# CMS Physics Analysis Summary

Contact: cms-pag-conveners-susy@cern.ch 2011/11/28

# Search for Physics Beyond the Standard Model in $Z$ + Jets + $E_{\mathrm{T}}^{\mathrm{miss}}$ events at the LHC

The CMS Collaboration


## Abstract

This paper presents an update of a search for Physics beyond the Standard Model in final states with a $Z$ boson, jets and missing transverse energy. The Jet-$Z$ Balance method is used to estimate the total background expected in the signal region from data control samples. The size of the data corresponds to an integrated luminosity of 2.1 fb$^{-1}$. In the absence of any significant excess beyond the SM expectation, upper limits are set in the context of simplified models of supersymmetry, taking into account signal contribution to the control samples.

This paper describes a search for Physics beyond the Standard Model (BSM) in a sample of pp collisions collected by the Compact Muon Solenoid (CMS) detector at the Large Hadron Collider (LHC), at a center-of-mass energy of 7 TeV. The size of the data sample corresponds to 2.1 $\text{fb}^{-1}$. We search for events with final states containing a $Z$ boson, jets and missing energy ($E_{\text{T}}^{\text{miss}}$), where the $Z$ boson decays to electron or muon pairs. This final state is a clean and distinct signature present in many models of BSM Physics, in particular supersymmetry (SUSY) [1, 2].

The most significant Standard Model (SM) backgrounds in this final state are $Z+$jets processes and $\bar{\text{t}}\text{t}$ production. In $Z+$jets events, while the $Z$ boson momentum is accurately measured from its leptonic decay products and there is no intrinsic $E_{\text{T}}^{\text{miss}}$ the imperfect measurement of the jet energy leads to instrumental $E_{\text{T}}^{\text{miss}}$ mimicking signal events. The ability to observe an excess of signal over background therefore relies on the ability to accurately predict the missing energy "tail" of this background, since it has the exact same topology as the signal. The jet-$Z$ balance (JZB) method has been devised to predict the contribution from $Z+$jets events [3]. The flavour symmetry of leptonic $\bar{\text{t}}\text{t}$ decays, as well as dilepton invariant mass side-band regions, are then used to predict the contribution from $\bar{\text{t}}\text{t}$ events.

The Jet-$Z$ Balance (JZB) method has been described in detail elsewhere [4] and a first update was carried out using 191 $\text{pb}^{-1}$ of data collected by the CMS detector in early 2011 [5]. Here we present a new update of the same analysis using a significantly larger data sample. In addition, the present update benefits from improvements to the background estimation methods described below, as well as interpretation in the context of simplified models.

The data were collected with a set of triggers requiring the presence of at least two leptons (electrons or muons) with $p_{\text{T}}$ thresholds between 8 GeV and 17 GeV. We further select events with at least three central jets with $p_{\text{T}} > 30$ GeV and two opposite-sign isolated leptons with $p_{\text{T}} > 20$ GeV. The jets are reconstructed using a particle flow technique [6], which combines information from all subdetectors. Jet clustering is performed using the anti-$k_T$ clustering algorithm [7] with a distance parameter $D = 0.5$.

The dilepton invariant mass distribution in preselected events is shown in Fig. 1. The signal region in the invariant mass distribution is defined as $|m_{\ell\ell} - m_Z| < 20$ GeV.

The JZB variable is defined as the difference between the $p_{\text{T}}$ of the vectorial sum of all jets and the $p_{\text{T}}$ of the $Z$ boson candidate. In order to be inclusive in the selection of the jets, we use $E_{\text{T}}^{\text{miss}}$ measured with a particle flow technique [8] and we define the JZB observable as

$$JZB = \left|\sum_{\text{jets}} \vec{p}_T\right| - \left|\vec{p}_T^{\,(Z)}\right| = \left|-\vec{E}_{\text{T}}^{\text{miss}} - \vec{p}_T^{\,(Z)}\right| - \left|\vec{p}_T^{\,(Z)}\right|, \quad (1)$$

where $\vec{p}_T^{\,(Z)}$ is the transverse momentum of the $Z$ boson candidate.

The JZB distribution of events in the Z peak region in the Monte Carlo simulation is shown in Fig. 2. Monte Carlo simulation was rescaled to describe the probability distribution of multiple interactions observed in data (referred to as "pile-up reweighting"). A low-, a mid-, and a high-JZB search region are defined by requiring JZB $> 50$ GeV, JZB $> 100$ GeV, and JZB $> 150$ GeV, respectively. These requirements efficiently suppress the $Z+$jets background, leaving $\bar{\text{t}}\text{t}$ as the dominant background. As already mentioned, however, the main challenge of this search is the precise determination of the remaining $Z+$jets background, since this background has the exact same topology as the signal and has an instrumental source, while the $\bar{\text{t}}\text{t}$ background is

