SUSY sensitivity and discovery potential using dileptons at $\sqrt{s} = 7$ TeV

D. Barge, C. Campagnari, P. Kalavase, D. Kovalskyi, V. Krutelyov, J. Ribnik

University of California, Santa Barbara

W. Andrews, D. Evans, F. Golf, J. Mülmenstädt, S. Padhi, Y. Tu, F. Würthwein, A. Yagil

University of California, San Diego

L. Bauerdick, I. Bloch, K. Burkett, I. Fisk, O. Gutsche, B. Hooberman

Fermi National Accelerator Laboratory, Batavia, Illinois

Abstract

SUSY sensitivity and discovery discovery potential using dileptons at $\sqrt{s} = 7$ TeV is presented in this note. The mass reach as well as CMS sensitivity to several regions in parameter space is evaluated within the framework of minimal supergravity and assuming R-parity conservation. These studies are performed for inclusive searches involving same and opposite sign dileptons. The new physics is characterized by large $\not{E}_T$ and significant hadronic activity. The study shows significant sensitivity in several regions in the parameters space with 100 pb^{-1} and 1 fb^{-1} of integrated luminosity.

1 Introduction

In this note, we present the estimates of CMS sensitivity to SUSY for 95% C.L. limits as well as 5σ discover reach in the mSUGRA framework with R-parity conservation at a proton proton center of mass of 7TeV. This model is characterized by five free parameters described as follows:

- m_0: the common scalar mass at the GUT scale;
- $m_{1/2}$: the common gaugino mass at the GUT scale;
- A_0: the common soft trilinear SUSY breaking parameter at the GUT scale;
- tanβ: the ratio of the Higgs vacuum expectation values at the electroweak scale;
- sign μ: the sign of the Higgsino mass term.

We set $A_0 = 0$, sign$\mu > 0$ and tan$\beta = 3$ in order to be able to directly compare with the recent Tevatron results[**?**][**?**]. The gluino-squark mass plane is then scanned via variations of m_0 and $m_{1/2}$ parameters. In this framework, all supersymmetric particles except the neutralino are unstable and thus will decay into their SM counterparts right after being produced. This cascade decay will result in dilepton final states associated with several jets, plus missing transverse energy ($\not\!\!E_T$) from the LSP.

The note is organized as follows. In Section 2 we list the Monte Carlo data samples, as well as the software tags used in this analysis. In Section 3 we describe the same (SS) and opposite (OS) sign dilepton event selection used in this study. The statistical procedure used for exclusion as well as the discovery potential is summarized in Section 4. In Section 5 we discuss the exclusion limits and mass reach using SS dileptons followed by similar study involving OS dileptons in Section 6. Finally, in Section 7 we summarize the results.

The work presented here updates work previously documented in[10] and [9] from 10TeV center of mass to 7TeV, and from 2 series to 3 series MC.

2 Data Samples

This study is based on the 3_1_X reco full simulation SM background samples listed in Table 1. The Standard Model (SM) data sets have been normalized to the cross-sections [1]; for the SUSY parameter scan studies, we have used 3_3_6 reco fast simulation data set [3]. In this data set several benchmark point are produced using fixed $A_0 = 0$, tan$\beta = 3$, sign $\mu > 0$ but with varying $m_0(0 - 2000)$ and $m_{1/2}(100 - 600)$ on a grid of 50 and 20 GeV steps, respectively.

/WW/Summer09-MC_31X_V3_7TeV-v1/GEN-SIM-RECO
/WZ/Summer09-MC_31X_V3_7TeV-v1/GEN-SIM-RECO
/ZZ/Summer09-MC_31X_V3_7TeV-v1/GEN-SIM-RECO
/Wenu/Summer09-MC_31X_V3_7TeV-v1/GEN-SIM-RECO
/Wmunu/Summer09-MC_31X_V3_7TeV-v1/GEN-SIM-RECO
/Wtaunu/Summer09-MC_31X_V3_7TeV-v1/GEN-SIM-RECO
/Zee/Summer09-MC_31X_V3_7TeV-v1/GEN-SIM-RECO
/Zmumu/Summer09-MC_31X_V3_7TeV-v1/GEN-SIM-RECO
/Ztautau/Summer09-MC_31X_V3_7TeV-v1/GEN-SIM-RECO
/TTbar/Summer09-MC_31X_V3_7TeV-v1/GEN-SIM-RECO
/SingleTop_sChannel-madgraph/Summer09-MC_31X_V3_7TeV-v1/GEN-SIM-RECO
/SingleTop_tChannel-madgraph/Summer09-MC_31X_V3_7TeV-v1/GEN-SIM-RECO
/SingleTop_tWChannel-madgraph/Summer09-MC_31X_V3_7TeV-v1/GEN-SIM-RECO
/TANB3_CMSW336FASTv0JetID/spadhi-TANB3_CMSW336FASTv0JetID-*/USER

Table 1: The data sets used in this study.

