chargino mass in this model is fixed to be half-way between the squark and the neutralino mass. For this model, only the acceptance times efficiency is presented, as its cross-section is too low to allow the exclusion of any part of the parameter space with the current luminosity.

4 Event Selection

Except for a lower $\not\!\!E_T$ criterion and additional b-jet requirements, the selection of physics objects and events is identical to that applied in the CMS search for new physics with a single lepton based on 2011 data [24].

The primary vertex must satisfy a set of quality requirements, including the longitudinal and transverse distances of the primary vertex with respect to the nominal interaction point, $|z_{PV}| <$ 24 cm and $\rho_{PV} < 2$ cm, respectively.

Muon candidates [25] are required to satisfy $p_T(\mu) > 20$ GeV and $|\eta| < 2.1$. The reconstructed track of a muon candidate must have an impact parameter $d_0 < 0.02$ cm in the transverse plane with respect to the beam spot and an impact parameter $|d_z| < 1.0$ cm with respect to the primary vertex along the z direction. To suppress background contributions from muons originating from heavy flavor quark decays, the muon is required to be isolated within a cone of size $\Delta R = 0.3$. The relative combined isolation of the muon, defined as the sum of the transverse energy E_T (as measured in the electromagnetic and hadron calorimeters) and the transverse momentum p_T (as measured in the silicon tracker) of all reconstructed objects within this cone, excluding the muon, $I_{rel}^{comb} = \sum_{\Delta R<0.3}(E_T + p_T)/p_T(\mu)$, is required to satisfy $I_{rel}^{comb} <$ 0.1.



(a) (b)

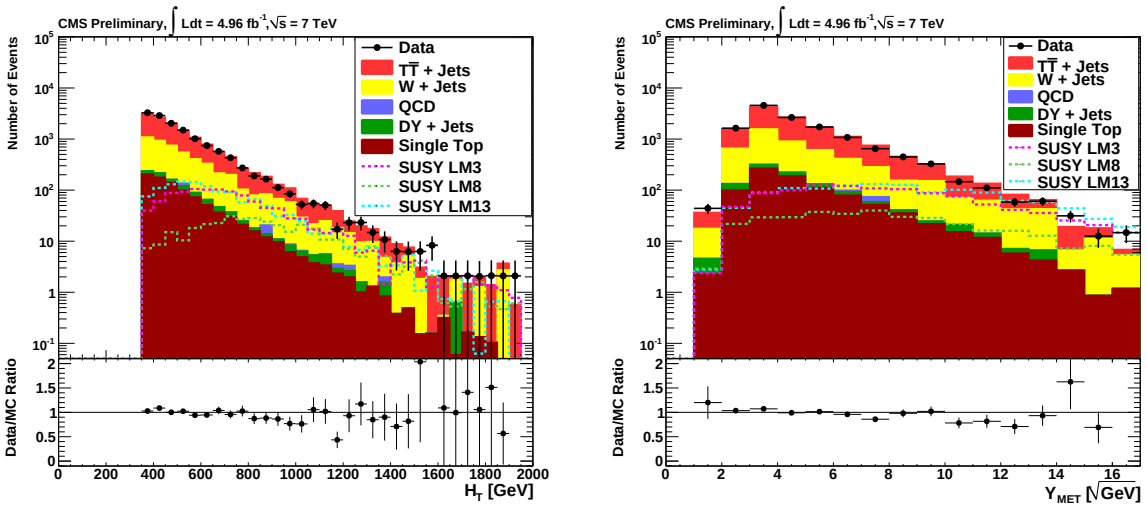
Figure 1: The number of b-tagged jets for data and simulated SM events after all selection steps except the b-jet requirement (a) and the p_{T} of the leptons in our signal region D, as defined in Section 5, including the requirement of at least one b-tagged jet (b). The muon and electron channels are combined. The CMS data are represented by solid points and the simulated SM events by stacked histograms. The three dashed lines represent possible signal scenarios. The simulation is normalized to the integrated luminosity.

Electron candidates [26] are required to have $p_{\mathrm{T}} > 20$ GeV and $|\eta| < 2.4$, excluding the barrel-endcap transition region ($1.44 < |\eta| < 1.57$). The reconstructed track of an electron candidate has to fulfill the same impact parameter requirements as the muon tracks described above. The relative combined isolation variable, similar to that defined in the muon case, must satisfy $I_{\mathrm{rel}}^{\mathrm{comb}} < 0.07$ in the barrel region and $I_{\mathrm{rel}}^{\mathrm{comb}} < 0.06$ in the endcaps, as well as a set of quality and photon-conversion rejection criteria. Exactly one selected muon or electron are required to be in the event. Events with a second lepton passing looser selection criteria, e.g. $p_{\mathrm{T}} > 15$ GeV and looser isolation requirements, are rejected.

The reconstruction of jets is based on the CMS particle-flow algorithm. Extra energy clustered into jets due to pileup is taken into account with an event-by-event correction to the jet four-vectors. Jets are reconstructed from particle-flow candidates using the anti-k_{T} clustering algorithm [27] with a distance parameter of 0.5. Different corrections are applied on the raw jet energy to obtain a relative uniform response across the detector in η and an absolute calibrated response in p_{T} [28]. Each event is required to contain at least four jets with $p_{\mathrm{T}} > 40$ GeV and $|\eta| < 2.4$ which are spatially separated from a selected muon or electron by $\Delta R > 0.3$ and pass different quality criteria in order to suppress noise and spurious energy deposits.

The missing transverse energy is reconstructed by summing up the transverse momentum vectors of all particle-flow objects and required to be larger in magnitude than 60 GeV. A measure of the hadronic activity, H_{T}, calculated by summing up the p_{T} of selected jets, is required to be larger than 375 GeV.

Except for the $\not\!\!E_{\mathrm{T}}$ requirement, the offline selection criteria are designed to be well above the trigger thresholds, where the efficiency reaches a plateau. For events with $\not\!\!E_{\mathrm{T}} < 80$ GeV the efficiency of the triggers with a $\not\!\!E_{\mathrm{T}}^{\mathrm{trigger}}$ threshold of 40 GeV drops down to 80%, which is corrected for in this analysis, as discussed in Section 6.



(a) (b)

Figure 2: H_T (a) and Y_{MET} (b) for data compared to the different SM processes. At least one b-tag is required. The muon and electron channels are combined. The CMS data are represented by solid points and the simulated SM events by stacked histograms. The three dashed lines represent possible signal scenarios. The simulation is normalized to the integrated luminosity.

Jets are tagged as b-jets if they have at least two tracks with impact parameter significance greater than 3.3 (Track Counting algorithm). B-tag efficiency and mistag rate (the efficiency of tagging a c-jet, light quark or gluon jet as b-jet) have been measured up to transverse momenta of $p_{\mathrm{T}} = 670$ GeV [29].