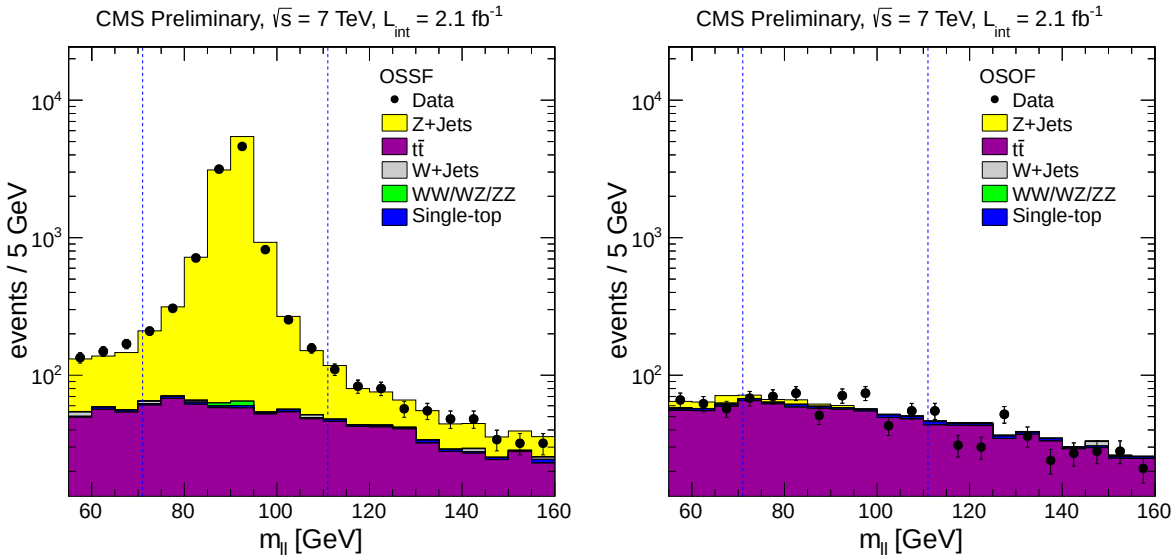


Figure 1: Dilepton invariant mass distribution in preselected events, for same-flavour (left) and opposite-flavour (right) pairs in data (points) and in SM Monte Carlo simulation (histograms). The dashed vertical lines indicate the signal and side-band regions.

almost flat in the invariant mass distribution and stems from physical $E_{\mathrm{T}}^{\mathrm{miss}}$ from neutrinos. The signal region, denoted $\mathrm{JZB}_{\mathrm{pos}}^{\mathrm{SFZP}}$, consists of events with same-flavour dileptons in the Z peak region and JZB $> 0$.

Three control regions are used to predict the contribution of flavour-symmetric backgrounds, which come mainly from $\mathrm{t\bar{t}}$ processes: opposite-flavour events compatible with the $Z$ boson mass hypothesis ($\mathrm{JZB}^{\mathrm{OFZP}}$), opposite-flavour events in the side-band of the $Z$ boson mass ($\mathrm{JZB}^{\mathrm{OFSB}}$), and same-flavour events in the same side-band region ($\mathrm{JZB}^{\mathrm{SFSB}}$). The side-band region is defined as $55\,\mathrm{GeV} < m_{\ell\ell} < 70\,\mathrm{GeV} \cup 112\,\mathrm{GeV} < m_{\ell\ell} < 160\,\mathrm{GeV}$ and was chosen to include the same number of events as the Z-peak region in $\mathrm{t\bar{t}}$ Monte Carlo simulation. This was confirmed in data by comparing the number of events in the region $30\,\mathrm{GeV} < |\mathrm{JZB}| < 50\,\mathrm{GeV}$, in the $\mathrm{JZB}^{\mathrm{OFZP}}$ and $\mathrm{JZB}^{\mathrm{OFSB}}$ control regions. The total contribution from flavour-symmetric backgrounds in the signal region is computed as the mere average of the three control regions, as they provide an independent estimates of the same background process.

The JZB distribution in the signal region and the three control regions are shown in Fig. 3, in data and Monte Carlo simulation of Standard Model events. Note that the signal region and the three JZB control regions are orthogonal to each other: an event can only fall in one of the four regions. The JZB distributions in the three control regions show good agreement above a JZB value of 50 GeV, where $\mathrm{t\bar{t}}$ dominates. The differences observed below 50 GeV are compatible with the expected differences in jet energy resolution. Detailed agreement is however not required by this analysis, since it only relies on control samples in data to estimate the SM background in the signal region.

SM backgrounds with a reconstructed $Z$ boson, mainly coming from $Z+$jets processes, are estimated using the negative JZB region after subtraction of flavour-symmetric backgrounds. This relies on the fact that $Z+$jets events with three or more jets evenly populate the negative and the positive sides of the JZB distribution (see Fig. 2). The total background prediction in the signal region is then computed as:

