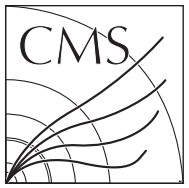
Available on CMS information server

CMS NOTE 2010/000



The Compact Muon Solenoid Experiment

CMS Note

Mailing address: CMS CERN, CH-1211 GENEVA 23, Switzerland

June 21, 2011

A Search For New Physics in Z + Jets + MET using MET Templates

The RA6Z Team

Abstract

We search for new physics in the dilepton final state of Z plus two or more jets plus missing transverse energy (MET). The search is performed using LHC data at a center of mass energy $\sqrt{s} = 7$ TeV recorded by the CMS detector corresponding to an integrated luminosity of 204 pb^{-1}. The Z boson is reconstructed in its decay to e^+e^- or $\mu^+\mu^-$, and the signal regions are defined using requirements of large MET. We use data driven techniques to predict the standard model backgrounds. No evidence for an excess in the signal regions is observed, and the results are used to place upper limits on the non-SM contributions to the yields in the signal regions.

1 Introduction

In this note we describe a search for physics beyond the standard model (BSM) in a sample of proton-proton collisions at a centre-of-mass energy of 7 TeV. The data sample was collected with the Compact Muon Solenoid (CMS) detector [1] at the Large Hadron Collider (LHC) in 2011 and corresponds to an integrated luminosity of 204 pb^{-1}. We search for new physics in events containing opposite sign isolated lepton pairs (ee, $e\mu$, and $\mu\mu$). The main sources of high p_T isolated dileptons at CMS are Drell Yan and $t\bar{t}$. Here we concentrate on dileptons with invariant mass consistent with $Z \to ee$ and $Z \to \mu\mu$. A separate search for new physics in the non-Z sample is described in [2].

We search for new physics in the final state of Z plus two or more jets plus missing transverse energy (MET). We reconstruct the Z boson in its decay to e^+e^- or $\mu^+\mu^-$. Our signal regions are defined as MET > 100 GeV (loose signal region) and MET > 200 GeV (tight signal region). We use data driven techniques to predict the standard model (SM) backgrounds in these search regions. Contributions from Drell-Yan production combined with detector mis-measurements that produce fake MET are modeled via the MET templates method [3, 4] based on control samples of photon plus jets and QCD multijets events. Top pair production backgrounds, as well as other backgrounds for which the lepton flavors are uncorrelated such as W^+W^- and DY$\to \tau^+\tau^-$, are modeled via $e^{\pm}\mu^{\mp}$ subtraction.

As leptonically decaying Z bosons provide a signature that has very little background, they provide a clean final state in which to search for new physics. Because new physics is expected to be connected to the SM Electroweak sector, it is likely that new particles will couple to W and Z bosons. For example, in mSUGRA, low $M_{1/2}$ can lead to a significant branching fraction for $\chi_2^0 \to Z\chi_1^0$. In addition, we are motivated by the existence of dark matter to search for new physics with MET. Enhanced MET is a feature of many new physics scenarios, and R-parity conserving SUSY again provides a popular example. The main challenge of this search is therefore to understand the tail of the fake MET distribution in Z plus jets events.

No specific BSM physics scenario, e.g. a particular SUSY model, has been used to optimize the search. In order to illustrate the sensitivity of the search, a simplified and practical model of SUSY breaking, the constrained minimal supersymmetric extension of the standard model (CMSSM) [5, 6], is used. The CMSSM is described by five parameters: the universal scalar and gaugino mass parameters (m_0 and $m_{1/2}$, respectively), the universal trilinear soft SUSY breaking parameter A_0, the ratio of the vacuum expectation values of the two Higgs doublets ($\tan\beta$), and the sign of the Higgs mixing parameter μ. Throughout the note, two CMSSM parameter sets, referred to as LM4 and LM8 [7], are used to illustrate possible CMSSM yields. In these scenarios, opposite sign leptons are produced in the decays of Z bosons, which are produced in the cascade decays of heavy, colored objects. The parameter values defining LM4 (LM8) are $m_0 = 210\ (500)$ GeV/c^2, $m_{1/2} = 285\ (300)$ GeV/c^2, $A_0 = 0\ (-300)$ GeV; both LM4 and LM8 have $\tan\beta = 10$ and $\mu > 0$. In this analysis, the LM4 and LM8 scenarios serve as benchmarks which may be used to allow comparison of the sensitivity with other analyses.