The multiplicity of b-jets in the selected events and the p_{T} spectrum of the lepton in the signal region D (as defined in Section 5) are shown in Fig. 1.

5 Background Estimation from Data

Most SUSY events are expected to have several high-energetic jets leading to high H_T, while the lightest SUSY particles are expected to lead in many scenarios to a significant amount of missing energy in the detector. Therefore, the signal region is defined as the region with $H_T > 650$ GeV and $Y_{\mathrm{MET}} > 5.5\sqrt{\mathrm{GeV}}$, with $Y_{\mathrm{MET}} \equiv \not{E}_{\mathrm{T}} / \sqrt{H_T}$. Y_{MET} is sometimes called $\not{E}_{\mathrm{T}}$ significance, as the denominator is proportional to the uncertainty on $\not{E}_{\mathrm{T}}$ arising from jet mismeasurements. This signal region is populated with events described by the tails of SM distributions, and smearing effects related to the finite detector resolution. Since a reliable simulation of such contributions is subtle, the background is estimated from data.

H_T and the Y_{MET} are nearly uncorrelated for the main background, $\mathrm{t\bar{t}}$ events. Therefore, a factorization ansatz in the Y_{MET} - H_T plane can be used to estimate the background contribution, namely from control regions with low H_T and/or Y_{MET}. This method has been described in an earlier paper [30]. H_T and Y_{MET} are shown for the inclusive 1 b-tag selection for data and simulated SM events in Fig. 2.

The control regions A, B, C and the signal region D used for the factorization method with H_T and Y_{MET} are defined in Table 1.

The number of events expected in region D $\hat{N}_D$ is estimated from the three control regions:

Table 1: Definition of the different regions used for the factorization method with H_T and the $\not\!\!E_T$ significance Y_{MET}.

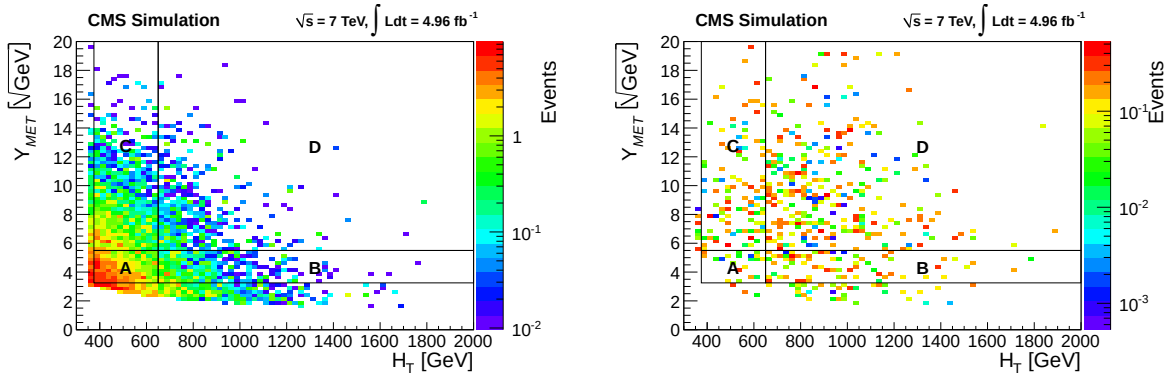
Region	H_T / GeV	Y_{MET} / $\sqrt{\mathrm{GeV}}$
A	$375 < H_T < 650$	$3.25 < Y_{\mathrm{MET}} < 5.5$
B	$650 < H_T$	$3.25 < Y_{\mathrm{MET}} < 5.5$
C	$375 < H_T < 650$	$5.5 < Y_{\mathrm{MET}}$
D	$650 < H_T$	$5.5 < Y_{\mathrm{MET}}$

$$\hat{N}_D = \kappa N_B \frac{N_C}{N_A}. \tag{1}$$

If the two variables are completely uncorrelated, the correlation factor κ is equal to unity. As Y_{MET} and H_T have a small correlation, κ differs slightly from one, which is discussed in Section 6.

The analysis is performed in three channels according to the number n of b-tags: exactly one b-tag, exactly two b-tags and three or more b-tags. In addition we include the exclusive no b-tag bin for cross checks. For the limit setting these bins are combined into one single channel with n or more b-tags, where n is selected to give the best expected limit. The best limits in the cMSSM scan and for the "T2ttww" simplified model are obtained with the inclusive one b-tag selection ($n \geq 1$), while for the "T1tttt" simplified model the inclusive three b-tag bin ($n \geq 3$) is favored. Our measurements, combined for muons and electrons, are sumarized in Table 5.

The distribution of events in the Y_{MET} - H_T plane after the event selection in the muon channel with the requirement of exactly 1 b-tag is presented in Fig. 3 for simulated SM events (a) compared to simulated signal events in the LM8 benchmark scenario (b). Corresponding distributions in the electron channel look very similar.



(a) (b)

Figure 3: Y_{MET} vs H_T for the SM background (a) compared to the low mass signal point LM8 (b) for the muon selection. Exactly one b-tag is required.

6 Efficiencies and Systematic Uncertainties

The uncertainties can be divided into two categories, the uncertainty on acceptance and efficiencies of the signal, and the uncertainty on the background prediction. They have been determined for the inclusive 1 b-tag requirement and for the exclusive b-tag selections.

The determination of the trigger efficiency is done separately for each part of the trigger; the leptonic, the H_T, and the $\not\!\!E_{\mathrm{T}}$ selection. The leptonic trigger selection is always found to be 97% to 98% efficient after the offline requirements.

The H_T requirements are even more efficient ($> 99\%$), as well as the requirement on $\not\!\!E_{\mathrm{T}}$ in earlier data taking period ($\not\!\!E_{\mathrm{T}}{}^{\mathrm{trigger}} > 20$ GeV). The $\not\!\!E_{\mathrm{T}}$ requirement in the later part of data taking ($\not\!\!E_{\mathrm{T}}{}^{\mathrm{trigger}} > 40$ GeV) has an efficiency of about 80% at the offline reconstructed $\not\!\!E_{\mathrm{T}}$ of 60 GeV.

The prediction in data obtained for region D is corrected for the trigger efficiency based on a fit with an error function to the turn-on curve. Triggers are not used in the simulation, but scale factors to account for the lower trigger efficiencies are applied to the simulation when compared to data.

The lepton efficiencies have been measured with a tag-and-probe method on the Z peak [31]: a sample of events containing two leptons with an invariant mass close to the mass of the Z boson is selected. One of the leptons has to fulfill tight selection criteria ("tag" lepton), while the 2$^{\mathrm{nd}}$ has to fulfill looser criteria ("probe" lepton), not including the requirement for which the efficiency is measured.