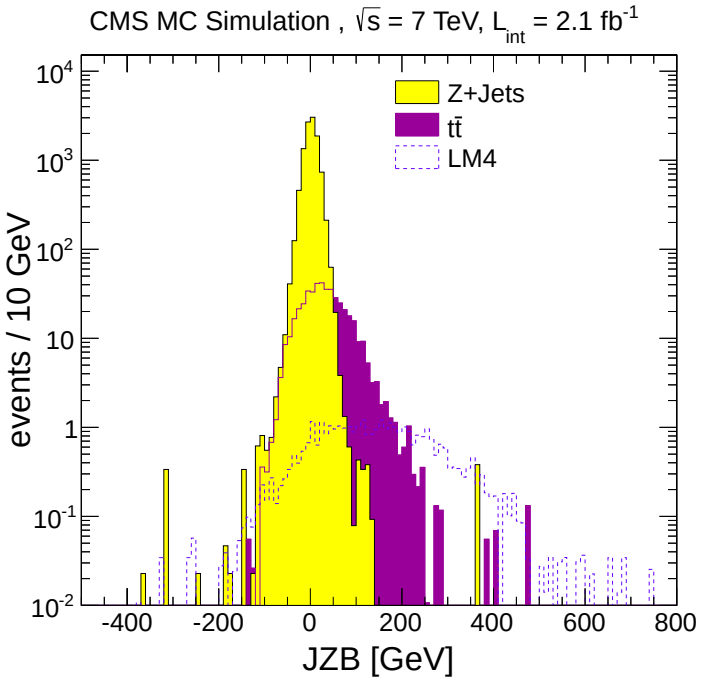


Figure 2: $JZB$ distribution in $Z+$jets, $\mathrm{t\bar{t}}$ and signal Monte Carlo simulation scaled to 2.1 fb$^{-1}$, after preselection and dilepton invariant mass requirement. The signal distribution corresponds to the SUSY LM4 scenario. Events in the tail are affected by the pile-up reweighting of Monte Carlo simulation.

$$JZB^{\mathrm{pred}}_{\mathrm{bkgd}} = |\mathrm{JZB}^{\mathrm{SFZP}}_{\mathrm{neg}}| + \frac{1}{3}(|\mathrm{JZB}^{\mathrm{OFZP}}_{\mathrm{pos}}| - |\mathrm{JZB}^{\mathrm{OFZP}}_{\mathrm{neg}}|) + \frac{1}{3}(|\mathrm{JZB}^{\mathrm{SFSB}}_{\mathrm{pos}}| - |\mathrm{JZB}^{\mathrm{SFSB}}_{\mathrm{neg}}|) + \frac{1}{3}(|\mathrm{JZB}^{\mathrm{OFSB}}_{\mathrm{pos}}| - |\mathrm{JZB}^{\mathrm{OFSB}}_{\mathrm{neg}}|).$$

The differences between the JZB shapes in the control and signal regions are studied in Monte Carlo simulation. The $\mathrm{JZB}^{\mathrm{SFZP}}_{\mathrm{neg}}$ and $\mathrm{JZB}^{\mathrm{SFZP}}_{\mathrm{pos}}$ distributions in $Z+$jets Monte Carlo simulation are found to agree within statistical uncertainties. We assign a systematic error of 25% to the corresponding estimates in order to cover observed discrepancies above 70 GeV, where the Monte Carlo simulation statistics is low. The three control regions for flavour-symmetric background show reasonable agreement in $\mathrm{t\bar{t}}$ Monte Carlo simulation. A systematic uncertainty of 50% is assigned to the corresponding estimates in order to cover the statistical fluctuation of the Monte Carlo simulation in the regions $|\mathrm{JZB}| > 50$ GeV (see Fig. 4), as well as possible shape differences between data and Monte Carlo simulation.

The background estimation methods are validated in Monte Carlo simulation, in a mixture of all SM backgrounds, with and without the inclusion of signal Monte Carlo simulation. The ratio between observed and predicted distributions is shown in Fig. 5 for the two cases. We find that there is very good agreement in the background-only hypothesis, while good sensitivity to a possible signal remains.

The comparison between predicted and observed distributions in data is shown in Fig. 6. The predicted distribution has been fitted with an analytical function in order to display statistical uncertainty bands. The fit is not used in the results. The observed and predicted yields in these