2 CMS Detector

The central feature of the CMS apparatus 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 field volume are several particle detection systems. Charged particle trajectories are measured by silicon pixel and silicon strip trackers, covering $0 \leq \phi \leq 2\pi$ in azimuth and $|\eta| < 2.5$ in pseudorapidity, defined as $\eta = -\log[\tan\theta/2]$, where θ is the polar angle of the trajectory of the particle with respect to the counterclockwise proton beam direction. A crystal electromagnetic calorimeter and a brass/scintillator hadronic calorimeter surround the tracking volume, providing energy measurements of electrons and hadronic jets. Muons are identified and measured in gas-ionization detectors embedded in the steel return yoke outside the solenoid. The detector is nearly hermetic, allowing energy balance measurements in the plane transverse to the beam direction. A two-tier trigger system selects the most interesting pp collision events for use in physics analysis. A more detailed description of the CMS detector can be found elsewhere [1].

3 Selection

Samples of MC events are used to guide the design of the analysis. These events are generated using either the PYTHIA 6.4.22 [8] or MADGRAPH 4.4.12 [9] event generators. They are then simulated using a GEANT4-based model [10] of the CMS detector, and finally reconstructed and analyzed using the same software as is used to process collision data.

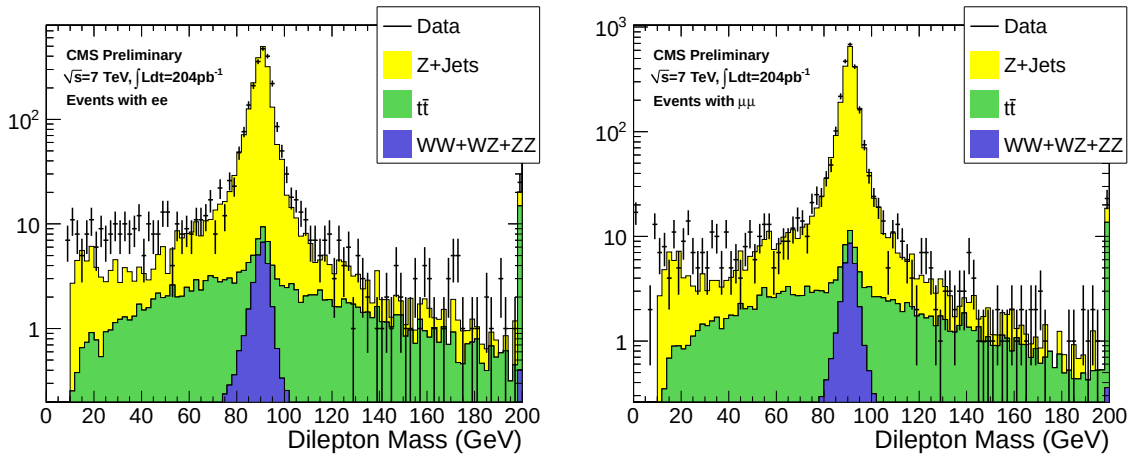


Figure 1: Dilepton mass distribution for events passing the pre-selection in the ee (left) and $\mu\mu$ (right) final states. Backgrounds from single top and W+jets are omitted since they are negligible.

