

18 May 2009 (v2, 19 May 2009)

# Discovery potential and measurement of a dilepton mass edge in SUSY events at $\sqrt{s}=10~{\rm TeV}$

N. Mohr

I. Physikalisches Institut B, RWTH Aachen University, Germany

#### Abstract

Within the minimal supergravity model (mSUGRA), the observability of the decay of the next to lightest neutralino into leptons has been studied using the full simulation of the CMS detector. The final state signature consists of two opposite sign leptons, several hard jets and missing transverse energy. Using three different minimal supergravity benchmark points the possible discovery of a mSUGRA signal is studied. The expected precision of the measurement of the dileptonic mass edge is reported for 200 pb<sup>-1</sup> and 1 fb<sup>-1</sup> of data, including systematic and statistic uncertainties and comparing different decay signatures.

## **1** Introduction

The standard model of particle physics (SM) leads to a number of unsolved issues like the hierarchy problem and it provides no solution for pressing questions arising from astrophysical observations, most notably dark matter. In Supersymmetry (SUSY) a natural candidate for dark matter can be found if R-parity conservation is assumed. Supersymmetric particles (sparticles) have not been observed up to now which implies that they have to be heavy. On the other hand to provide a solution to the hierarchy problem their masses have to be in the TeV range.

The long anticipated start of the Large Hadron Collider (LHC) in 2009 will allow to explore this new TeV range. With its center of mass energy of 10 TeV in 2010 it will allow to probe supersymmetric models very early on. A key point after discovery will be the determination of the sparticle properties. If R-parity is conserved the lightest neutralino escapes detection and no mass peaks can be observed in SUSY decay chains. Of special interest are robust signatures such as mass edges in leptonic final states which can be probed with the CMS experiment.

The purpose of this analysis is to observe a significant excess of opposite sign same flavour leptons over the various backgrounds and to measure the endpoint in the invariant mass distribution. All flavour symmetric background (including SUSY decays of this type) can be determined from data events with opposite sign opposite flavour leptons. The aim is to perform such an analysis already with the first LHC data which is expected to amount to roughly 200-300 pb<sup>-1</sup> in 2010.

## 2 Signal

Three minimal supergravity benchmark points (Tab.1) have been studied to cover different decay modes of the neutralinos within supersymmetry. The mass spectra of the three benchmark points have been calculated using the Softsusy code [1]. All branching ratios have been calculated with the SUSYHit program [2] and the events are simulated using Pythia[3]. The k-factor for the cross section at 10 TeV is calculated using a modified version of Prospino 2 [4]. In mSUGRA there are very long decay chains leading to several hard jets. The escaping neutralino leads to missing transverse energy. This fact allows to define a search region to observe an excess over the SM and is used as main event selection as described in Sec.5.

	$m_0$ [GeV]	$m_{1/2}  [{ m GeV}]$	$A_0$ [GeV]	$\tan\beta$	sign $\mu$	$\sigma_{LO}$ [pb]	$\sigma_{NLO}$ [pb]	$m_{ll,max}$ [GeV]
LM0	200	160	-400	10	+1	110.0	151.8	52,7
LM1	60	250	0	10	+1	16.1	21.7	78,2
LM9	1450	175	0	50	+1	11.1	18.2	62,9

Table 1: mSUGRA benchmark points LM0, LM1 and LM9.

Additionally the leptonic decay of the next to lightest neutralino leaves a characteristic signature. This decay itself can manifestate in different ways even in the mSUGRA model. A mass difference of the neutralinos smaller than the Z-boson mass and any slepton mass leads to a three body decay. In that case the endpoint in the lepton invariant mass represents directly the mass difference of the two lightest neutralinos

$$m_{ll,max} = m_{\tilde{\chi}_2^0} - m_{\tilde{\chi}_1^0}.$$
 (1)

The shape of the distribution depends on the mSUGRA parameters and is shown in Fig. 1(a) for the LM9 benchmark point at parton level.

A two body decay occurs via a real slepton and is allowed if at least one slepton is lighter than the mass difference of the neutralinos. In that case the endpoint can be expressed by

$$(m_{ll}^{max})^2 = \frac{\left(m_{\tilde{l}}^2 - m_{\tilde{\chi}_2^0}^2\right) \left(m_{\tilde{\chi}_2^0}^2 - m_{\tilde{\chi}_2^0}^2\right)}{m_{\tilde{l}}^2},\tag{2}$$

where  $m_{\tilde{l}}$  is the mass of the intermediate slepton. The shape of the mass edge results only from kinematics and is triangular as shown in Fig. 1(b) for LM1. At LM0 (Fig. 1(c)) the 3-body decay mode is present as for LM9.

If the mass difference matches the Z-boson mass this leads to an enhanced Z-boson production accompanied by large missing transverse energy and several hard jets. This signature is not in the focus of the present analysis. It has been studied for example in [5]. Another possibility in mSUGRA is a decay predominantly through the lightest Higgs-boson which leads to a different signature as well.