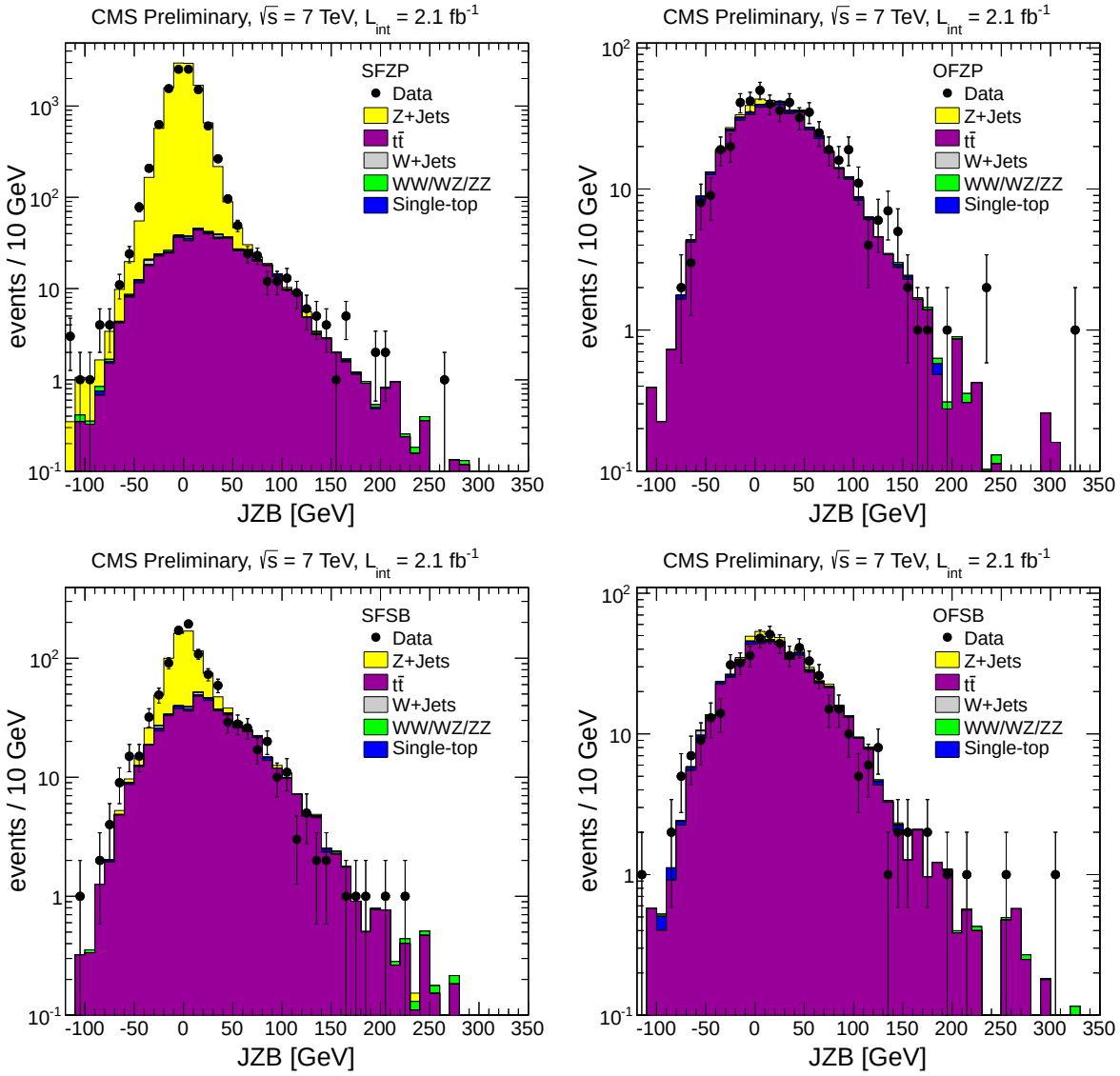


Figure 3: Comparison of the JZB distributions in the same-flavour Z-peak (top left), opposite-flavour Z-peak (top right), same-flavour side-band (bottom left), and opposite-flavour side-band (bottom right) regions, in data and Monte Carlo simulation.

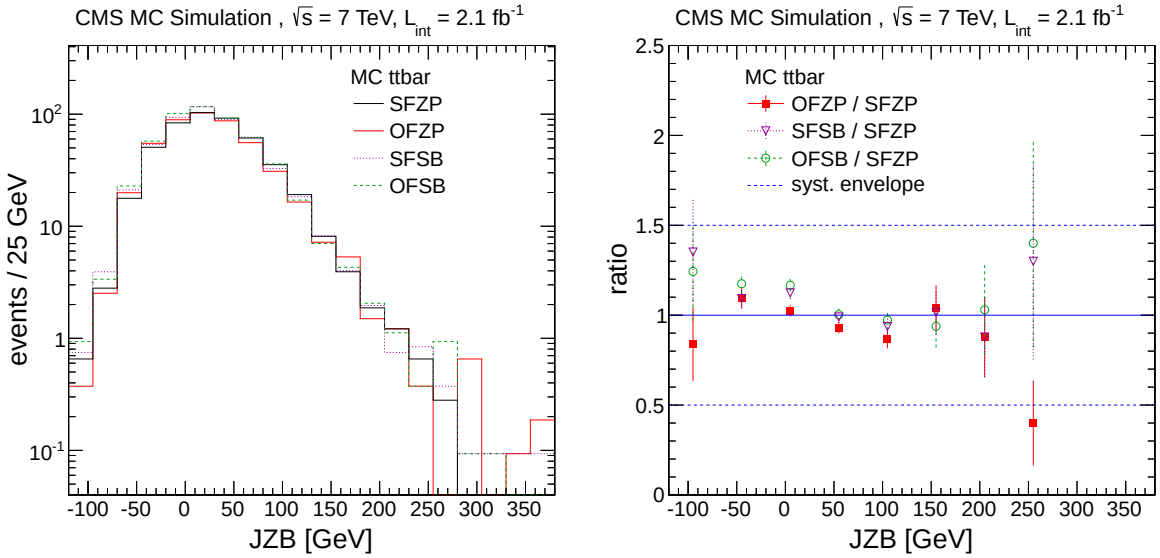


Figure 4: Comparison of the JZB distributions in the signal region and the control regions in $t\bar{t}$ Monte Carlo simulation. The plot on the right shows the ratio between the JZB distribution in each control region and that in the signal region. The assigned systematic error is also shown in this plot.