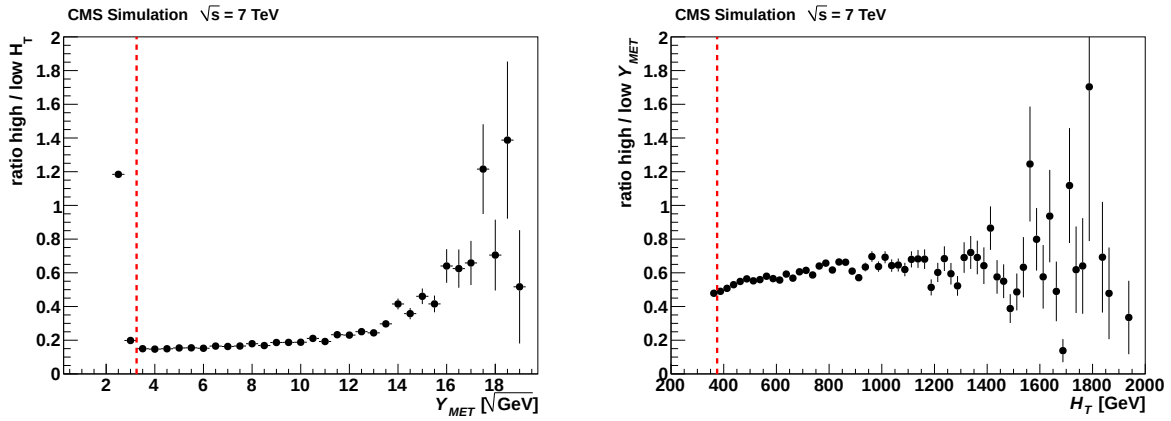
The overall muon selection efficiency including isolation is measured to be between 92% and 96%, depending on the detector region and the data taking period. It is found to be well described by the simulation except for one data taking period, where the forward detector had some inefficiencies, which are not included in the simulation. Here the efficiency between data and simulation differs by $\approx 3\%$.

The overall electron efficiency, which is divided into tracking and identification efficiency, is measured in a similar way. The electron tracking efficiency rises with the p_{T} of the electron. At the offline p_{T} threshold of 20 GeV the tracking efficiency is at the level of 96% and continues to increase with p_{T}, which is reasonably well reproduced by the simulation. The identification efficiency increases with p_{T} as well and saturates in data at about 87% (83%) in the central (forward) regions. This overall efficiency, including the p_{T} dependence, is described by the simulation within $\approx \pm 2\%$. The identification efficiency includes the isolation requirement, which depends on the number of vertices and jets. For high vertex and jet multiplicities an increasing deviation between data and simulation is observed, which is included in the systematic uncertainty given above.

We conclude that the total lepton selection efficiency in simulation is modeled to within 3%.

In the data-driven background determination most uncertainties cancel out. As an example, an upwards shift of the jet energy leads to larger H_T. Events in region A and C (with $H_T < 650$ GeV) might be shifted to the regions B and D, leading to a larger number of events in the signal region D. But, given that H_T and Y_{MET} are almost uncorrelated, the prediction for the events in region D, which is proportional to the number of events in region B, increases by the same amount.

Values of κ for the main SM background processes are shown in Table 2. For the dominant background, $\mathrm{t\bar{t}}$, as well as for Single Top production, it is slightly higher than one, indicating



(a) (b)

Figure 4: Ratio of events in the high H_T region to those in the low H_T region vs Y_{MET} for simulated $t\bar{t}$ events (a) and the ratio of events in the high Y_{MET} region to those in the low Y_{MET} region vs H_T. The behavior is linear except for the region with $Y_{\mathrm{MET}} > 13$, where less than 1% of all events are located. The dashed red line indicates the lower border of the control regions.

a small remaining correlation. Besides $t\bar{t}$ and Single Top we expect small contributions from W+jets and Z+jets events, for which the correlation factor can only be determined with large uncertainties due to very small statistics of the simulated samples, especially in the signal region D.

Table 2: Correlation factor κ of H_T and Y_{MET} for the main SM background processes and a different number of b-tags. For illustration, also the value for exclusive 0 b-tags is included. While the 0 b-tag bin is dominated by W+jets events, the channels which include b-tags contain mainly $t\bar{t}$ events. Only statistical uncertainties are displayed.

No. of b-tags	$\kappa(t\bar{t})$	κ(SingleTop)	κ(W+jets)	κ(All SM bkgs)
0 b-tags	1.14 ± 0.06	1.36 ± 0.14	1.01 ± 0.17	1.08 ± 0.13
1 b-tag	1.20 ± 0.05	1.29 ± 0.10	1.03 ± 0.20	1.19 ± 0.06
2 b-tags	1.19 ± 0.06	1.34 ± 0.12	1.42 ± 0.58	1.23 ± 0.06
≥ 3 b-tags	1.19 ± 0.08	1.36 ± 0.18	1.23 ± 0.91	1.21 ± 0.07
≥ 1 b-tags	1.19 ± 0.04	1.31 ± 0.07	1.08 ± 0.18	1.20 ± 0.04

The stability of the correlation factor κ has been tested extensively, e.g by analyzing the ratio of events in the high and low H_T (Y_{MET}) regions as defined in Table 1, binned as a function of Y_{MET} (H_T). This is demonstrated in Fig. 4 (a), where the ratio of simulated $t\bar{t}$ events with $H_T > 650$ GeV to events with 375 GeV$< H_T < 650$ GeV is shown vs Y_{MET}. The behavior is linear up to a value of $Y_{\mathrm{MET}} > 13$, less than 1% of all events are above that value. The analogous plot for the ratio of high over low Y_{MET} values vs H_T is presented in Fig. 4 (b). In addition, the shape of the H_T distribution has been compared for different intervals of Y_{MET} (and vice versa). The small correlation is accounted for by the value of κ.