Figure 1: Invariant mass of lepton pairs originating from signal decays in case of a 3-body decay at LM9 is shown in (a). Triangular shaped invariant mass of lepton pairs in case of a 2-body decay at LM1 in Fig. (b). The invariant mass distribution at LM0 (3-body decay) is shown in Fig. (c).

## **3** Physics objects

## 3.1 Monte Carlo Datasets

The datasets with a Vector boson or a  $t\bar{t}$ -pair have been simulated using the Madgraph matrix element generator [6]. The parton shower and hadronisation is modelled in Pythia. The QCD, Quarkonia, W $\gamma$  and Z $\gamma$  samples are simulated using Pythia.

The samples have been scaled to next to leading order cross sections. For the SUSY samples the cross-section calculated with Prospino 2 has been used. The k-factor for the  $t\bar{t}$  sample of 1.3 has been derived from [7]. The k-factors of the Z+jets and the W+jets samples of 1.14 have been derived from [8]. For the QCD sample no k-factors are available.

The datasets where produced during the Summer 08 Monte Carlo (MC) production. The MC production was targeted of an integrated luminosity of  $200 \text{ pb}^{-1}$  and used ideal calibration and alignment constants. The simulation samples are reconstructed using version 2\_1\_X of the CMS software. The simulation and digitisation is carried out in CMSSW 2\_0\_X. All samples are listed in App. A. From these samples so called Pat-Tuples have been created using the following tags of the Physics Analysis Toolkit (PAT) in CMSSW\_2\_2\_X.

```
PhysicsTools/PatAlgos V04-14-19
PhysicsTools/PatUtils V03-05-02
DataFormats/PatCandidates V03-18-04
TopQuarkAnalysis/TopObjectProducers/python V04-07-00
```

All physics objects necessary for this analysis are in included in the Pat-Tuples. The physics object selection is based on the V+jets [9] recommendation.

## 3.2 Muons

As muon objects the standard muon collection [10] of the CMSSW reconstruction has been used via PAT. Each muon has to be identified as a global muon. The track of the muon in the inner tracker has to have at least 11 hits and a  $\chi^2/ndf$  below 10. Additionally a  $p_T > 10$  GeV and  $|\eta| < 2$  is required for each muon. The impact parameter of the muon track which is corrected for the beam-spot position is required to be below 2 mm and this cut could be tightened if necessary (the current set of events does not include misalignment so a too tight cut could reduce the efficiency in real data).

#### 3.3 Electrons

The electron collection [11] is derived from the PIXELMATCHGSFELECTRONS and provided by PAT. Each electron has to fulfill the tight electron identification provided in the standard reconstruction sequence of CMSSW 2\_1\_X. Additionally a  $p_T > 10$  GeV and  $|\eta| < 2$  is required for each electron. The impact parameter of the electron track which is corrected for the beam-spot position is required to be below 2 mm as in the muon case.

#### 3.4 Lepton isolation

A combined relative lepton isolation has been used. The isolation uses information from both calorimeters and the silicon tracker. The isolation value (*Iso*) is given by the ratio of the sum of all  $E_T$  or  $p_T$  objects within a cone in  $\eta$ - $\phi$ -space of  $\Delta R = \sqrt{\Delta \eta^2 + \Delta \phi^2} < 0.3$  around the lepton and the lepton  $p_T$ . It has been pre-calculated in PAT using

$$Iso = \frac{\sum_{ECAL} E_T + \sum_{HCAL} E_T + \sum_{tracks} p_T}{p_T}$$
(3)

where the first sum is runs over the transverse energy in the electromagnetic calorimeter, the second sum runs over the transverse energy in the hadronic calorimeter and the third sum runs over the transverse momentum deposited in the tracker within the cone.

The isolation for muons is shown Fig. 2(a) and the cut value is chosen to be Iso < 0.2. The distribution for electrons is displayed in Fig. 2(b) and the cut is placed at Iso < 0.4 to obtain a similar rejection and efficiency for electrons and muons.



Figure 2: Isolation value for muons (a) and electrons (b) passing the acceptance and identification cuts in  $t\bar{t}$  and SUSY LM0 events. The red curve shows all leptons which can be matched onto a prompt lepton. The blue curve represents leptons matched onto leptons from decays of heavy resonances. Magenta are unmatched leptons, e.g. fake leptons from jets.

## 3.5 Jets and missing transverse energy

The jet algorithm is a seedless infrared safe cone algorithm (SIScone) [12] with a cone size of 0.5 in  $\Delta R$  from the reconstruction sequence in CMSSW\_2\_1\_X. The jet collection is corrected up to level 3 using MC jet energy corrections [13]. Each corrected jet is required to have a  $p_T$  above 30 GeV and the jet axis has to be within  $|\eta| < 2.5$ . Additionally the overlap of the jets with the electrons is checked and a jet is vetoed if an electron shares its supercluster with the jet. The missing transverse energy (MET) is based on the calorimeter information and is corrected for muon energy deposition and jet energy scale [14].