Table 1: Data and MC yields passing preselection. The MC uncertainties are statistical only.

sample	ee	$\mu\mu$	$e\mu$	tot
$Z+$jets	1840.6 ± 21.2	2088.0 ± 22.6	1.5 ± 0.6	3930.1 ± 31.0
$t\bar{t}$	24.5 ± 0.9	26.0 ± 0.9	51.3 ± 1.2	101.8 ± 1.8
$W+$jets	0.9 ± 0.6	0.0 ± 0.0	0.4 ± 0.4	1.3 ± 0.7
W^+W^-	0.2 ± 0.0	0.4 ± 0.1	0.6 ± 0.1	1.3 ± 0.1
$W^\pm Z^0$	13.9 ± 0.2	16.0 ± 0.2	0.1 ± 0.0	30.1 ± 0.2
Z^0Z^0	10.0 ± 0.1	11.4 ± 0.1	0.0 ± 0.0	21.4 ± 0.1
single top	0.7 ± 0.1	0.8 ± 0.1	1.7 ± 0.1	3.2 ± 0.1
total MC	1890.8 ± 21.2	2142.6 ± 22.6	55.6 ± 1.4	4089.0 ± 31.1
data	2051	2277	66	4394

Events with two opposite-sign, isolated leptons (e^+e^-, $e^\pm\mu^\mp$, or $\mu^+\mu^-$) are selected. Both leptons must have $p_T > 20$ GeV/c and the electrons (muons) must have $|\eta| < 2.5$ ($|\eta| < 2.4$). In events with more than two such leptons, the two leptons with invariant mass closest to the Z mass are selected. The leptons are required to be consistent with originating from the Z by requiring the invariant mass to satisfy $81 < \mathrm{M}(\ell\ell) < 101$ GeV/c^2.

Events are required to pass at least one of a set of ee, $e\mu$ or $\mu\mu$ double-lepton triggers. The efficiency for events containing two leptons passing the analysis selection to pass at least one of these triggers is measured to be approximately 100%, 95%, and 90% for ee, $e\mu$ or $\mu\mu$ double-lepton triggers, respectively. In the following, the MC yields are weighted by these trigger efficiencies.

Because leptons produced in the decays of low-mass particles, such as hadrons containing b and c quarks, are nearly always inside jets, they can be suppressed by requiring the leptons to be isolated in space from other particles that carry a substantial amount of transverse momentum. The details of the lepton isolation measurement are given in Ref. [11]. In brief, a cone is constructed of size $\Delta R \equiv \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} = 0.3$ around the lepton momentum direction. The lepton relative isolation is then quantified by summing the transverse energy (as measured in the calorimeters) and the transverse momentum (as measured in the silicon tracker) of all objects within this cone, excluding the lepton, and dividing by the lepton transverse momentum. The resulting quantity is required to be less than 0.15, rejecting the large background arising from QCD production of jets.

We require the presence of at least two jets with $p_T > 30$ GeV/c and $|\eta| < 3.0$, separated by $\Delta R > 0.4$ from leptons passing the analysis selection. The anti-k_T clustering algorithm [12] with $\Delta R = 0.5$ is used for jet clustering. The jets and MET are reconstructed with the Particle Flow technique [13].

The resulting dilepton mass spectra for the ee and $\mu\mu$ final states for events passing the above requirements except for the Z mass requirement are shown in Figure 1. The data and MC yields passing the full preselection above are displayed in Table 1. The MC yields are normalized to 204 pb^{-1}, assuming 100% trigger efficiency. As anticipated, the MC predicts that the preselection is dominated by $Z+$jets in the same-flavor case and by $t\bar{t}$ in the opposite-flavor case. The data yield is in reasonable agreement with the predictions for the ee, $\mu\mu$ and $e\mu$

channels.

4 Definition of the signal regions

We define signal regions to look for possible new physics contributions by adding the requirement of large MET to the preselection.

We define two signal regions for our search:

- MET $>$ 100 GeV (loose signal region):

  In this region there is a small contribution from the tail of the MET distribution in Z plus jets events. The bulk of the events in this region are from $t\bar{t}$ events where the leptons happen to be in the Z mass window.

  The dilepton mass distributions for data and MC events in the loose signal region are displayed in Fig. 2.

- MET $>$ 200 GeV (tight signal region): This signal region was selected by picking a region where the SM expectation is very low. At this kinematical region the dominant background contribution is expected to be from $t\bar{t}$.