The following experimental systematic uncertainties have been investigated. The effect of the jet energy scale (JES) uncertainty is determined by shifting the energy of jets with $p_{\mathrm{T}} > 10$ GeV and $|\eta| < 4.7$ in simulated events up and down by p_{T} and η dependent uncertainties that have been measured using dijet and γ/Z+jets events [10]. The applied uncertainties, which are 1-3%

for jets with $p_T > 40$ GeV and $|\eta| < 2.0$ and increase towards lower p_T and higher $|\eta|$, are propagated to the missing transverse energy. The simulated jet energy resolution (JER) [10] of jets with $p_T > 10$ GeV and $|\eta| < 4.7$ is globally increased by 10% to provide a more realistic description of the data. The uncertainty on the jet energy resolution is then determined by variation of the corrected simulated JER up and down by 10%, and propagated to $\not{E}_T$. The uncertainty on the energy of jets with $p_T < 10$ GeV, referred to as unclustered energy, is assumed to be 10%. This uncertainty is also propagated to the missing transverse energy. The lepton transverse momenta are shifted according to their resolution of 1% for the selected muon and 1% (2.5%) for the selected electron in the barrel (endcap) region, and propagated to the missing transverse energy calculation as well. The number of pileup (PU) interactions in data is estimated from the measured luminosity per bunch-crossing and the total inelastic cross-section. Corresponding uncertainties are investigated by scaling the distribution of the multiplicity of PU interactions in a way that the mean value gets shifted by ± 0.6. The efficiencies to tag b, c and light flavor jets were corrected in the simulation using scale factors (SF). These p_T and η dependent scale factors were obtained by comparing the simulated tagging efficiencies to those in data [29]. The applicability of the measured b-tag scale factors within this analysis has been cross-checked in different ways. The uncertainties on the measured SF are of the order of a few percent. Their effect is investigated by scaling up and down the scale factors in simulated events. This is done separately for the b-tag efficiency SF and the mistag rate SF. To account for possible uncertainties in the cross-section of the main SM processes, each cross-section has been scaled up and down by 50% and the corresponding uncertainty on κ has been determined. An additional 0.2% accounts for the uncertainty of the trigger efficiency correction for the prediction in data.

As the studies above are based on simulation, a cross-check is performed on data in the zero b-tag channel, which can be considered as signal-free, since previous analyses have already excluded this part of phase space [24]. From this channel a value of $\kappa = 1.15 \pm 0.13$ is observed in data, while for the SM simulation a value of 1.08 ± 0.13 is extracted in the same channel. Although the values are consistent within their statistical uncertainties, a smaller value of κ cannot be excluded with the present statistics in data and simulation. This is accounted for by an additional systematic uncertainty on κ of 10%. The uncertainties for the different selections are summarized in Table 3.

In summary, the value of κ is found to be consistent for all channels within their statistical fluctuations. Therefore, for the prediction in all channels the value of the inclusive one b-tag channel is chosen, $\kappa = 1.20 \pm 0.04_{\text{stat}}$. The corresponding systematic uncertainty for each channel is taken from Table 3. The sum of the statistical and systematic uncertainty on κ corresponds to the systematic uncertainty for the prediction $\hat{N}_D$.

For the comparison with simulation, the absolute uncertainties for the signal and SM background and the scale factors between data and simulation have to be taken into account. These scale factors correct for the differences in the lepton identification efficiency, trigger efficiencies, b-tag efficiency and pileup. A combination of these scale factors is typically within 10% from unity per event.

In addition, the impact of model uncertainties has been taken into account. For the dominant background, $\mathrm{t\bar{t}}$, the uncertainties for the total cross section are determined using MCFM [32]. The uncertainties associated with scales are determined by varying both factorisation and renormalisation scale by a factor of 2 up and down. Including PDF uncertainties we apply a total uncertainty of 16%.

The uncertainties for SM simulation in the signal region D, shown in Table 4, are needed for the

Table 3: Overview of the uncertainties on the correlation factor κ for the different selections. All systematic uncertainties are added in quadrature. The variation in JES, JER, p_{T} lepton and unclustered energy is propagated to the $\not{E}_{\mathrm{T}}$. The statistical uncertainty in simulation is relatively small, as the b-tagging is applied in the simulation by event weights.

Variation	$\Delta\kappa$ (0 b-tags)	$\Delta\kappa$ (1 b-tag)	$\Delta\kappa$ (2 b-tags)	$\Delta\kappa$ ($\geq$ 3 b-tags)	$\Delta\kappa$ ($\geq$ 1 b-tags)
JES	$\pm$7.5%	$\pm$2.2%	$\pm$1.4%	$\pm$4.0%	$\pm$1.5%
JER	$\pm$4.2%	$\pm$1.7%	$\pm$1.8%	$\pm$5.5%	$\pm$1.1%
p_{T} lepton	$\pm$0.6%	$\pm$1.5%	$\pm$0.7%	$\pm$1.2%	$\pm$0.7%
Uncl. energy	$\pm$3.1%	$\pm$0.3%	$\pm$0.7%	$\pm$0.8%	$\pm$0.4%
Pile-up	$\pm$1.7%	$\pm$0.5%	$\pm$1.1%	$\pm$0.9%	$\pm$0.8%
B-tag SF	$\pm$0.3%	$\pm$0.1%	$\pm$0.1%	$\pm$0.1%	$\pm$0.0%
Mis-tag SF	$\pm$0.0%	$\pm$0.1%	$\pm$0.0%	$\pm$0.1%	$\pm$0.1%
Cross-sect. var.	$\pm$3.4%	$\pm$1.0%	$\pm$2.0%	$\pm$1.4%	$\pm$0.4%
0b-data	$\pm$10.0%	$\pm$10.0%	$\pm$10.0%	$\pm$10.0%	$\pm$10.0%
Total syst. uncert.	$\pm$14.1%	$\pm$10.6%	$\pm$10.5%	$\pm$12.3%	$\pm$10.3%
Stat. error	$\pm$11.8%	$\pm$4.9%	$\pm$4.6%	$\pm$6.2%	$\pm$3.3%

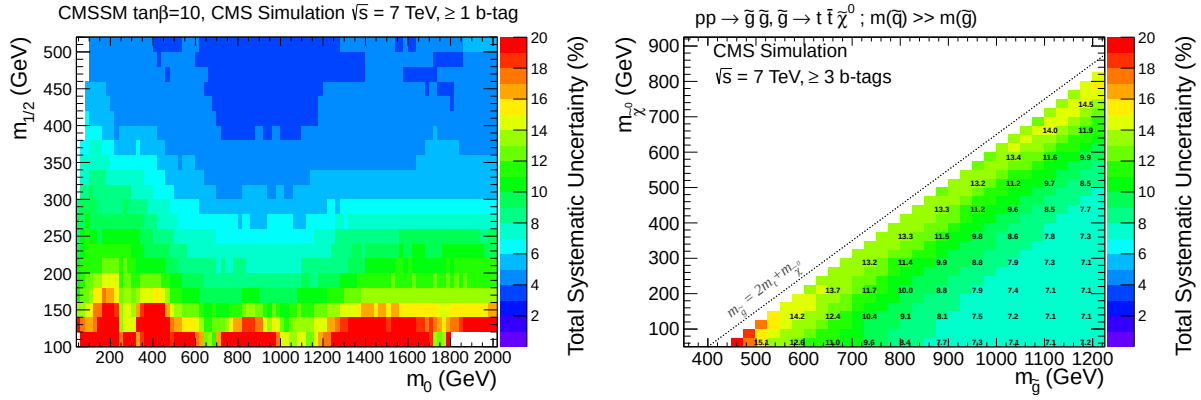
Table 4: Systematics uncertainties in the signal region for the different selections for the SM simulation, needed for the comparison with the SM simulation (as in Table 5). All uncertainties are summed in quadrature. The variation in JES, JER, p_{T} lepton and unclustered energy is propagated to the $\not{E}_{\mathrm{T}}$.