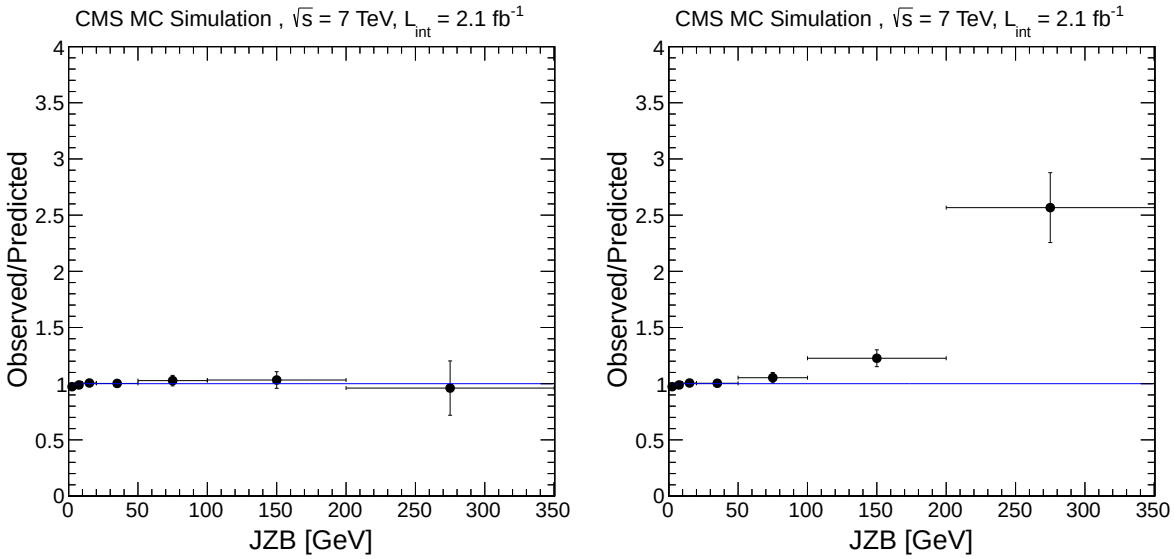


Figure 5: Ratio between observed and predicted JZB distributions in Monte Carlo simulation, without (left) and with (right) the inclusion of signal Monte Carlo simulation. The errors correspond to the statistical uncertainty of the Monte Carlo simulation.

regions are summarised in Tab. 1. As an illustration of the computation of the predicted yields, the background composition of the $JZB > 150\,\text{GeV}$ region is as follows:

$$JZB^{\text{pred}}_{\text{bkgd}} = 0 + \frac{1}{3}(8-0) + (5-0) + (8-0).$$

The systematic uncertainties assigned to the background estimates from the different control regions are considered uncorrelated enough to be added quadratically. The absence of strong correlation is confirmed by the validation of the method in Monte Carlo simulation (Fig. 5), as well as the aforementioned comparison of the number of events in the $30\,\text{GeV} < |JZB| < 50\,\text{GeV}$ region.

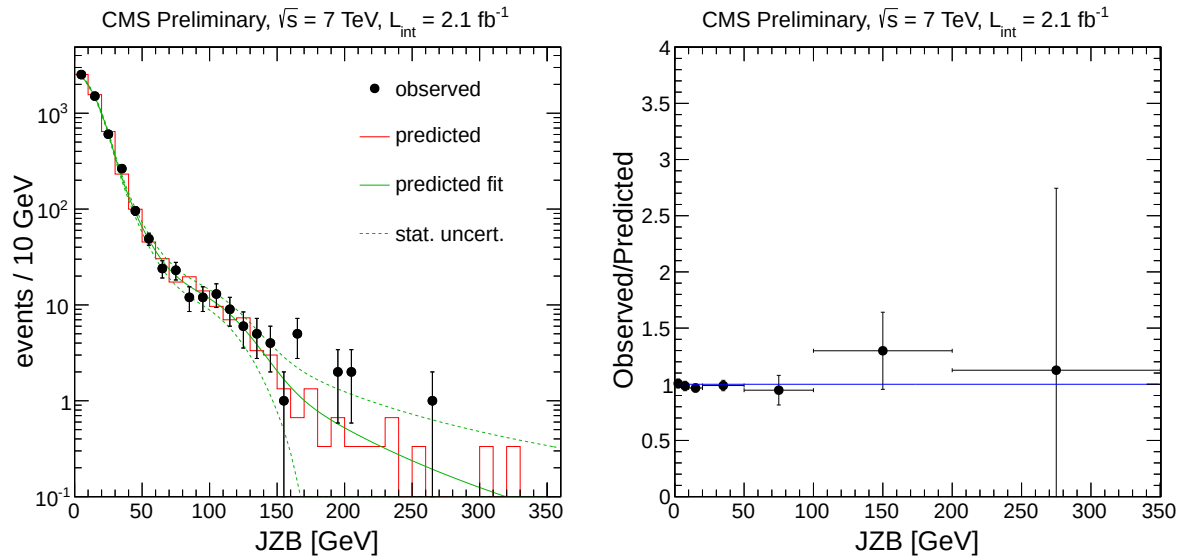


Figure 6: Comparison between the predicted and observed JZB distributions in the search region. The predicted distribution has been fitted to display 1-$\sigma$ error bands. The right plot shows the ratio between the two distributions on the left.