In the two signal regions above, the dominant background is $t\bar{t}$. However, it is still essential to have a data driven estimation of the Z contribution in the signal regions, since if we were to observe an excess we would need to demonstrate that this excess is not due to SM Z production accompanied by artificial MET.

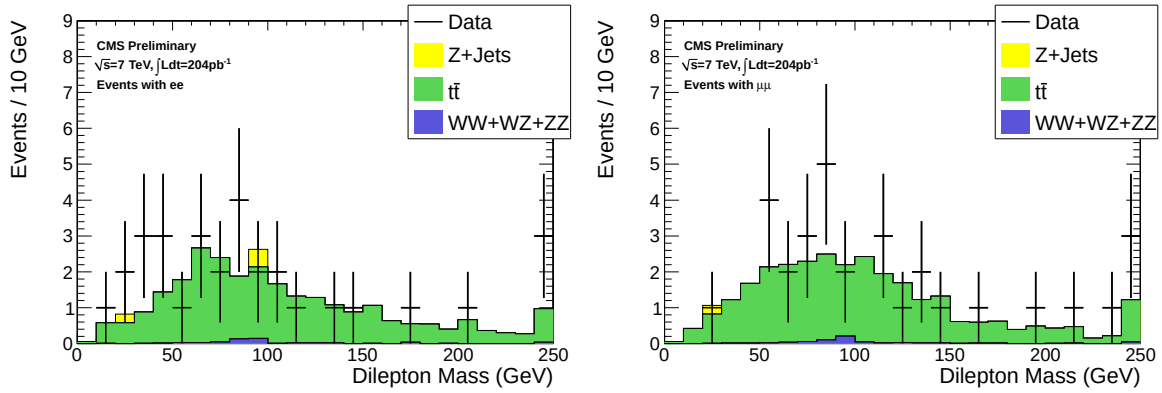


Figure 2: Dilepton mass distribution for events passing the pre-selection and MET $>$ 100 GeV in the ee (left) and $\mu\mu$ (right) final states. Backgrounds from single top and W+jets are omitted since they are negligible.

5 MET Templates

The background from SM Z production accompanied by artificial MET from detector mismeasurements is estimated using the MET templates technique. The premise of this data driven technique is that MET in Z plus jets events is produced by the hadronic recoil system and *not* by the leptons making up the Z . Therefore, the MET distribution in these events can be modeled using a control sample which has no true MET and is accompanied by a similar hadronic system as in Z plus jets events. We use two complementary control samples: one consisting of photon plus jets events, and one consisting of QCD multijet events.

In the Z plus jets events as well as in both control samples, the MET is dominated by mismeasurements of the hadronic system. To account for kinematic differences between the hadronic systems in the control vs. signal samples, the expected MET distribution for a Z plus jets event is obtained from the MET distribution for photon plus jets or QCD multijet events of the same jet multiplicity and scalar sum of jet transverse energies. We use 2 separate control samples, since each has relative advantages. The photon plus jets events have a topology which is more similar to the Z plus jets events, since both consist of a well-measured object recoiling against a system of hadronic jets. Possible contributions of the photon to the MET in the event are eliminated in the QCD multijet sample. We have verified that the predictions from the 2 control samples are consistent within their uncertainties, and choose to use the prediction from the photon plus jets sample. By using two independent control samples, we

are able to illustrate the robustness of the MET templates method and to cross check the data driven background prediction.

The systematic uncertainty in the prediction from the photon plus jets MET templates originates from 2 sources. Possible contributions to the MET from the photon are assessed by varying the photon selection, which leads to a relative difference in the predicted background of 15%. The effect of the difference between the distributions of hadronic recoil p_T in the control vs. signal samples is estimated by reweighting the photon plus jets events such that the hadronic recoil p_T distribution matches that in Z plus jets events, leading to a relative difference in the predicted background of 20%. The total uncertainty in the predicted background from the MET templates method is 25%.