Variation	ΔN_D (0 b-tags)	ΔN_D (1 b-tag)	ΔN_D (2 b-tags)	ΔN_D ($\geq$ 3 b-tags)	ΔN_D ($\geq$ 1 b-tags)
JES	$\pm$26.3 %	$\pm$20.9 %	$\pm$17.9 %	$\pm$17.1 %	$\pm$19.6 %
JER	$\pm$7.7 %	$\pm$6.1 %	$\pm$7.0 %	$\pm$9.5 %	$\pm$6.7 %
p_{T} lepton	$\pm$2.3 %	$\pm$1.8 %	$\pm$2.0 %	$\pm$2.2 %	$\pm$1.9 %
Uncl. energy	$\pm$2.1 %	$\pm$0.3 %	$\pm$0.3 %	$\pm$0.3 %	$\pm$0.3 %
Pile-up	$\pm$0.2 %	$\pm$0.7 %	$\pm$0.3 %	$\pm$0.2 %	$\pm$0.5 %
B-tag SF	$\pm$2.3 %	$\pm$0.9 %	$\pm$3.8 %	$\pm$7.4 %	$\pm$1.5 %
Mis-tag SF	$\pm$1.9 %	$\pm$0.6 %	$\pm$1.1 %	$\pm$5.7 %	$\pm$1.2 %
Model uncert.	$\pm$16.0 %	$\pm$16.0 %	$\pm$16.0 %	$\pm$16.0 %	$\pm$16.0 %
Lep. trig. & ID	$\pm$3.0 %	$\pm$3.0 %	$\pm$3.0 %	$\pm$3.0 %	$\pm$3.0 %
Lumi. uncert.	$\pm$2.2 %	$\pm$2.2 %	$\pm$2.2 %	$\pm$2.2 %	$\pm$2.2 %
Total uncert.	$\pm$32.2 %	$\pm$27.3 %	$\pm$25.7 %	$\pm$27.3 %	$\pm$26.6 %
Stat. error	$\pm$8.4 %	$\pm$3.4 %	$\pm$3.1 %	$\pm$4.3 %	$\pm$2.2 %

comparison with the SM simulation (as in Table 5), but are not used in the limit determination with the scans.

The systematic uncertainties for the signal have been determined for each point of the scans, as they depend on the signal parameters. They are shown in Fig. 5 (a) for the cMSSM mass scan with $\tan\beta = 10$ for the inclusive one b-tag selection and in Fig. 5 (b) for the simplified model "T1tttt" with the inclusive three b-tag selection. The shape of the total systematic uncertainty

distributions are dominantly driven by the uncertainty in the jet energy corrections, which depend upon the mass scale of the signal process.



(a) (b)

Figure 5: Systematic uncertainties for the cMSSM scan with $\tan\beta = 10$ (a) with the inclusive one b-tag selection and for the simplified model "T1tttt" with the inclusive three b-tag selection (b).

7 Results

The number of events in the signal region N_D and the predicted value $\hat{N}_D$ are shown in Fig. 6. The measured number of events N_D and the predicted value $\hat{N}_D$ are in good agreement; we do not observe an excess. The small systematic uncertainties demonstrate nicely the advantage of the data-driven background estimation, while for the comparison of data and simulation, several scale factors have been taken into account, as described in Section 6. The SM simulation is able to describe the measurements. The measured number of events in bin D and the predicted value $\hat{N}_D$ are in good agreement also for the SM simulation. This demonstrates the validity of the factorization ansatz for the background estimation.

The information displayed in Fig. 6 is also given in Table 5, which additionally includes expectations for several possible signal scenarios.

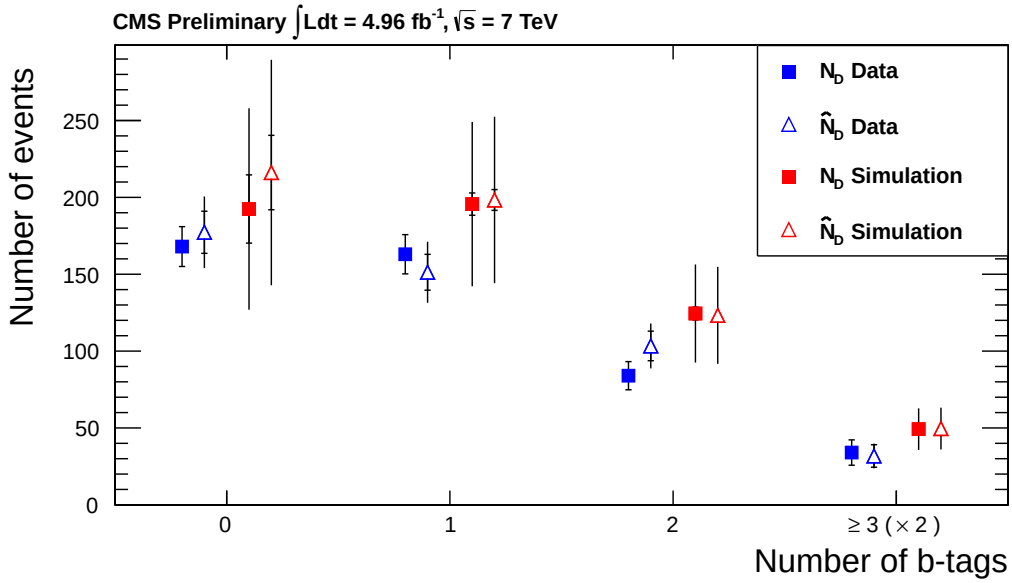


Figure 6: The number of events N_D and the predicted number of events $\hat{N}_D$ in signal region D for data and simulation. In data the systematic uncertainty on $\hat{N}_D$ is given by the uncertainty on κ only and thereby benefits from the data-driven background estimation. The systematic uncertainty in simulation includes in addition the uncertainty on the absolute rate of simulated events, see text. The exclusive 0 b-tag selection is shown for comparison as well. The number of events in the $\geq$ 3-btag channel have been multiplied by a factor of two in order to be better visible on this plot.