## 4 Efficiency measurement

Since the muon and electron reconstruction efficiencies are not equal and can not be determined using MC only, one needs to measure these efficiencies from data. Therefore a tag and probe method using events with a Z boson is used. To select clean Z events a tight selection is applied to one lepton (tag) and only loose criteria are used on the probe side. Similar studies are presented in [15]. For this analysis a Z sample including all backgrounds with an integrated luminosity of  $100 \text{ pb}^{-1}$  has been used. All samples with a weight larger than 100 have been excluded because single events with a large weight would distort the lepton efficiency measurement.

## 4.1 Muon tag and probe method

The tag has to fullfil the muon selection criteria described in Sec. 3.2. Additionally a  $p_T$  of at least 20 GeV is required and the tag has to be reconstructed within  $|\eta| < 1$ . As probes all tracks in the event which have a  $p_T$ above 10 GeV within  $|\eta| < 2$  are used. The invariant mass of the tag and the probe is required to be in the range of the Z boson mass (80 <  $m_{tag,probe} < 100$  GeV). The efficiency is the calculated simply by counting number of events where the probe can be matched on a muon with a  $p_T$  of 10 GeV (within  $\Delta R < 0.1$ ), which passes the muon selection criteria described in Sec. 3.2.

The distribution of all tag and probe pairs is shown in Fig. 3(a). As expected the pairings do not consist purely of Z events and include also backgrounds mainly from di-jet and W events. These events would distort the efficiency measurement and therefore this background is measured from data by selection of same sign tag and probe pairs. The number of background events is simply two times the number of same sign pairs since in background events the number of same sign pairs should be the same as the number of opposite sign pairs  $N_B = 2 \cdot N_{SS}$ . We are aware of the potential bias due to a charge correlation of the tag and probe pair in W boson events and we do not account for a possible charge misreconstruction efficiency because it should give only a small correction to the obtained results. We include a potential bias of  $\pm 5\%$  in the efficiency calculation in the study of the systematic uncertainties.



Figure 3: The invariant mass of all muon tag and probe pairs is shown in (a). The black points represent the background estimation from the same sign tag and probe pairs. All pairs where the probe could be matched onto a reconstructed muon are shown in (b).

The efficiency can simply be calculated using

$$\epsilon = \frac{N_{pass} - N_{B,pass}}{N_{probe} - N_{B,probe}},\tag{4}$$

where  $N_{pass}$  is the number of passing probes,  $N_{probe}$  the number of probes and  $N_B$  the number of background events from the same sign selection.



Figure 4: Efficiency measurement by the tag and probe method for muons in comparison to MC truth, versus  $p_T$  (a) and  $\eta$  (b).

The efficiency is calculated in bins of  $\eta$  and  $p_T$ . Since the statistic is largest in the bulk region of the Z sample (lepton  $p_T$  around 50 GeV the binning is chosen accordingly. The efficiency versus  $p_T$  for muons is shown in Fig. 4(a). The  $\eta$  distribution is displayed in Fig. 4(b). A good agreement between the tag and probe results and MC expectation is found.

## 4.2 Electron tag and probe method

The tag has to fullfil the electron selection criteria described in Sec. 3.3. Additionally a  $p_T$  of at least 20 GeV is required and the tag has to be reconstructed within  $|\eta| < 1$ . As probes all tracks in the event which have a  $p_T$  above 10 GeV within  $|\eta| < 2$  are used. The invariant mass of the tag and the probe is required to be in the range of the Z-boson mass ( $80 < m_{tag,probe} < 100$  GeV).

In Fig. 5(a) the invariant mass of all electron tag and probe pairs is shown. The larger number of background events (especially from di-jet events) in this selection can be explained by the fact that the single electron selection leads



Figure 5: The invariant mass of all electron tag and probe pairs is shown in (a). The black points represent the background estimation from the same sign tag and probe pairs. All pairs where the probe could be matched onto a reconstructed electron are shown in (b).

to a less pure sample (compared to the muon selection). The asymmetry in the distribution is due to bremsstrahlung of electrons in the CMS tracker, which is not corrected for in the general track collection. The distribution of the passed tag and probe pairs shows a very pure Z sample as in the muon case.



Figure 6: Efficiency measurement by the tag and probe method for electrons in comparison to MC truth, versus  $p_T$  (a) and  $\eta$  (b).

The efficiencies are calculated using the same technique as described in Sec. 4.1. The efficiency versus  $p_T$  (Fig. 6(a)) shows a good agreement in the bulk region of the  $p_T$  from Z events. The  $\eta$  distribution of the efficiency agrees with the MC expectation as displayed in Fig. 6(b).