6 Top Background Estimation

The $t\bar{t}$ contribution to the signal region is estimated using an opposite-flavor (OF) subtraction technique. This technique takes advantage of the fact that the $t\bar{t}$ yield in the OF final state ($e\mu$) is the same as in the same-flavor (SF) final state ($ee + \mu\mu$), modulo differences in efficiency in the e vs. μ selection. Hence the $t\bar{t}$ yield in the same-flavor final state can be estimated using the corresponding yield in the opposite-flavor final state. Other backgrounds for which the lepton flavors are uncorrelated (for example, W^+W^- and DY$\rightarrow \tau^+\tau^-$) are also included in this estimate.

To predict the SF yield in a signal region defined by a requirement on the MET, we take the OF yield passing the same MET requirement. This yield is corrected using the ratio of muon to electron selection efficiencies $R_{\mu e} = 1.07 \pm 0.03$. This quantity is evaluated from studies of $Z \rightarrow \mu^+\mu^-$ and $Z \rightarrow e^+e^-$ events in data. To improve the statistical precision of the background estimate, we do not require the OF events to lie in the Z mass region, and we apply a scale factor $K = 0.16 \pm 0.01$ accounting for the fraction of $t\bar{t}$ events which lie in the region $81 < \mathrm{M}(\ell\ell) < 101\,\mathrm{GeV}/c^2$, extracted from MC.

Backgrounds from pair production of vector bosons are negligible compared to $t\bar{t}$. Backgrounds from fake leptons are negligible due to the requirement of two $p_T > 20$ GeV leptons in the Z mass window, accompanied by jets and large MET.

7 Results

The data and SM predictions are shown in Fig. 3 and summarized in Table 2. In the loose signal region (MET $>$ 100 GeV), we observe 14 events (7 ee and 7 $\mu\mu$) compared to a data-driven prediction of 13.2 $\pm$ 1.5 (stat) $\pm$ 0.6 (syst). For the tight signal region (MET $>$ 200 GeV), we observe 2 events (both in the $\mu\mu$ channel) compared to a data-driven prediction of 1.2 $\pm$ 0.5 (stat) $\pm$ 0.04 (syst). We conclude that no evidence for an anomalous event yield above the data-driven prediction is observed.

8 Acceptance and Efficiency Systematic Uncertainties

The acceptance and efficiency, as well as the systematic uncertainties in these quantities, depend on the signal under consideration. For some of the individual uncertainties, it is reasonable to quote values based on SM control samples with kinematic properties similar to the SUSY benchmark models. For others that depend strongly on the kinematic properties of the event, the systematic uncertainties must be quoted model by model.

The systematic uncertainty in the lepton acceptance consists of two parts: the trigger efficiency uncertainty and the identification and isolation uncertainty. The trigger efficiency for two leptons of $p_T > 20\,\mathrm{GeV}/c$ is measured using samples of $Z \rightarrow \ell\ell$, with an uncertainty of 2%. We verify that the MC reproduces the lepton identification and isolation efficiencies in data using samples of $Z \rightarrow \ell\ell$; the data and MC efficiencies are found to be consistent within 2%.

Another significant source of systematic uncertainty is associated with the jet and MET energy scale. The impact of this uncertainty is final-state dependent. Final states characterized by very large MET are less sensitive than final states where the MET is typically close to the minimum requirement applied to this quantitys. To be more quantitative, we have used the method of Ref. [11] to evaluate the systematic uncertainties in the acceptance for $t\bar{t}$ and for the two benchmark SUSY points using a 5% uncertainty in the hadronic energy scale [16]. The signal efficiency uncertainty for the region MET $>$ 100 GeV is 16%, 3% and 4% for $t\bar{t}$, LM4 and LM8, respectively.

The signal efficiency uncertainty for the region MET $>$ 200 GeV is 33%, 8% and 9% for $t\bar{t}$, LM4 and LM8, respectively.

The uncertainty in the integrated luminosity is 4% [17].

9 Limits on New Physics

As discussed in Sec. 7, we do not observe any excess in the signal regions. We use these results to place Bayesian 95% confidence level upper limits [18] on the non-SM contributions to the yields in the signal regions, using a log-normal model of nuissance parameter integration. The results are summarized in Table 2.