The Monte Carlo events were analyzed with CMSSW_3_3_6 with the additional tags listed in Table 2.

V00-03-04	RecoEgamma/EgammaTools
V03-00-12-13	RecoMET/METProducers
V00-02-07-15	RecoMET/METAlgorithms
V00-06-10-02	RecoMET/Configuration
V03-01-01-04	DataFormats/METReco
V00-05-38	RecoEcal/EgammaCoreTools
V01-08-23-05	JetMETCorrections/Configuration
V01-08-08-09	CondFormats/JetMETObjects
V02-06-03	HLTrigger/HLTcore

Table 2: Additional software tags used in this study.

3 Event Selection

The event selection used is not optimized for any specific SUSY scenario. It is based on small modifications to the dilepton event selections that we used in approved WW[4] and $t\bar{t}$ [5] cross-section analyses. The OS and SS analyses summarized here are carbon copies of AN-XXX and AN-YYY. The only difference to those analysis notes is that we now use the 3 series MC at 7TeV instead of the 2 series MC at 10TeV. A quick summary of the event selection is:

- We require inclusive lepton triggers with no isolation, $i.e.$, the logical OR of `HLT_Ele15_SW_L1R` and `HLT_Mu9`. The combined trigger efficiency is $\sim 99\%$ for dilepton events that pass the event selection.
- Two isolated, same or opposite sign leptons (ee, $e\mu$, and $\mu\mu$).
- Leptons must have $P_T > 10$ GeV, $|\eta| < 2.4$ and at least one of them must have $P_T > 20$ GeV.
- We consider L2L3 corrected caloJets with $P_T > 30$ GeV and $|\eta| < 2.4$ for both analyses.
- The scalar sum of the P_T of all jets passing the requirements above should be > 200 GeV.
- For SS analysis:
  - we veto the candidate lepton, if an extra lepton in the event, pairs with the candidate lepton to form a Z within the mass range between $76 < m_{\ell\ell}$ (GeV) < 106. This requirement is designed to reject WZ events.
  - At least three jets.
  - We require $\not\!\!E_T > 80$ GeV.
- For the OS analysis:
  - We remove ee and $\mu\mu$ pairs consistent with a Z by requiring mass($\ell\ell$) < 76 GeV or mass($\ell\ell$) > 106 GeV.
  - At least two jets.
  - we have a general requirement that $\not\!\!E_T > 50$ GeV *. We additionally define a "tight" $\not\!\!E_T$ requirement of $\not\!\!E_T > 175$ GeV that is used for reporting predicted and observed event yields. This latter cut is intended to allow just a few Standard Model events to pass in 100 pb^{-1}. For $\not\!\!E_T$, we use tcMET [14] corrected for μ.

More details on the lepton selections are given below.

3.1 Electron Selection

- In our corresponding 10TeV analyses with 2 series MC, we used "e-gamma category based tight" for electron ID. An exact equivalent does not exist in the 3 series, and we thus tried out a variety of different options, including "e-gamma category based loose". We find differences in efficiency and background at the 10% levels, and consider them negligible for the purpose of this study.
- No muon candidate within $\Delta R < 0.1$.

- $|d_0| < 200\ \mu m$ (corrected for beamspot).
- Iso $<$ 0.1, where Iso=Sum/Max(20 GeV, P_T), and Sum = tkIso + hcalIso + Max(0 GeV, ecalIso - 2GeV). All isolation sums are the standard sums used in release 3_1_X from the egamma group (cone of 0.4 for ecal, jurassic, rec-hit based; cone of 0.3 for tracker, and cone of 0.4 for hcal).
- Conversion rejection [**?**] using tracks within cone of 0.3 of the candidate electron for SS studies:
  - $|\Delta \cot\theta| < 0.02$; the difference between cotangent polar angles of tracks parallel to each other.
  - $|d_{2d}| < 0.02$ cm; the two dimensional distance between points within nearest tracks.
  - The charge of the associated GSF and CTF tracks must be consistent. If the CTF track is not reconstructed, the electron is kept.