## 4.3 Efficiency correction

The efficiencies obtained from the tag and probe method are measured in bins of  $\eta$  and  $p_T$  which are chosen as shown in the projections in Sec. 4.1 and Sec. 4.2. These efficiencies have been used to correct the invariant mass distributions of the lepton pairs. Each lepton in the distributions gets a weight of  $1/\epsilon$ . The tag and probe method could be refined if necessary at a higher luminosity but for this analysis the expected precision is sufficient (Tab. 2 shows the global results from both methods).

## 5 Observability of the decay

In this part we discuss the observability of the signal decay including all backgrounds for 200  $pb^{-1}$  of MC data at the benchmark point LM0.

Table 2: Glo	bal efficiencies	obtained with	the tag and	probe method	compared to MC truth.

	$N_{pass}$	$N_{B,pass}$	$N_{probe}$	$N_{B,probe}$	$\epsilon_{TnP}$	$\epsilon_{MC}$
Muons	$60028 \pm 245$	$4\pm 2$	$68679 \pm 262$	$6662 \pm 81$	$0.968 \pm 0.013$	$0.964 \pm 0.003$
Electrons	$34550 \pm 185$	$690 \pm 26$	$50164 \pm 223$	$12400 \pm 111$	$0.897 \pm 0.010$	$0.881 \pm 0.003$

## 5.1 Trigger

To trigger the events leptonic triggers have been used because they will probably be more reliable in the beginning of CMS data taking. On the other hand if hadronic or MET triggers are working within the early data they provide a good cross check.

In this analysis we require two single leptonic high level trigger (HLT) paths

- HLT\_IsoMu11
- HLT\_IsoEle15\_LW\_L1I

to select the events. Since the leptons originating from the signal decay have a very soft  $p_T$  spectrum we propose to use the triggers with the lowest available threshold for electrons and muons.

Due to the SUSY signature of hard jets and missing transverse energy there exists the possibility to use a single hadronic trigger path to recover possible inefficiencies of the leptonic triggers and we compare their efficiencies to hadronic trigger path efficiencies.

Table 3: High level trigger efficiencies at LM0. The efficiency on the inclusive sample is  $\epsilon_1$  and the inclusive efficiency with respect to the final selection is  $\epsilon_2$ . The efficiency  $\epsilon_S$  is the trigger efficiency of signal events with respect to the final selection.

HLT path	Thresh. [GeV]	Pathname	$\epsilon_1$	$\epsilon_2$	$\epsilon_S$
Single $\mu$	9	HLT_Mu9	$28.9\pm0.1\%$	$74.9 \pm 1.2\%$	$65.8\pm2.8\%$
Single e	15	HLT_Ele15_SW_L1R	$42.6\pm0.2\%$	$76.2\pm1.2\%$	$67.1\pm2.7\%$
Single $e + \mu$	(9,15)	Mu9 OR Ele15	$58.6\pm0.1\%$	$99.3\pm0.2\%$	$99.8\pm0.2\%$
Single isolated $\mu$	11	HLT_IsoMu11	$13.4\pm0.1\%$	$65.0\pm1.3\%$	$58.2\pm2.9\%$
Single isolated e	15	HLT_IsoEle15_LW_L1I	$13.3\pm0.1\%$	$55.6 \pm 1.4\%$	$45.5\pm2.9\%$
Single iso. $e + \mu$	(11,15)	IsoMu11 OR IsoEle15	$26.9\pm0.1\%$	$96.1\pm0.5\%$	$97.2\pm0.7\%$
e+ $\mu$ Cross trigger	(10,10)	HLT_IsoEle10_Mu10_L1R	$3.7\pm0.1\%$	$32.8 \pm 1.3\%$	$13.7\pm2.0\%$
Single jet	110	HLT_Jet110	$80.3\pm0.1\%$	$97.9\pm0.4\%$	$99.0\pm0.6\%$
Di-jet	(70,70)	HLT_DiJetAve70	$74.8\pm0.1\%$	$94.2\pm0.7\%$	$97.6\pm0.9\%$
MET	50	HLT_MET50	$74.6\pm0.1\%$	$96.8\pm0.5\%$	$96.9 \pm 1.0\%$

The efficiencies on the signal events and on the inclusive sample are listed in Tab. 3. While hadronic triggers are more efficient on the inclusive sample without event selection, leptonic triggers become very efficient when the final lepton selection is applied. We observe an efficiency of  $96.1 \pm 0.5\%$  for the isolated leptonic paths and an efficiency of  $99.3 \pm 0.2\%$  for the non isolated leptonic paths in the final selection. The most efficient hadronic trigger path yields an efficiency of  $97.9 \pm 0.4\%$  using the single jet trigger with a threshold of 110 GeV. For this analysis we used the leptonic paths but with real data the hadronic paths can be used as cross-check.

All used trigger paths (apart from the single electron trigger which is now prescaled) are included in the proposed lean trigger menu V0.4 [16] for an instantaneous luminosity of  $10^{31}$  cm<sup>-2</sup> s<sup>-1</sup>. Therefore we use the single isolated leptonic trigger paths.