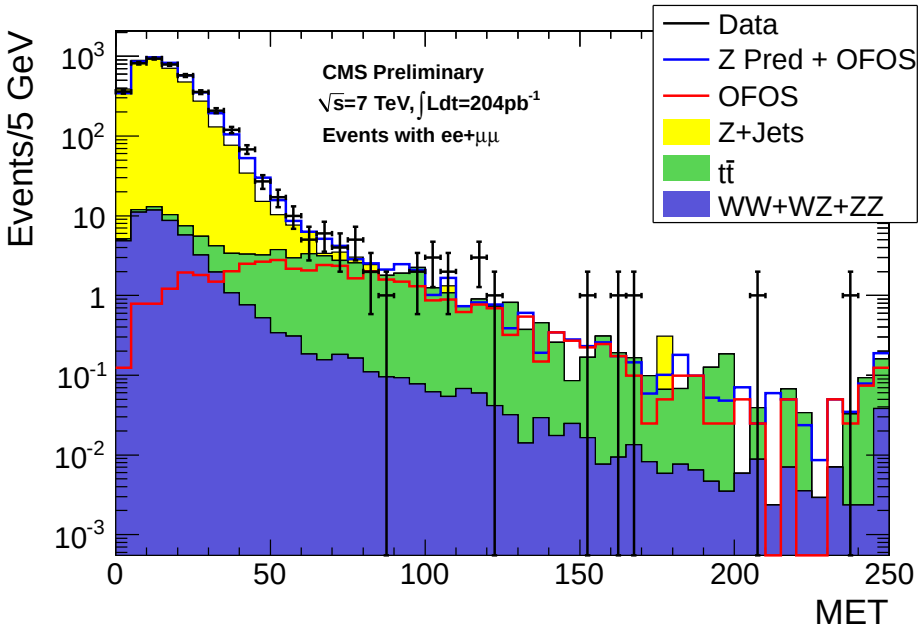


Figure 3: The observed MET distribution for data in the (black points), predicted $t\bar{t}$ MET distribution (red line), the sum of predicted $t\bar{t}$ MET distribution and Z MET distribution predicted from photon MET templates (solid blue line), and MC (solid histograms).

Table 2: Summary of the yields in the regions MET $>$ 30, 60, 100 and 200 GeV. The total predicted yield is the sum of the predicted background from Z plus jets from the MET templates method (Z prediction) plus the $t\bar{t}$ contribution predicted from OF subtraction ($t\bar{t}$ prediction). Here the first uncertainty is statistical, the second uncertainty is systematic. The 95% CL Bayesian UL is indicated, as well as the expected NLO yields for the LM4 and LM8 scenarios, including the uncertainties from lepton identification and isolation efficiency, trigger efficiency, hadronic energy scale, and integrated luminosity.

	N(MET $>$ 30) GeV	N(MET $>$ 60) GeV	N(MET $>$ 100) GeV	N(MET $>$ 200) GeV
Z prediction	$406.2 \pm 7.1 \pm 101.6$	$13.1 \pm 1.2 \pm 3.3$	$1.4 \pm 0.6 \pm 0.4$	$0.05 \pm 0.02 \pm 0.01$
$t\bar{t}$ prediction	$54.8 \pm 3.0 \pm 2.0$	$34.7 \pm 2.4 \pm 1.3$	$11.8 \pm 1.4 \pm 0.4$	$1.1 \pm 0.5 \pm 0.04$
total prediction	$461.0 \pm 7.7 \pm 101.6$	$47.9 \pm 2.7 \pm 3.5$	$13.2 \pm 1.5 \pm 0.6$	$1.2 \pm 0.5 \pm 0.04$
observed	488	39	14	2
UL	179	11.5	10.0	5.3
LM4	5.6 ± 0.3	5.0 ± 0.3	4.4 ± 0.3	2.7 ± 0.3
LM8	2.6 ± 0.2	2.4 ± 0.2	1.9 ± 0.1	1.1 ± 0.1