3.2 Muon Selection

- Must be a global muon **and** a tracker muon [6].
- GlobalMuonPromptTight (global χ^2/ndof$<$10) [7].
- At least 11 valid hits for the silicon track [7].
- $|d_0| < 200\ \mu m$ (from silicon track, corrected for beamspot).
- Global fits must have hits in the muon chambers.
- Minimum ionizing: EcalVetoEnergy $<$ 4 GeV and HcalVetoEnergy < 6 GeV [15].
- Iso $<$ 0.1, where Iso=Sum/Max(20 GeV, P_T), and Sum = tkIso + hcalIso + ecalIso. All isolation sums are the standard sums stored in the muon object in release 3_1_X, and are calculated in a cone of 0.3.

4 Statistical Methods

5 Same Sign Dileptons

This section summarizes the results of the SUSY parameter space scan in the same sign dilepton channel. The measurement technique is described in detail in a CMS note [9]. The technique utilizes a data-driven method to estimate background characterized by the presence of two high P_T, isolated, same sign leptons, $\not\!\!E_T$, and significant jet activity. This generic signature is sensitive to many new physics scenarios such as SUSY. For the purposes of this note we restrict ourselves to the ee, $e\mu$, and $\mu\mu$ final states, *i.e.*, we do not consider τ's, except in the case that the τ decays leptonically.

As we will show in Section 6.1, for a reasonable event selection the main background is $t\bar{t}$ decays. The data-driven background prediction is based on a combination of estimating "fake leptons"[8] and electrons reconstructed with the wrong sign[9]. The probability for muons to be reconstructed with the wrong sign is so small at the relevant momenta that it is negligible.

5.1 Event Yields

The expected event yields in 100 pb^{-1} after applying the event selections described in Section 3 to the data sets described in Section 2 are detailed below: the SM yields are listed in Table 3, and the mSUGRA scan point yields are illustrated in Fig. **??**.

5.2 Procedure for Determining 5σ Discovery Reach

We determine the 5σ discovery reach in the $m_0 - m_{1/2}$ plane by performing our analysis at each of the mSUGRA scan points. For each point, we determine expected yield based on the number of events passing all cuts scaled by the LO cross section for that point assuming 100/pb or 1/fb of data respectively. We then in addition, perform our data driven background estimation procedure at each point to determine the "signal contamination" at that point. We find the latter to be small, less than 10% of the expected signal yield.

We quantify the significance of the discrepancy between the observed yield and predicted background yield using two significance estimators: Z_{Bi} [16] and Z_N [**?**]. The quantities required to calculate these estimators are:

Sample	expected bkg yields
$t\bar{t}$	0.36 ± 0.10
WW	0.0
WZ	0.02 ± 0.01
ZZ	0.0
W+jets	0.0
Z+jets	0.0
Single top	0.0

Table 3: Expected SM event yields in 100 pb^{-1}. Errors are MC statistics only.

- The predicted background yield
- The relative systematic uncertainty on the predicted background yield (set to 50%)
- The statistical uncertainty on the predicted background yield (Z_N only, set to 0)
- The observed yield

5.3 Procedure for Excluding a Region of the mSUGRA Parameter Space

Next we determine the region of the mSUGRA parameter space which we expect to exclude at 95% confidence level (CL) if we see the standard model (SM) expected yields in data. We assume that we find the same predicted background yield and observed yield in data that we expect to find based on our SM MC. However, as the latter is 0.4, we work the math for both an observation of 0 or 1 events in 100/pb. We use this information to exclude a subset of the mSUGRA points using the following procedure.

The first step is to determine the 95% CL upper limit (UL) on the signal yield using a Bayesian method from John Conway, implemented in the program bayes.f. The required inputs are: the observed yield, the relative uncertainty in the signal acceptance (set to 15%), the predicted background yield, and the total error on the predicted background yield. We assume 0.4 ± 0.6 for background yield and its uncertainty.

These values lead to to 95% CL ULs of 3.2 and 4.7 signal events respectively for an assumed observation of 0 or 1 events.

Next, we wish to exclude mSUGRA points based on the signal yield UL derived above. The most obvious way to do so is to exclude points which lead to a difference between observed yield and predicted background yield which exceeds the UL on the signal yield. As the effect of signal contamination is small, we stick to this simplest of possible ways.