Table 5: Number of observed (N_D) vs predicted ($\hat{N}_D$) events in the signal region with statistical and systematic uncertainties for SM, several possible signal scenarios and data.

b-tags	**Sample**	N_D	$\hat{N}_D$
0	Σ SM	$192.4 \pm 22.2 \pm 61.6$	$216.2 \pm 24.2 \pm 75.6$
	Σ SM+LM3	$295.1 \pm 25.5 \pm 87.6$	$252.7 \pm 27.0 \pm 83.0$
	Σ SM+LM8	$219.8 \pm 23.3 \pm 69.0$	$223.7 \pm 24.8 \pm 77.0$
	Σ SM+LM13	$263.4 \pm 25.1 \pm 79.7$	$248.0 \pm 27.0 \pm 82.7$
	Data	168	$177.3 \pm 13.7 \pm 24.9$
1	Σ SM	$195.6 \pm 7.2 \pm 53.0$	$198.3 \pm 6.8 \pm 57.6$
	Σ SM+LM3	$340.3 \pm 9.6 \pm 81.1$	$258.6 \pm 8.3 \pm 67.4$
	Σ SM+LM8	$257.5 \pm 8.0 \pm 67.3$	$214.4 \pm 7.1 \pm 60.4$
	Σ SM+LM13	$337.4 \pm 9.4 \pm 82.4$	$265.9 \pm 8.6 \pm 70.8$
	Data	163	$151.3 \pm 11.7 \pm 16.0$
2	Σ SM	$124.5 \pm 4.2 \pm 31.6$	$123.2 \pm 3.8 \pm 33.9$
	Σ SM+LM3	$218.9 \pm 6.2 \pm 47.1$	$165.0 \pm 4.9 \pm 39.6$
	Σ SM+LM8	$179.2 \pm 5.2 \pm 43.3$	$138.2 \pm 4.1 \pm 36.5$
	Σ SM+LM13	$242.8 \pm 6.7 \pm 54.1$	$184.2 \pm 5.5 \pm 45.4$
	Data	84	$103.4 \pm 9.6 \pm 10.9$
≥ 3	Σ SM	$24.6 \pm 1.1 \pm 6.7$	$24.8 \pm 1.1 \pm 7.4$
	Σ SM+LM3	$67.1 \pm 3.4 \pm 13.2$	$48.0 \pm 2.6 \pm 11.1$
	Σ SM+LM8	$58.9 \pm 2.9 \pm 13.7$	$35.5 \pm 1.7 \pm 9.3$
	Σ SM+LM13	$68.1 \pm 3.2 \pm 14.1$	$59.0 \pm 3.3 \pm 14.2$
	Data	17	$15.9 \pm 3.7 \pm 1.9$
≥ 1	Σ SM	$344.7 \pm 8.4 \pm 90.7$	$346.9 \pm 7.8 \pm 98.0$
	Σ SM+LM3	$626.4 \pm 11.9 \pm 139.7$	$470.2 \pm 10.0 \pm 115.4$
	Σ SM+LM8	$495.7 \pm 10.0 \pm 124.0$	$388.7 \pm 8.4 \pm 105.1$
	Σ SM+LM13	$648.3 \pm 12.0 \pm 149.5$	$505.6 \pm 10.7 \pm 127.6$
	Data	264	$272.2 \pm 15.6 \pm 27.9$

8 Interpretation

Using the results presented in Section 7, limits are set upon the parameters of several supersymmetric models, including the constrained MSSM and the "T1tttt" simplified model (SMS) as described in Section 3.

Limits are set using the CL_s method [33] with a test statistic given by a profile likelihood ratio. The likelihood $\mathcal{L}$ is built to model the important aspects of the background estimation method. For brevity it is not presented in its full form, but the most important components are given:

$$\mathcal{L} = P(\kappa, \epsilon, \mathcal{A}, L)\Pi_{j=A}^{D} Pois(N_j | b_j + s\epsilon_j \mathcal{A}_j \sigma L) \ , \tag{2}$$

where L is the integrated luminosity, σ is the cross-section for the production of supersymmetric states, and s is a signal strength parameter. Each bin is modelled by a Poisson distribution $Pois(N|\langle N \rangle)$. The N_i, $\epsilon_i \mathcal{A}_i$, and b_i represent, for each control and signal region, the number of data events, the acceptance times efficiency for detecting the supersymmetric events (as calculated in simulation), and the expected SM contribution, respectively. The SM contributions b_i are nuisance parameters that are determined by maximizing the likelihood for each model, subject to the constraint:

$$b_D = \kappa \frac{b_C b_B}{b_A} \ . \tag{3}$$

In this way, the possible contamination of the regions A, B and C by SUSY events is taken into account. The factor $P(\kappa, \epsilon, \mathcal{A}, L)$ in Eq. 2 represents the additional constrains due to systematic uncertainties on nuisance parameters. For all systematic uncertainties, the correlation between all four regions are taken into account in the likelihood model and are modeled separately for each point in the scan. The primary sources of the systematic uncertainties are discussed in the context of each model below.

8.1 Constrained MSSM

Within the cMSSM limits are set upon two of the five possible parameters: the scalar mass, m_0, and gaugino mass, $m_{1/2}$. The remaining three parameters are fixed to the following set of values: $\{\tan\beta = 10, A_0 = 0\,\text{GeV}, \mu > 0\}$.

For the cMSSM limits the inclusive one b-tag bin is used. The acceptance and efficiency factors $\epsilon_i \mathcal{A}_i$ are calculated in a scan over the parameters m_0 and $m_{1/2}$. This is done with LO simulation, combined with NLO k-factors which are calculated separately for each SUSY sub-process. The systematic uncertainties from sources discussed in Section 6 are evaluated for each point. The dominant contributions to these systematic uncertainties are found to arise from the JES, which has a 5 – 15% effect depending upon the point in the m_0–$m_{1/2}$ plane, and the uncertainty in the tagging of b-jets, which contributes $\sim 5\%$. Models with larger masses were generically found to have a smaller JES uncertainty.

The 95% CL limit using the CL_s technique obtained for the model with $\tan\beta = 10$ is presented in Fig. 7. Theoretical uncertainties on cross-sections, arising from scale and PDF uncertainties are illustrated by bands of the expected and observed limits with these uncertainties added or subtracted.

8.2 Simplified models

In simplified models a limited set of hypothetical particles are introduced to produce a given topological signature. Two simplified models which produce final states containing a lepton and b-jets as described in Section 3 are studied. Each model contains two free parameters: the