10 Additional Information for Model Testing

UPDATE NUMBERS BELOW

Other models of new physics in the dilepton final state can be confronted in an approximate way by simple

generator-level studies that compare the expected number of events in 204 pb^{-1}with the upper limits from Section 9. The key ingredients of such studies are the kinematic requirements described in this note, the lepton efficiencies, and the detector responses for and $\not\!\!E_T$. The muon identification efficiency is $\approx 96\%$; the electron identification efficiency varies approximately linearly from $\approx 60\%$ at $p_T = 10\,\text{GeV}/c$ to 90% for $p_T > 30\,\text{GeV}/c$. The lepton isolation efficiency depends on the lepton momentum, as well as on the jet activity in the event. In $t\bar{t}$ events, it varies approximately linearly from $\approx 73\%$ (muons) and $\approx 82\%$ (electrons) at $p_T = 10\,\text{GeV}/c$ to $\approx 97\%$ for $p_T > 60\,\text{GeV}/c$. In LM4 (LM8) events, this efficiency is decreased by ≈X% (≈X%)over the whole momentum spectrum. The average detector responses (the reconstructed quantity divided by the generated quantity) for $\not\!\!E_T$ is consistent with 1 within the 5% jet energy scale uncertainty. The experimental resolutions on this quantity is X%.

11 Conclusion

We have performed a search for BSM physics in the Z + jets + MET final state. Backgrounds from SM Z production were estimated using the data-driven MET templates method, and backgrounds from $t\bar{t}$ were estimated using the data-driven opposite-flavor subtraction technique. We found no evidence for anomalous yield beyond SM expectations and placed Bayesian 95% CL upper limits on the non SM yields in the signal regions. We also provided information on the detector response and efficiencies to allow testing whether specific models of new physics are excluded by our results.

References

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

[2] arXiv:1103.1348v1 [hep-ex], "Search for Physics Beyond the Standard Model in Opposite-Sign Dilepton Events at $\sqrt{s} = 7$ TeV."

[3] V. Pavlunin, Phys. Rev. **D81**, 035005 (2010).

[4] V. Pavlunin, CMS AN-2009/125

[5] G. Kane, "Study of constrained minimal supersymmetry"", *Phys. Rev. D* **49** (1994), no. 11, 6173–6210. doi:10.1103/PhysRevD.49.6173.

[6] 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.

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

[8] T. Sjöstrand, S. Mrenna, and P. Skands, "PYTHIA 6.4 Physics and Manual", *JHEP* **0605** (2006) 026, arXiv:hep-ph/0603175.

[9] J. Alwall, P. Demin, S. de Visscher et al., "MadGraph/MadEvent v4: The New Web Generation", *JHEP* **0709** (2007) 028, arXiv:0706.2334.

[10] **GEANT4** Collaboration, S. Agostinelli et al., "GEANT4: A simulation toolkit", *Nucl. Instrum. Meth.* **A506** (2003) 250–303. doi:10.1016/S0168-9002(03)01368-8.

[11] **CMS** Collaboration, V. Khachatryan et al., "First Measurement of the Cross Section for Top-Quark Pair Production in Proton-Proton Collisions at sqrt(s)=7 TeV", *Phys.Lett.* **B695** (2011) 424–443. doi:10.1016/j.physletb.2010.11.058.

[12] 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.

[13] 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).

https://twiki.cern.ch/twiki/bin/view/CMS/ProductionSpring2011

[14] CMS Collaboration, "Jet performance in pp collisions at $\sqrt{s} = 7$ TeV", *CMS-PAS* **JME 10-003** (2010).

[15] CMS Collaboration, "CMS MET Performance in Events Containing Electroweak Bosons from pp Collisions at sqrt(s)=7 TeV", *CMS-PAS* **JME 10-005** (2010).

[16] CMS Collaboration, "Jet Energy Corrections determination at $\sqrt{s} = 7$ TeV", *CMS-PAS* **JME 10-010** (2010).

[17] CMS Collaboration, "Measurement of CMS luminosity", *CMS-PAS* **EWK-10-004** (2010).

[18] G. Landsberg, https://twiki.cern.ch/twiki/pub/CMS/EXOTICA/cl95cms.c