6 Opposite Sign Dileptons

This section summarizes the results of the SUSY parameter space scan in the opposite sign dilepton channel. The measurement technique is described in detail in a CMS note [10]. The technique utilizes a data-driven method to estimate background characterized by the presence of two high P_T, isolated, opposite sign leptons, large $\not\!\!E_T$, and significant jet activity. This generic signature is sensitive to many new physics scenarios such as SUSY. For the purposes of this note we restrict ourselves to the ee, $e\mu$, and $\mu\mu$ final states, *i.e.*, we do not consider τ's, except in the case that the τ decays leptonically.

As we will show in Section 6.1, for a reasonable event selection the main background is $t\bar{t}$ decays. The data-driven background prediction is based on a suggestion by Victor Pavlunin [11]. The idea is that in dilepton $t\bar{t}$ events the leptons and neutrinos from W decays have on average the same P_T spectrum (modulo effects of $V - A$). One can then use the **observed** $P_T(\ell\ell) \equiv |\vec{P}_T(\ell_1) + \vec{P}_T(\ell_2)|$ distribution to model the sum of neutrino P_T's which is identified with $\not\!\!E_T$.

6.1 Event Yields

The expected event yields in 100 pb^{-1} after applying the event selections described in Section 3 to the data sets described in Section 2 are detailed below: the SM yields are listed in Table 3, and the mSUGRA scan point yields are illustrated in Fig. 1.

In Fig. 2 one sees, in blue triangles, the SM $\not\!\!E_T$ distribution predicted by the data-driven background estimation technique described earlier, and, in red open circles, the actual SM $\not\!\!E_T$ distribution. The predicted $\not\!\!E_T$ distribution is the $P_T(\ell\ell)$ distribution scaled to account for the $\not\!\!E_T$ >50 GeV cut; the scale factor is about 1.6. The predicted event yield for a given $\not\!\!E_T$ cut is then obtained by integrating over this distribution starting from the corresponding $P_T(\ell\ell)$ value. For $\not\!\!E_T$ >175 GeV, the method predicts 3.9±0.51 SM events. The true yield is 4.3±0.27 SM events, which agrees well with the prediction. The reported errors are statistical only.

Sample	tcMET > 175
$t\bar{t}$	3.99
WW	0.19
WZ	0.01
ZZ	0.01
W+jets	0
Z+jets	0
Single top	0.05

Table 4: Expected SM event yields in 100 pb^{-1}.

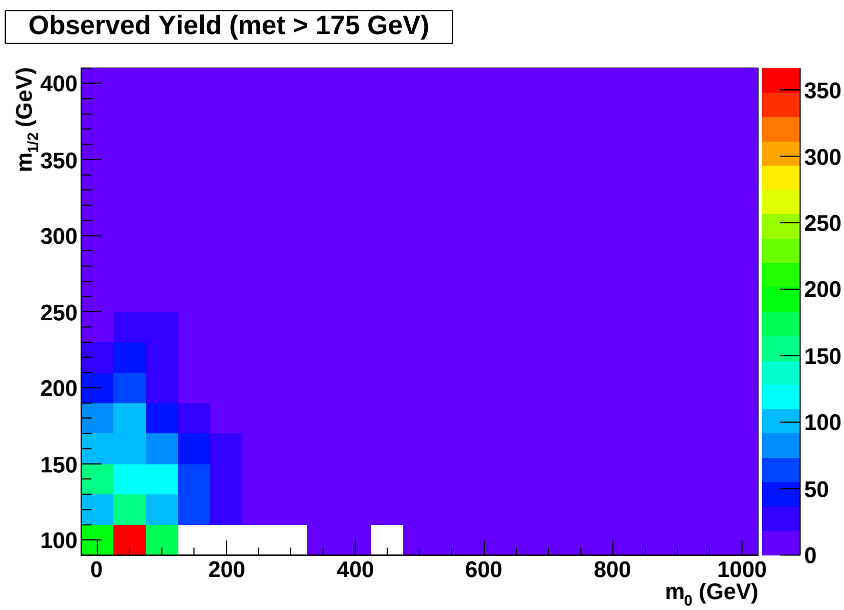

Figure 1: Expected mSUGRA scan point event yields in 100 pb^{-1}

6.2 Procedure for Determining 5σ Discovery Reach

We determine the 5σ discovery reach in the $m_0 - m_{1/2}$ plane by performing our analysis at each of the mSUGRA scan points. For each point, we determine the distributions of $p_T(\ell\ell)$ and $\not\!\!E_T$ and add these to the corresponding SM distributions. Then we perform the data-driven background estimate by using the $p_T(\ell\ell)$ to predict the number of events with $\not\!\!E_T > 175$ GeV (predicted background yield), as well as directly count the number of events which pass this $\not\!\!E_T$ cut (observed yield). We quantify the significance of the discrepancy between the observed yield and predicted background yield using two significance estimators: Z_{Bi} [16] and Z_N [**?**]. The quantities required to calculate these estimators are:

- The predicted background yield
- The relative systematic uncertainty on the predicted background yield (set to 25%)
- The statistical uncertainty on the predicted background yield (Z_N only, set to 0)
- The observed yield

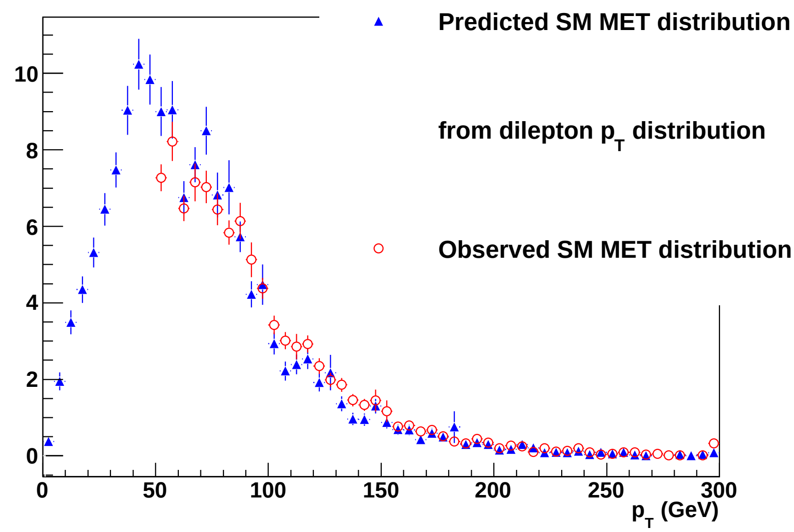

Figure 2: Rescaled $P_T(\ell\ell)$ distribution in the SM cocktail used to predict the $\not\!\!E_T$ in blue, compared with the $\not\!\!E_T$ in red.

In Figs. 3 and 4 we display the Z_{Bi} significance assuming integrated luminosities of $100\ \mathrm{pb}^{-1}$ and $1\ \mathrm{fb}^{-1}$, respectively. In Figs. 5 and 6 we display the Z_N significance assuming the same integrated luminosities.

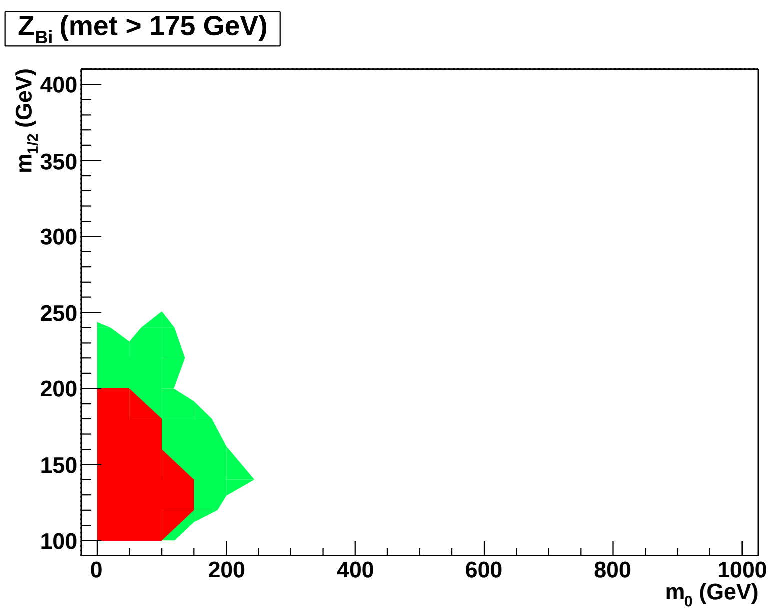

Figure 3: The Z_{Bi} significance in the $m_0 - m_{1/2}$ plane assuming an integrated luminosity of $100\ \mathrm{pb}^{-1}$. Only points with $Z_{Bi} \geq 5$ are displayed.

6.3 Procedure for Excluding a Region of the mSUGRA Parameter Space

Next we determine the region of the mSUGRA parameter space which we expect to exclude at 95% confidence level (CL) if we do not see evidence for signal in data. We assume that we find the same predicted background yield and observed yield in data that we expect to find based on our SM MC. We use this information to exclude a subset of the mSUGRA points using the following procedure.