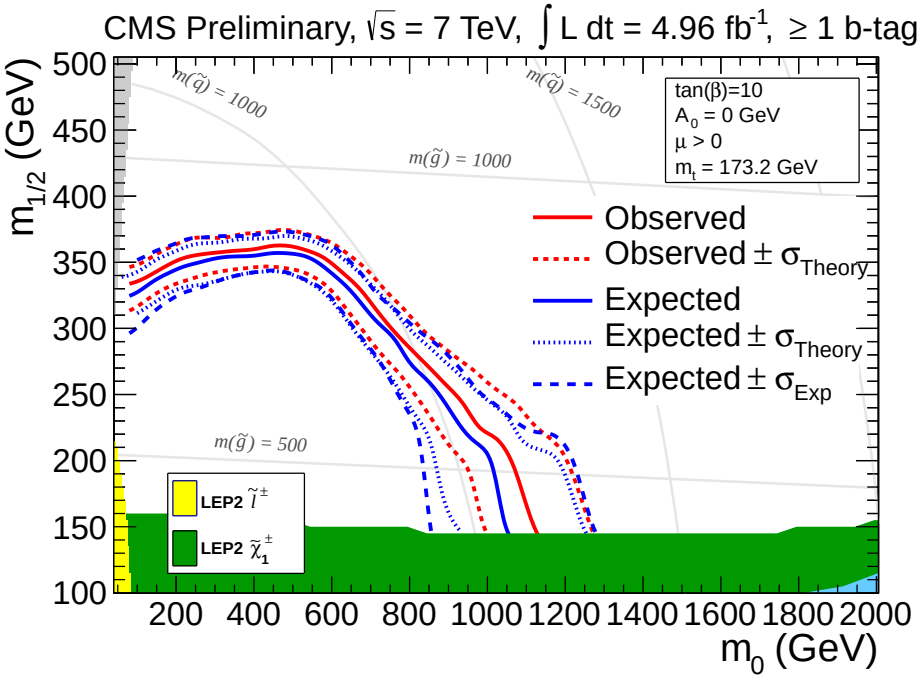


Figure 7: The 95% CL limit using the CL_s technique for the CMSSM model with $\tan\beta = 10$, $A_0 = 0$ GeV and $\mu > 0$. The area below the solid red line is excluded.

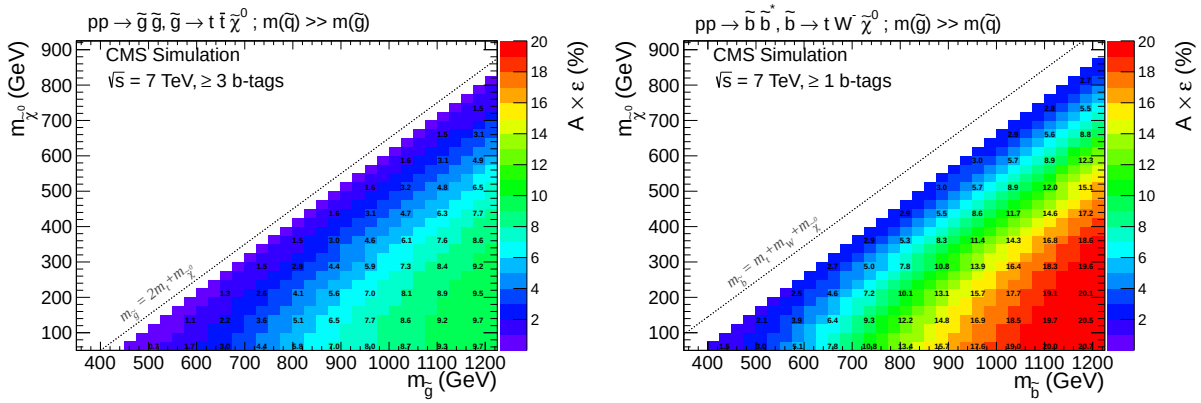
mass of the LSP $m_{\tilde{\chi}^0}$ and the mass of the particle that is initially produced. For each point in the parameter plane, the acceptance times efficiency and a cross-section upper-limit is calculated. The systematic uncertainties are, as in the cMSSM case, determined for each point.

The acceptance times the efficiency for the "T1tttt" and the "T2ttww" model are presented in Fig. 8. For the "T2ttww" model the inclusive one b-tag (n$\geq$ 1) selection is used. Despite the better $\mathcal{A}_i\epsilon_i$ factors compared to the "T1tttt" model, it is not possible to exclude any point, as the cross-section for direct squark production is much lower than for gluino production. For the "T1tttt" model the inclusive three b-tag selection (n$\geq$ 3) is used. In this case acceptance times efficiency in the control regions is near the exclusion region below 3%. The 95% CL upper-limit on the cross-section is presented in Fig. 9.

9 Conclusion

A search for light 3$^{\text{rd}}$ generation squarks and light gluinos has been performed on data collected in 2011 by the CMS experiment in proton-proton collisions at a center-of-mass energy of 7 TeV, corresponding to an integrated luminosity of 4.96 fb^{-1}. The search is based on a selection including a single lepton, missing transverse energy and b-tagged jets.

No deviation from the SM has been found and upper limits have been set on production cross-sections in the cMSSM model, and in the "T1tttt" and "T2ttww" simplified models. Regions of parameter space of the cMSSM and T1tttt models are excluded based on these cross-section limits. Upper limits on the production cross-section of the "T2ttww" do not yet constrain the parameter space of this model.



(a) (b)

Figure 8: Acceptance times efficiency for two simplified models. For the "T1tttt" model (a) the inclusive three b-tag selection is used and for the "T2ttww" model (b) the inclusive one b-tag selection.

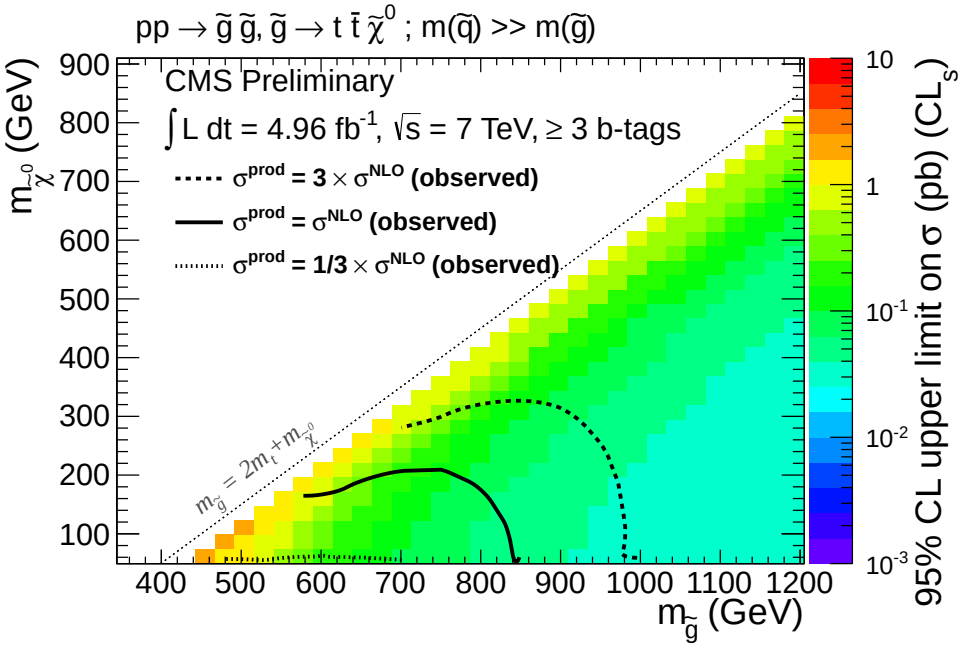


Figure 9: The 95% cross-section upper limit using the CL_s technique for the simplified model "T1tttt" for the inclusive three b-tag selection. The area below the solid line is excluded.