Table 1: Total number of events observed in the search regions, and corresponding background prediction.

| Region | Observed events | Background prediction |
|---|---|---|
| JZB $> 50\,\text{GeV}$ | 168 | $164 \pm 10(\text{stat}) \pm 42(\text{sys})$ |
| JZB $> 100\,\text{GeV}$ | 48 | $37 \pm 4(\text{stat}) \pm 10(\text{sys})$ |
| JZB $> 150\,\text{GeV}$ | 11 | $7.0 \pm 1.5(\text{stat}) \pm 2.1(\text{sys})$ |

In the absence of any significant excess, we use this result to set upper limits on the production cross-section of simplified models of SUSY (SMS) [9, 10]. Figure 7 illustrates the process considered in this study: two gluinos are generated, decaying to jets and a neutralino, which finally decays to a $Z$ boson and the lightest SUSY particle (LSP). The parameters of the model are the masses of the gluino ($m_{\tilde{g}}$) and the LSP ($m_{\tilde{\chi}^0_1}$). The mass of the intermediate neutralino ($m_{\tilde{\chi}^0_2}$) is fixed to the mean of the two other masses: $m_{\tilde{\chi}^0_2} = (m_{\tilde{\chi}^0_1} + m_{\tilde{g}})/2$. In addition, one of the two $Z$ bosons is required to decay leptonically, in electron, muon or tau pairs (the latter actually contributes less than 0.5% of the events to this analysis after the invariant mass requirement).

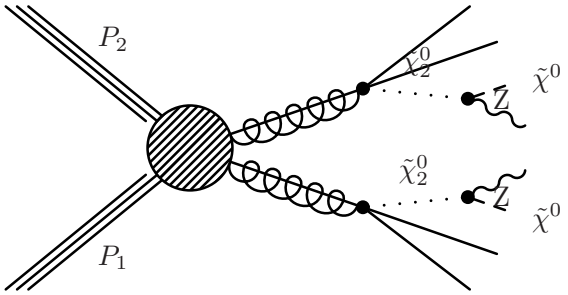


Figure 7: Simplified model for the production of two gluinos decaying into jets and a $Z$ boson.

It should be emphasised that the interpretation of the analysis results in terms of upper limits in the signal cross-section requires special caution when the background is estimated from control regions in data. The signal always contaminates the control region and leads to an over-estimate of the prediction in the signal region. Failure to account for this effect would lead to unrealistic upper limits and exclusion of models that the analysis is actually not sensitive to. The situation is complicated in the SMS case, in which only a specific topology is considered, especially if the background is estimated from an orthogonal control sample. In this analysis, however, the main control region, $\text{JZB}_{\text{neg}}^{\text{SFZP}}$, comes from the same sample and the number of signal events in this control region can simply be subtracted from the number of signal events in the signal region. The signal contamination in the other control regions, used for flavour-symmetric background, would cancel out the corresponding signal in the $\text{JZB}_{\text{pos}}^{\text{SFZP}}$ region. Signal contamination is therefore fully accounted for and the obtained limits are reliable.

Additional systematic uncertainties on the signal efficiency are considered to account for differences between data and Monte Carlo simulation (see Tab. 2). The JZB scale uncertainty corresponds to the $E_{\text{T}}^{\text{miss}}$ scale uncertainty originating from differences in calorimeter response; the JZB resolution uncertainty corresponds to differences in the $E_{\text{T}}^{\text{miss}}$ resolution as observed, *e.g.*, in Fig. 3. The uncertainty on the parton density function (PDF) is also taken into account (following [11]). The model-dependent efficiencies are estimated for each point in the SMS parameter space.

Table 2: Summary of systematic uncertainties on the Monte Carlo simulation signal efficiency. Uncertainties on model-dependent efficiencies are indicated as ranges.

| Source | Uncertainty [%] |
|---|---|
| Trigger efficiency | 5 |
| Lepton selection efficiency | 2 |
| Luminosity | 4.5 |
| Jet energy scale | $3-6$ |
| JZB scale uncertainty | $3-14$ |
| JZB Resolution | $2-15$ |
| PDF uncertainty | $<3$ |

Figure 8 shows the selection efficiency (including acceptance) for the three JZB regions. The 95% CL upper limits on the cross-section for the topology described above are shown in Fig. 9. The upper limits are computed using a hybrid frequentist-bayesian $\text{CL}_{\text{S}}$ method [12]. Here the cross-section is scaled down to 20% to take into account the requirement on one of the $Z$ boson decays. For each parameter point, we calculate the observed and expected upper limits using the results in the three JZB regions, keeping the observed limit corresponding to the best

expected limit. The region with the best sensitivity in each point of the parameter space is also shown in Fig. 9. In order to interpret these limits in terms of gluino pair production cross-section, we use a reference cross-section $\sigma_{\text{ref}}$ and draw the 95% CL exclusion contours at 0.3, 1 and 3 times $\sigma_{\text{ref}}$ in order to give a sense of how the limit changes with the cross-section. $\sigma_{\text{ref}}$ corresponds to gluino pair production in the limit of infinitely heavy squarks, calculated at next to leading order using PROSPINO [13] and CTEQ6 [14] parton distribution functions.