The first step is to determine the 95% CL upper limit (UL) on the signal yield using a Bayesian method from John Conway, implemented in the program bayes.f. The required inputs are: the observed yield, the relative uncertainty in the signal acceptance (set to 15%), the predicted background yield, and the total error on the predicted back-

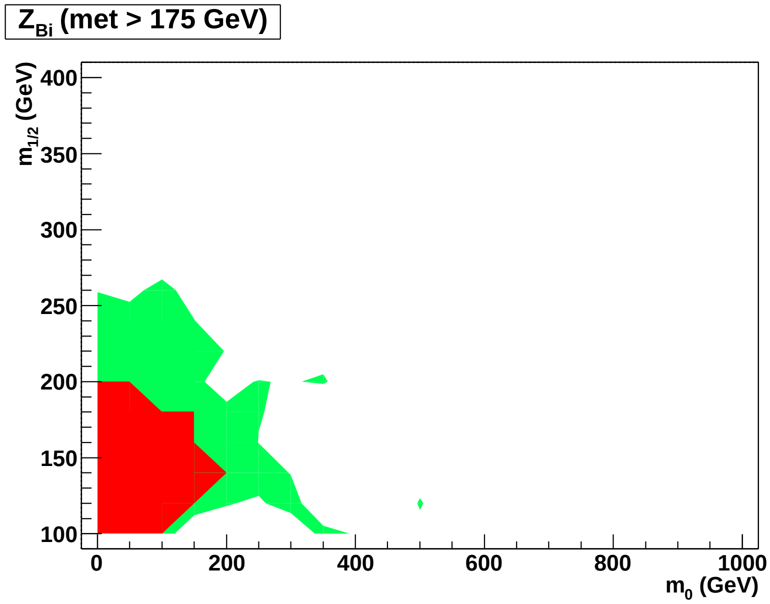

Figure 4: The Z_{Bi} significance in the $m_0 - m_{1/2}$ plane assuming an integrated luminosity of $1\ \mathrm{fb}^{-1}$. Only points with $Z_{Bi} \geq 5$ are displayed.

ground yield. We evaluate this error as the quadrature sum of the systematic error (set to 25% of the predicted background yield) and the statistical uncertainty, equal to $k\sqrt{N_{BKG}}$, where $k \approx 1.6$ is the scaling factor applied to the $p_T(\ell\ell)$ distribution to account for the $\not{E}_T > 50$ GeV cut, and N_{BKG} is the predicted background yield. We find an error of $\sigma_{BKG} = 3.4$ for 100 pb^{-1} ($N_{BKG} = 4$) and $\sigma_{BKG} = 14.2$ for 1 fb^{-1} ($N_{BKG} = 40$). These values lead to 95% CL ULs of 7.6 signal events and 33.1 signal events for 100 pb^{-1} and 1 fb^{-1}, respectively. It should be noted that bayes.f assumes a Gaussian error distribution, while for our analysis (especially at 100 pb^{-1} where the background yield is 4), our errors are Poisson-distributed.

Next, we wish to exclude mSUGRA points based on the signal yield UL derived above. The most obvious way to do so is to exclude points which lead to a difference between observed yield and predicted background yield which exceeds the UL on the signal yield. However, due to the effects of signal contamination, which in our analysis can lead to large biases in the background prediction, one cannot rely on this difference for exclusion. Consider an mSUGRA point which leads to an observed yield of 110 and a predicted background yield of 100 for an integrated luminosity of 100 pb^{-1}. Since $110 - 100 > 7.6$ we would exclude this point using as our metric the difference between observed and predicted yields. However, the statistical difference between these yields is only at the $\approx 1\sigma$ level and hence this point should not be excluded at 95% CL. Instead, we use as our metric the *significance* of the discrepancy between observed yield and predicted background yield, quantified by Z_{Bi}. First, we determine the Z_{Bi} significance corresponding to the UL on the total yield (*ie.* the observed yield plus the UL on the signal yield) compared to the predicted background yield, again assuming a relative systematic uncertainty of 25% on the predicted background yield. For 100 pb^{-1} we find $Z_{Bi} = 2.5$ for an observed yield of 4+7.6=11.6 and predicted background of 4, while for 1 fb^{-1} we find $Z_{Bi} = 2.2$ for an observed yield of 40+33.1=73.1 and predicted background of 40. Finally, we exclude those mSUGRA points which lead to a larger Z_{Bi} significance between the observed yield and predicted background yield than these values, as shown in Figs. 7 and 8 for 100 pb^{-1} and 1 fb^{-1}, respectively.