References

[1] S. P. Martin, "A Supersymmetry Primer", arXiv:hep-ph/9709356v6.

[2] J. Wess and B. Zumino, "Supergauge transformations in four dimensions", *Nucl. Phys. B* **70** (1974) 39, doi:10.1016/0550-3213(74)90355-1.

[3] H. P. Nilles, "Supersymmetry, Supergravity and Particle Physics", *Phys. Reports* **110** (1984) 1, doi:10.1016/0370-1573(84)90008-5.

[4] H. E. Haber and G. L. Kane, "The Search for Supersymmetry: Probing Physics Beyond the Standard Model", *Phys. Reports* **117** (1987) 75, doi:doi:10.1016/0370-1573(85)90051-1.

[5] R. Barbieri, S. Ferrara, and C. A. Savoy, "Gauge Models with Spontaneously Broken Local Supersymmetry", *Phys. Lett. B* **119** (1982) 343, doi:10.1016/0370-2693(82)90685-2.

[6] S. Dawson, E. Eichten, and C. Quigg, "Search for Supersymmetric Particles in Hadron - Hadron Collisions", *Phys. Rev. D* **31** (1985) 1581, doi:10.1103/PhysRevD.31.1581.

[7] A. H. Chamseddine, R. Arnowitt, and P. Nath, "Locally Supersymmetric Grand Unification", *Phys. Rev. Lett.* **49** (Oct, 1982) 970–974, doi:10.1103/PhysRevLett.49.970.

[8] L. Hall, J. Lykken, and S. Weinberg, "Supergravity as the messenger of supersymmetry breaking", *Phys. Rev. D* **27** (May, 1983) 2359–2378, doi:10.1103/PhysRevD.27.2359.

[9] CMS Collaboration, "The CMS experiment at the CERN LHC", *JINST* **0803** (2008) S08004, doi:10.1088/1748-0221/3/08/S08004.

[10] CMS Collaboration, "Determination of Jet Energy Calibration and Transverse Momentum Resolution in CMS", *JINST* **6** (2011) P11002, doi:10.1088/1748-0221/6/11/P11002, arXiv:1107.4277.

[11] CMS Collaboration, "Particle–Flow Event Reconstruction in CMS and Performance for Jets, Taus, and $E_{\mathrm{T}}^{\mathrm{miss}}$", CMS Physics Analysis Summary CMS-PAS-PFT-09-001, (2009).

[12] 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).

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

[14] J. Alwall et al., "MadGraph/MadEvent v4: The New Web Generation", *JHEP* **09** (2007) 028, doi:10.1088/1126-6708/2007/09/028, arXiv:0706.2334.

[15] T. Sjostrand, S. Mrenna, and P. Z. Skands, "PYTHIA 6.4 Physics and Manual", *JHEP* **05** (2006) 026, doi:10.1088/1126-6708/2006/05/026, arXiv:hep-ph/0603175.

[16] R. Field, "Early LHC underlying event data - findings and surprises", (2010). arXiv:1010.3558.

[17] Z. Was, "TAUOLA the library for tau lepton decay, and KKMC/KORALB/KORALZ/... status report", *Nucl. Phys. Proc. Suppl.* **98** (2001) 96–102, doi:10.1016/S0920-5632(01)01200-2, arXiv:hep-ph/0011305.

[18] S. Alioli, P. Nason, C. Oleari et al., "NLO single-top production matched with shower in POWHEG: s- and t-channel contributions", *JHEP* **0909** (2009) 111, doi:10.1088/1126-6708/2009/09/111,10.1007/JHEP02(2010)011, arXiv:0907.4076.

[19] E. Re, "Single-top Wt-channel production matched with parton showers using the POWHEG method", *Eur.Phys.J.* **C71** (2011) 1547, doi:10.1140/epjc/s10052-011-1547-z, arXiv:1009.2450.

[20] B. Allanach, "SOFTSUSY: a program for calculating supersymmetric spectra", *Comput.Phys.Commun.* **143** (2002) 305–331, doi:10.1016/S0010-4655(01)00460-X, arXiv:hep-ph/0104145.

[21] CMS Collaboration, "CMS technical design report, volume II: Physics performance", *J. Phys. G* **34** (2007) 995, doi:10.1088/0954-3899/34/6/S01.

[22] W. Beenakker, R. Hopker, and M. Spira, "PROSPINO: A Program for the production of supersymmetric particles in next-to-leading order QCD", arXiv:hep-ph/9611232.

[23] N. Arkani-Hamed, P. Schuster, N. Toro et al., "MARMOSET: The Path from LHC Data to the New Standard Model via On-Shell Effective Theories", arXiv:hep-ph/0703088.

[24] CMS Collaboration, "Search for new physics with single-leptons at the LHC", CMS Physics Analysis Summary CMS-PAS-SUS-11-015, (2011).

[25] CMS Collaboration, "Performance of muon identification in pp collisions at $\sqrt{s}$ = 7 TeV", CMS Physics Analysis Summary CMS-PAS-MUO-10-002, (2010).

[26] CMS Collaboration, "Electron Reconstruction and Identification at $\sqrt{s} = 7$ TeV", CMS Physics Analysis Summary CMS-PAS-EGM-10-004, (2010).

[27] M. Cacciari, G. P. Salam, and G. Soyez, "The anti-k_T jet clustering algorithm", *JHEP* **0804** (2008) 063, doi:10.1088/1126-6708/2008/04/063.

[28] CMS Collaboration, "Determination of jet energy calibration and transverse momentum resolution in CMS", CMS Physics Analysis Summary CMS-PAS-JME-11-011, (2011).

[29] CMS Collaboration, "b-Jet Identification in the CMS Experiment", CMS Physics Analysis Summary CMS-PAS-BTV-11-004, (2011).

[30] CMS Collaboration Collaboration, "Search for supersymmetry in pp collisions at sqrt(s)=7 TeV in events with a single lepton, jets, and missing transverse momentum", *JHEP* **1108** (2011) 156, doi:10.1007/JHEP08(2011)156, arXiv:1107.1870.

[31] CMS Collaboration, "Measuring Electron Efficiencies at CMS with Early Data", CMS Physics Analysis Summary CMS-PAS-EGM-07-001, (2007).

[32] J. M. Campbell and R. K. Ellis, "MCFM home page, http://mcfm.fnal.gov",.

[33] A. L. Read, "Presentation of search results: The CL(s) technique", *J.Phys.G* **G28** (2002) 2693–2704, doi:10.1088/0954-3899/28/10/313.