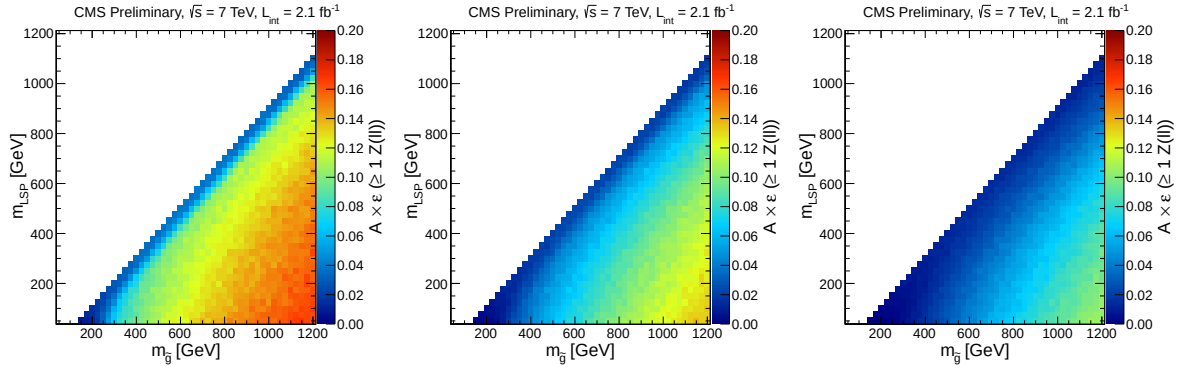


Figure 8: SMS signal efficiency (including acceptance) in the $(m_{\tilde{g}}, m_{\text{LSP}})$ parameter space, for the JZB $> 50$ GeV (left), JZB $> 100$ GeV (centre) and JZB $> 150$ GeV (right) regions.

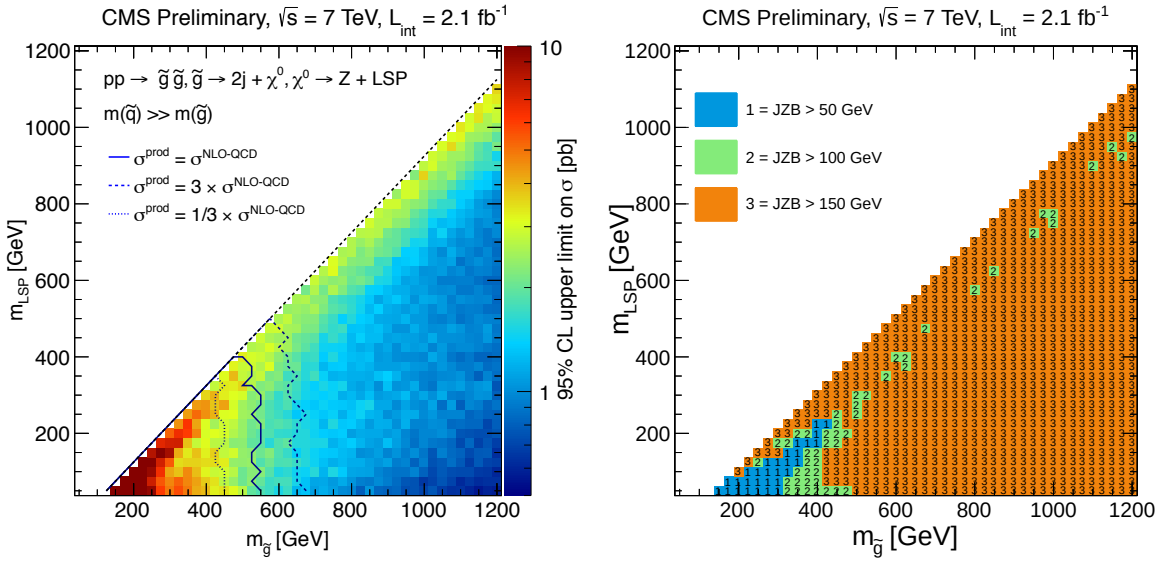


Figure 9: Left: 95% CL upper limits on the cross-section of the inclusive $Z$ boson decay mode in the $(m_{\tilde{g}}, m_{\text{LSP}})$ SMS parameter space. In each point, the best expected limit from the low-, mid-, and high-JZB regions is used. The exclusion curves for reference cross-sections are also shown in this plot. Right: map of the JZB region giving the best limit in each point of the parameter space.

In this simplified model, the JZB analysis is mostly sensitive to topologies in which the $Z$ boson and the LSP have a momentum in the rest frame of the neutralino that is small with respect to the momentum of the neutralino [4]. In the set of simplified models we consider, this topology is realised in the region where the difference between the LSP mass and the gluino mass is small: in this region, the JZB distribution for signal is very asymmetric and signal contamination has little impact on the sensitivity (see Fig. 10). This explains the better upper limit close

to the diagonal in the $(m_{\tilde{g}}, m_{\mathrm{LSP}})$ plane. The loss of sensitivity on the diagonal and in the low mass region are due to the requirement on the jet multiplicity, which translates into a minimum required mass difference between the gluino and the neutralino.