7 Conclusion

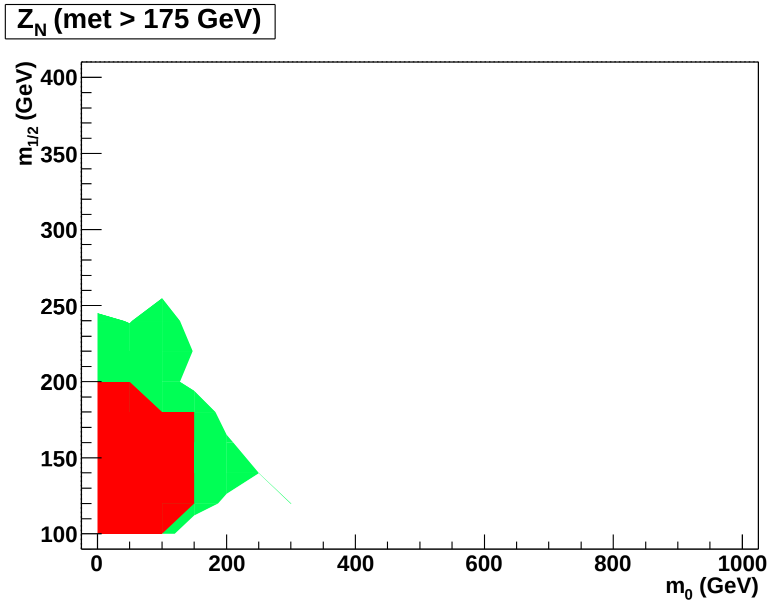

Figure 5: The Z_N significance in the $m_0 - m_{1/2}$ plane assuming an integrated luminosity of $100\ \mathrm{pb}^{-1}$. Only points with $Z_N \geq 5$ are displayed.

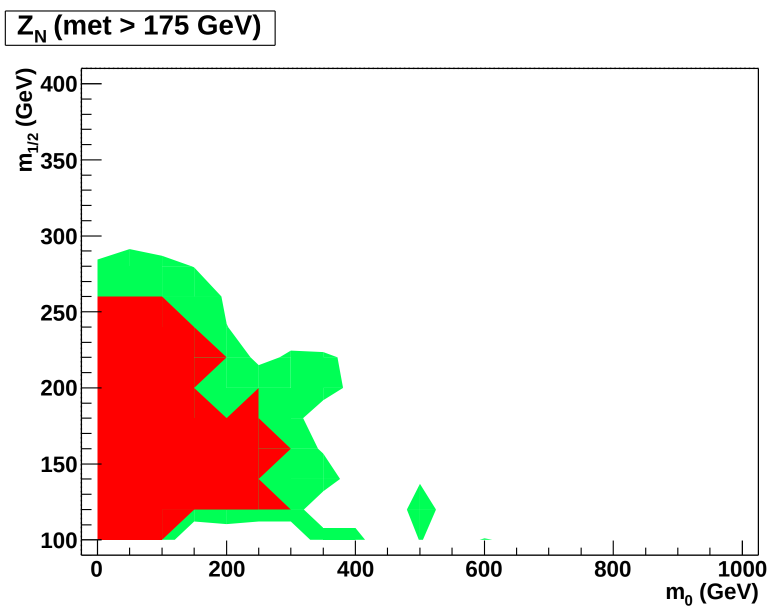

Figure 6: The Z_N significance in the $m_0 - m_{1/2}$ plane assuming an integrated luminosity of $1\ \mathrm{fb}^{-1}$. Only points with $Z_N \geq 5$ are displayed.