## 5.2 Event selection

The base selection requires two leptons of opposite sign. We do not require two same flavour leptons in order to measure the background events directly from this dataset as described in Sec. 5.3. The main SUSY selection is

based on jets and missing transverse energy. The cuts have not been optimised at a certain benchmark point, but should reflect the general SUSY signature. The selection requires three jets with  $p_T^{j1} > 100$  GeV,  $p_T^{j2} > 50$  GeV, and  $p_T^{j3} > 50$  GeV. Additionally a missing transverse energy of at least 100 GeV is required.

	$\sigma_{LO}$ [pb]	k-factor	Events analysed	HLT	$\geq 2$ leptons	$\geq 3$ jets	MET
LM0 signal	1.0	1.38	370	362	226	128	86
LM0 inclusive	110.0	1.38	88872	8067	1021	546	366
tt+jets	319.0	1.3	1005253	25655	2474	245	84
Z+jets	3700.0	1.14	1221378	541013	190032	409	1
W+jets	40000.0	1.14	6230138	3108397	358	5	2
Diboson	51.9	1.0	404155	5444	911	2	0
Di-jets	2003572.9	15368755	1.0	2801134	1116	4	0

Table 4: Number of selected events using the described event selection for an integrated luminosity of 200  $pb^{-1}$ .

The number of events obtained after each cut is listed in Tab. 4. After HLT selection the sample is still dominated by di-jet and W events. The requirement of two isolated and well identified opposite sign leptons rejects most of the di-jet events and the selection is dominated by events with a Z boson. After requirement of three hard jets a SUSY inclusive signal to background ratio of roughly one can be reached. After requirement of missing transverse energy the main background from the standard model consists of  $t\bar{t}$ -events. With the described selection an efficiency of 25% on the signal events is obtained. At the studied benchmark point LMO a high number of SUSY background events is found (flavour symmetric background from chargino decays) which is irreducible. This complicates the discovery of this decay at this point as described in Sec. 6.4.

## 5.3 Statistical measurement of the background events

All background which leads to uncorrelated lepton pairs can be measured directly from data [17]. Therefore we select the opposite sign opposite flavour lepton pairs and use this distribution (Fig. 7(b)) to extrapolate to the same flavour opposite sign lepton pair distribution shown in Fig. 7(a). One can see that the shape of the flavour symmetric background events is nicely reproduced for  $m_{ll} > m_Z$ .



Figure 7: The same flavour opposite sign lepton pair distribution without any event selection cuts is shown in (a). Black points represent the extrapolation from (b), where the opposite flavour opposite sign lepton pair distribution without event selection cuts is displayed.

With this method one is able to predict all backgrounds which produce uncorrelated leptons such as W,  $t\bar{t}$ , di-jet and WW events. In case of di-jet events the remaining events originate from quarkonia and the associated samples include only decays into muons.

The invariant mass distribution of all opposite sign same flavour leptons for 200 pb<sup>-1</sup> is shown in Fig. 8(a). In this plot no scaling has been applied but only the number of expected events in 200 pb<sup>-1</sup> has been analysed. The opposite sign same flavour distribution used to extrapolate the background is displayed in Fig. 8(b). One can see that at a low luminosity the statistical fluctuations are relatively large. Therefore these background is modelled and fitted as described in Sec. 6.



Figure 8: The same flavour opposite sign lepton pair distribution including all cuts is shown in (a). Black points represent the extrapolation from (b), which displays the opposite flavour opposite sign lepton pairs. No scaling has been applied but exactly 200  $pb^{-1}$  of MC events have been analysed.

## 6 Determination of the mass edge

The model used for the fit of the mass edge consists of three parts. To model the signal a quadratic term convoluted with a gaussian has been used in case of a 3-body decay

$$S(m_{ll}) = \frac{1}{\sqrt{2\pi\sigma}} \int_{0}^{m_{cut}} dy \cdot y^2 e^{\frac{-(m_{ll}-y)^2}{2\sigma^2}}.$$
 (5)

In case of the two-body decay the signal model consists of a triangle convoluted with a gaussian

$$T(m_{ll}) = \frac{1}{\sqrt{2\pi\sigma}} \int_{0}^{m_{cut}} dy \cdot y e^{\frac{-(m_{ll}-y)^2}{2\sigma^2}}.$$
 (6)

A curve parametrized as

$$B(m_{ll}) = m_{ll}^a \cdot e^{-b \cdot m_{ll}} \tag{7}$$

has been used to fit the opposite sign opposite flavour invariant mass distribution. Additionally the Z peak is fitted using a Breit-Wigner convoluted with a gaussian.