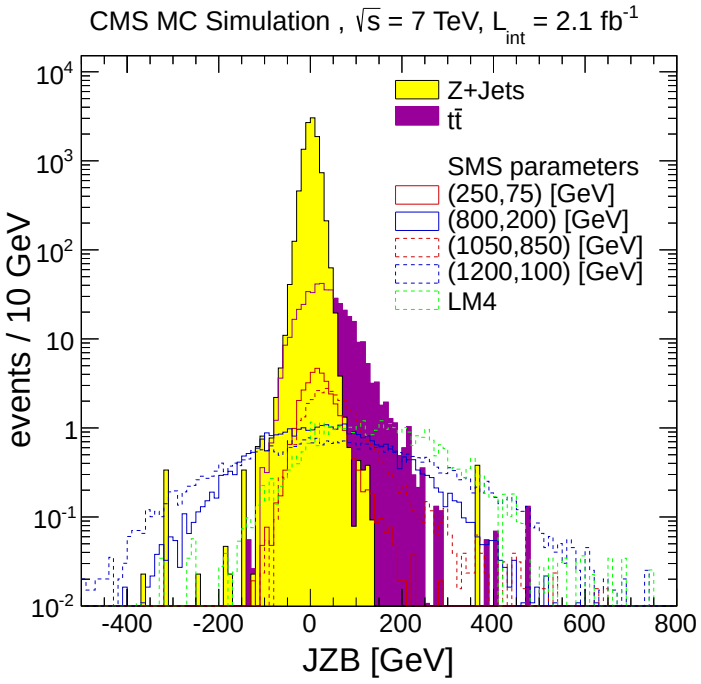


Figure 10: The JZB distribution for the main SM backgrounds, for typical examples of the considered simplified models, and for the LM4 CMSSM benchmark point. The SMS parameters are given in terms of $(m_{\tilde{g}}, m_{\mathrm{LSP}})$, and are scaled to the LM4 cross-section for illustration purpose.

We also estimate upper limits in the context of two CMSSM benchmark scenarios [15], the LM4 and LM8 scenarios (see Tab. 3). LM4 (and LM8) are defined as $m_0 = 210\ (500)$ GeV, $m_{1/2} = 285\ (300)$ GeV, $\tan\beta = 10$, $\mathrm{sign}(\mu) = +$, and $A_0 = 0\ (-300)$ GeV, respectively. In these scenarios, $\tilde{\chi}_2^0$ decays preferentially into a $Z$ boson and the LSP, and it is heavier than the LSP by about 25% of $m_{\tilde{g}} - m_{\mathrm{LSP}}$, in contrast with 50% for the above SMS. The signal contamination is therefore smaller in the LM4 and LM8 scenarios than in the SMS we tested, and the high-JZB region has good sensitivity. The LM4 scenario is excluded at 95% CL by our search.

Table 3: Observed upper limits on the NLO cross-section of LM4 and LM8 benchmark points. The last column indicates the NLO cross-section of the two scenarios.

| Scenario | JZB $>$ 50 GeV | JZB $>$ 100 GeV | JZB $>$ 150 GeV | cross-section |
|---|---|---|---|---|
| LM4 | 7.4 pb | 3.8 pb | 1.9 pb | 2.53 pb |
| LM8 | 7.9 pb | 4.2 pb | 2.0 pb | 1.03 pb |

In summary, we have presented an update of a search for Physics beyond the Standard Model in events with a $Z$, at least three jets and missing energy. The results were interpreted in the context of simplified models of supersymmetry.

## References


[1] K. T. Matchev and S. D. Thomas, "Higgs and $Z$ boson signatures of supersymmetry", *Phys. Rev.* **D62** (2000) 077702. doi:10.1103/PhysRevD.62.077702.

[2] J. T. Ruderman and D. Shih, "General Neutralino NLSPs at the Early LHC", arXiv:1103.6083.

[3] K. Theofilatos, "Supersymmetric particle detection techniques and electromagnetic calorimeter testbeam analysis with the CMS detector". PhD thesis, National Tech. University of Athens and N.C.S.R. Demokritos, 2009.

[4] CMS Collaboration, "Search for Physics Beyond the Standard Model in $Z$ + jets + $E_{\mathrm{T}}^{\mathrm{miss}}$ events at the LHC", *CMS-PAS* **SUS-10-010** (2010).

[5] CMS Collaboration, "Search for Physics Beyond the Standard Model in $Z$ + jets + $E_{\mathrm{T}}^{\mathrm{miss}}$ events at the LHC", *CMS-PAS* **SUS-11-012** (2011).

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

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

[8] CMS collaboration, "Missing transverse energy performance of the CMS detector", *Journal of Instrumentation* **6** (2011), no. 09, P09001. doi:10.1088/1748-0221/6/09/P09001.

[9] D. Alves et al., "Simplified Models for LHC New Physics Searches", arXiv:1105.2838.

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

[11] D. Bourilkov, R. C. Group, and M. R. Whalley, "LHAPDF: PDF use from the Tevatron to the LHC", arXiv:hep-ph/0605240.

[12] K. Nakamura *et al.* (Particle Data Group), "Review of Particle Physics", *J. Phys. G* **37** (2010) 075021. doi:10.1088/0954-3899/37/7A/075021.

[13] W. Beenakker, R. Hoepker, and M. Spira, "PROSPINO: A Program for the Production of Supersymmetric Particles in Next-to-leading Order QCD", 1996.

[14] J. Pumplin et al., "New generation of parton distributions with uncertainties from global QCD analysis", *JHEP* **07** (2002) 012. doi:10.1088/1126-6708/2002/07/012.

[15] CMS Collaboration Collaboration, "CMS technical design report, volume II: Physics performance", *J.Phys.G* **G34** (2007) 995–1579. doi:10.1088/0954-3899/34/6/S01.