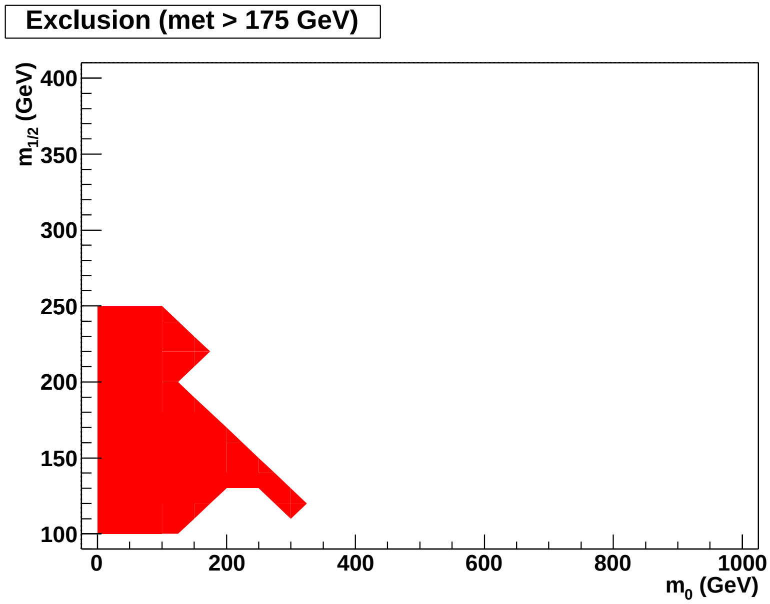

Figure 7: The excluded region (red shaded area) of the $m_0 - m_{1/2}$ plane assuming an integrated luminosity of $100\ \mathrm{pb}^{-1}$.

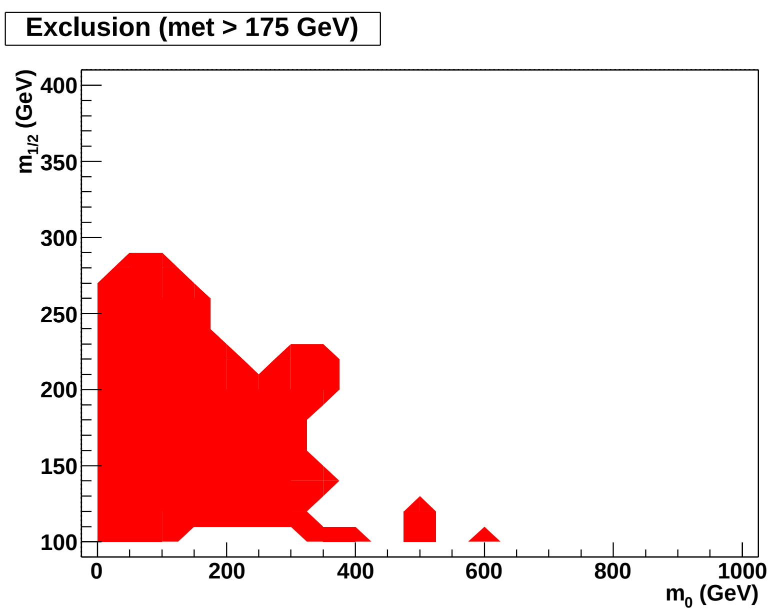

Figure 8: The excluded region (red shaded area) of the $m_0 - m_{1/2}$ plane assuming an integrated luminosity of $1\ \mathrm{fb}^{-1}$.

References

[1] https://twiki.cern.ch/twiki/bin/viewauth/CMS/SUSYMCRequirements0911.

[2] https://twiki.cern.ch/twiki/bin/view/CMS/SUSY31XProduction.

[3] https://twiki.cern.ch/twiki/bin/view/CMS/SUSY33XScan.

[4] "Prospects for measuring the WW production cross section in pp collisions at $\sqrt{s}$ =10 TeV", CMS AN-2009/042 and PAS EWK-09-002.

[5] "Expectations for observation of top quark pair production in the dilepton final state with the early CMS data", CMS AN-2009/050 and PAS TOP-09-002.

[6] https://hypernews.cern.ch/HyperNews/CMS/get/muon/258.html.

[7] "Muon Identification in CMS", CMS AN-2008/098.

[8] "Data-driven methods to estimate the electron and muon fake contributions to lepton analyses", CMS AN-2009/041.

[9] "Data driven background study for new physics searches with same sign dileptons at $\sqrt{s} = 10$ TeV", CMS AN-2009/138.

[10] "Data driven background estimate for a new physics search with opposite sign dileptons", CMS AN-2009/130.

[11] http://arxiv.org/pdf/0906.5016;
http://indico.cern.ch/contributionDisplay.py?contribId=2&confId=39042.

[12] http://www.ge.infn.it/ tosi/cms/topMC.html.

[13] https://twiki.cern.ch/twiki/bin/view/CMS/ProductionSummer2008.

[14] CMS AN-2009/022.

[15] https://twiki.cern.ch/twiki/bin/view/CMS/VplusJets.

[16] arXiv:physics/0702156 [physics.data-an]