The fits are performed within the RooFit package [18] based on an unbinned and extended maximum likelihood fit on the di-lepton invariant mass distribution

$$L = \frac{e^{-(N_{Sig} + N_{Bkg} + N_Z)}}{(N_{Sig} + N_{Bkg} + N_Z)!} \prod_i [N_{Sig} P_S(m_{ll})_i + N_{Bkg} P_B(m_{ll})_i + N_Z P_Z(m_{ll})_i].$$
(8)

Here  $P_S = S$  or  $P_S = T$  is the signal probability density function (triangle or quadratic term convolved with a gaussian),  $P_B$  is the background model and  $P_Z$  is the Breit-Wigner function convoluted with a gaussian. The number of signal  $N_{Sig}$ , background  $N_{Bkg}$  and Z events  $N_Z$  are fitted as well.

#### 6.1 Resolution measurement

The resolution smearing of the detector, which is used in the fit function (Eq. 5+6), is measured from Z events. Therefore the selection using two well identified opposite sign leptons is used without any additional event selection cuts. In the two distributions for electrons and muons a Bifurcated Gaussian (with different widths on each side of the mean)

$$G_{BF}(m_{ll}) = \begin{cases} \frac{1}{\sqrt{2\pi}\sigma_L} e^{\frac{-(m_{ll}-m)^2}{2\sigma_L^2}}, & m_{ll} \le m \\ \frac{1}{\sqrt{2\pi}\sigma_R} e^{\frac{-(m_{ll}-m)^2}{2\sigma_R^2}}, & m_{ll} > m \end{cases}$$
(9)

is fitted to account for the asymmetry in the distribution. The detector resolution is averaged from the two resolution values obtained from the fit subtracting the natural Z boson width. The values of the resolution parameter in the convolution is used in the final fit in Sec. 6.2.



Figure 9: Resolution measurement from Z events for muons (a) and electrons (b).

The resolution fit for muons is shown in Fig. 9(a) and it yields a resolution of

$$\sigma_{\mu\mu} = (1.33 \pm 0.07) \text{ GeV.}$$
(10)

The fit of the electron invariant mass distribution is displayed in Fig. 9(b). It yields a resolution value of

$$\sigma_{ee} = (1.96 \pm 0.09) \text{ GeV.}$$
(11)

For the combined fit a combined resolution has been determined. It yields a value of

$$\sigma_{ll} = (1.61 \pm 0.06) \text{ GeV.}$$
(12)

Since the leptons in Z boson events tend to have a harder  $p_T$  spectrum compared to SUSY events we include the uncertainty in the invariant mass resolution in the systematic uncertainty on the endpoint determination. The method could be improved by measuring the resolution from  $J/\Psi$  and  $\Upsilon$  decays into leptons.

## 6.2 Fit of the mass edge at LM0 for 200 $pb^{-1}$

Using all information derived up to this point one can do a fit of the invariant mass distribution of opposite sign lepton pairs. For this fit a dataset of exactly 200  $pb^{-1}$  has been analysed to perform one pseudo experiment. The final statistical error is then derived by a set of toy Monte Carlo experiments as described in the next section.



Figure 10: The combined fit at LM0 for 200 pb<sup>-1</sup> is shown in (a). The red curve represents the SUSY signal model, the light green curve is the background function and the dark green dashed line the Z contribution. The black points represent the MC events. In (b) the fit of the background function to the  $e\mu$  invariant mass distribution is shown.

The fit of the invariant mass distribution is shown in Fig. 10(a) and it yields a value of

$$m_{ll,max} = (48.0 \pm 1.4) \text{ GeV.}$$
 (13)

The derived number of signal events  $n_{sig} = 85 \pm 13$  agrees with the number of signal events from MC truth (Tab. 4). The theoretical endpoint  $m_{ll,theo} = 52.7$  GeV is 4,7 GeV away from the fitted value. This is taken into account in the systematic uncertainty which is discussed in Sec.6.6. Nevertheless the bias can be removed using a calibration as explained in App. D.

The background fit of the opposite sign opposite flavour lepton pairs is shown in Fig. 10(b). We obtain a total number of background events of

$$n_{Bkg} = 234 \pm 15,\tag{14}$$

which is in agreement with the expected number from MC truth. The number of Background events in the signal region from 0 to 80 GeV in  $m_{ll}$  yields a value of

$$n_{Bkg,0-80} = 144 \pm 12. \tag{15}$$

#### 6.3 Monte Carlo toy study

A MC toy study of each fit has been performed in the RooFit package. For each distribution 1000 toy Monte Carlo datasets have been simulated and fitted. With these 1000 fits the distribution of the mass edge and the error can be calculated as shown in Fig. 6.3. The mean value of the mass edge is centered at 48 GeV which is 4.5 GeV away from the theoretical endpoint. This bias is not introduced by the fit itself as indicated by the MC toy study of the fit model as discussed in App. C. The bias is included as an extra systematic uncertainty in Sec. 6.5.





The MC toy studies of the other fits are shown in App. C

## 6.4 Significance

The significance at each LM point has been calculated once including shape information and once from simple event counting. Two utilities from the RooStats package [19] have been used to calculate the significance of the signal decay only is quoted and not a general SUSY excess over the Standard Model. The significance of the signal including shape information ( $\sigma_{Shape}$ ) uses signal and background probability density functions and the pseudo datapoints. Given the background only hypothesis (fitted from the opposite sign opposite flavour lepton distribution) the significance of the signal plus background hypothesis given the pseudo datapoints is calculated based on the profile likelihood of both hypothesis. We observe a nominal significance Of

$$\sigma_{Shape} = 6.7 \tag{16}$$

This significance is compared to the calculation from the number of signal and background events only ( $\sigma_{Count}$ ). The number of background events is determined by integration over the background model in the region between 0 and 80 GeV. The number of signal events is determined by the fit. Since the number of background events has been fitted from data the method  $Z_{Bi}$  described in [20] has been used to calculate the significance. Here we use a scaling factor  $\tau = 1$  to predict the number of background events and calculate the excess of events over this background. Using the number of events only we observe a nominal significance of

$$\sigma_{Count} = 4.4 \tag{17}$$

## 6.5 Systematic uncertainty on significance

The main source for systematic uncertainties is the jet energy scale uncertainty. It is assumed to be 5% at an integrated luminosity of 200 pb<sup>-1</sup> [13]. Due to the way the missing transverse energy corrections are implemented it is anti-correlated with MET:

$$\vec{E}_{T}^{corr} = \vec{E}_{T} - \sum_{i=1}^{N_{jets}} \left[ \vec{p}_{T_{i}}^{corr} - \vec{p}_{T_{i}}^{raw} \right].$$
(18)

This correlation has been taken into account. All jets above the threshold of 30 GeV which do not overlap with an electron are shifted by 5% in energy which propagates into the MET calculation.

Another source for systematic uncertainty is the arbitrariness of the background model. To evaluate the impact of the background model on the significance a different background model is choosen. We use a Landau function to fit the flavour symmetric background and compare the outcome of the fit. We observe a significance of  $5.3\sigma$  in case of a fit with a Landau function as shown in Fig. 13 in App. B.

To evaluate the impact of theoretical uncertainties on the cross-section of the background processes each sample has been scaled by  $\pm 10\%$ . The impact of this variation is found to be negligible since the dominant background is SUSY itself at LM0. Other sources for systematic uncertainties as luminosity have not been taken into account because the background is measured directly from data.

The minimal observed significance for 200 pb<sup>-1</sup> is 5.6  $\sigma$  from shape information and 4.4  $\sigma$  from pure event counting.

Table 5: Systematic uncertainties on the jet energy scale and their impact on the number of events at LM0. The quoted significance is in terms of standard deviations.

JES	N <sub>Sig</sub>	$N_{Bkg,0-80}$	$N_{Bkg}$	N <sub>Z</sub>	$\sigma_{Count}$	$\sigma_{Shape}$
0.05	$76 \pm 12$	$107 \pm 10$	$170\pm13$	$0\pm7$	4.4	5.6
$\pm 0.00$	$85 \pm 13$	$144 \pm 12$	$234\pm15$	$0\pm 4$	4.3	6.6
-0.05	$114 \pm 14$	$169 \pm 13$	$272\pm15$	$5\pm4$	5.3	8.0

Evaluating the uncertainty on the number of events due to the uncertainty of the jet energy scale one obtains the numbers shown in Tab. 5. The variation in the number of events leads to different observed significances in the experiment.

#### 6.6 Systematic uncertainty on endpoint

To evaluate the systematic uncertainty on the dilepton invariant mass endpoint the same uncertainty on the jet energy scale as in Sec. 6.5 is assumed. Additionally an electron energy scale uncertainty of 0.3% is assumed.

The main source of systematic uncertainties originates from the fit model itself. To evaluate this uncertainty the impact of the resolution model has been studied by varying the obtained fit values within their errors. The fit range has been varied as well as the initial values. The impact of these variations on the dilepton endpoint is displayed in Tab. 6.

The main impact results from the bias which is introduced by the fit itself. The bias can be found in each dataset. Therefore we split the LM0 dataset into 6 pieces of 200 pb<sup>-1</sup> per piece. In each of the subsamples the fit is repeated. The outcome of the fits is shown in Tab. 7. All fits are statistically compatible but biased to lower values of  $m_{ll}$  compared to the theoretical value of the endpoint.

We use the mean (48.2 GeV) of the fits as an estimator of the systematic bias on the fitted endpoint which is found to be 4, 5 GeV (compared to the theoretical value) at the discovery of this signal decay. The bias is compatible with the result of the MC toy study (Sec. 6.3).

Variation	Nominal	+ Var.	-Var.
Jet energy scale	$48.0 \pm 1.4$	$48.7\pm1.4$	$49.5\pm1.3$
Electron energy scale	$48.0 \pm 1.4$	$47.9 \pm 1.4$	$48.1\pm1.4$
Resolution model	$48.0 \pm 1.4$	$47.5\pm0.4$	-
Muon Efficiency	$48.0 \pm 1.4$	$48.1\pm1.4$	$47.9 \pm 1.4$
Electron Efficiency	$48.0 \pm 1.4$	$48.1\pm1.4$	$47.9 \pm 1.4$
Background model	$48.0 \pm 1.4$	$47.6 \pm 1.8$	-
Lepton acceptance	$48.0 \pm 1.4$	$49.3 \pm 1.7$	-

Table 6: Systematic uncertainties on the determination of the dilepton endpoint  $m_{ll,max}$ .

Table 7: Systematic bias on the determination of the dilepton endpoint  $m_{ll,max}$ .

Fit No.	0	1	2	3	4	5
$m_{ll,max}$ [GeV]	$48.0 \pm 1.4$	$47.9 \pm 1.1$	$46.6\pm0.7$	$51.0 \pm 1.0$	$47.0\pm0.9$	$48.6 \pm 1.3$

This error can be reduced when more data is available or other input values are used. With more data one can distinguish between the different decay modes using the information of the whole invariant mass ditribution and not only the endpoint.

## 6.7 Higher integrated luminosity

At the benchmark points LM1 and LM9 a higher integrated luminosity is necessary to measure the endpoint. Nevertheless the points can be discovered allready at an integrated luminosity of 500 pb<sup>-1</sup>. For the evaluation of the significance the same method as for LM0 is used (Sec. 6.5). The observed significances at LM1 and LM9 for  $500 \text{ pb}^{-1}$  are listed in Tab. 8.

Table 8: Significance of the signal at LM1 and LM9 for 500  $pb^{-1}$  for both methods. Both the minimal observed significance and the nominal significance are quoted.

	Nom. $\sigma_{Shape}$	Nom. $\sigma_{Count}$	Min. $\sigma_{Shape}$	Min. $\sigma_{Count}$
LM1	6.3	6.1	4.7	5.6
LM9	7.8	4.8	6.8	4.5

The fit of the invariant mass distribution at LM9 using an integrated luminosity of 1 fb<sup>-1</sup> is shown in Fig. 12(a). It yields a value of

$$m_{ll,max} = (61.4 \pm 0.7_{stat.} \pm 0.9_{syst.}) \text{ GeV.}$$
 (19)

The theory value of  $m_{ll,max} = 62,7$  GeV is underestimated by 1.5 GeV.

At LM1 a triangle is used as signal model and the fit of the invariant mass distribution at LM1 is shown in Fig. 12(b). It yields a value of

$$m_{ll,max} = (77.2 \pm 0.9_{stat.} \pm 1.0_{syst.}) \text{ GeV.}$$
 (20)

The thoretical endpoint of  $m_{ll,max} = 78.2 \text{ GeV}$  is reproduced within the statistical error.

At LM0 the fit to the invariant mass distribution using an integrated luminosity of 1 fb<sup>-1</sup> is shown in Fig. 12(c). It yields a value of

 $m_{ll,max} = (48.0 \pm 0.7_{stat.} \pm 1.2_{syst.}) \text{ GeV.}$  (21)

One can see that the bias towards lower values in the endpoint is still visible.

## 7 Conclusion

A significant excess of SUSY opposite sign same flavour lepton pairs can be found within the first 200 pb<sup>-1</sup> at LM0. The signal provides a quite robust signature and the background determination directly from data is possible. The nominal observed significance of the signal using shape information is  $6.7 \sigma$  and the minimal observed significance using only the number of observed events yields a minimal value of  $4.3\sigma$ .



Figure 12: Final fits to the opposite sign same flavour invariant mass distribution at LM9 (a) where the signal model consists of a quadratic term, at LM1 (b) where the signal model is a triangle and at LM0 (c), for 1  $fb^{-1}$ .

At the other studied benchmark points we observe a nominal significance using shape information of  $6.3 \sigma$  (LM1) and  $7.8 \sigma$  (LM9) for 500 pb<sup>-1</sup>. The significance using only the number of observed events yields a value of  $6.1 \sigma$  for LM1 and  $4.8 \sigma$  for LM9, respectively.

We presented an unbinned maximum likelihood fit of the dilepton invariant mass distribution (corrected for the diefference in muon and electron reconstruction efficiency) with a data-driven resolution determination. At LM0 the combined fit of the diletonic endpoint is possible with 200 pb<sup>-1</sup>. We obtain a value of

$$m_{ll,max} = (48.0 \pm 1.4_{stat.} \pm 2.2_{syst.} \pm 4.5_{bias.}) \text{ GeV.}$$
 (22)

The main systematic bias arises from the fit model itself, which leads to a bias towards lower values since at LM0 the sharp endpoint is not present.

At the benchmark points LM9 and LM1 the endpoint can be measured with an integrated luminosity of 1  $fb^{-1}$ .

## 8 Acknowledgements

We would like to thank the SUSY conveners Oliver Buchmuller and Jeff Richman as well as the conveners of the leptonic SUSY subgroup Frederic Ronga and David Stuart for the useful discussion and comments during this work.

## References

- [1] B.C. Alanach, "SOFTSUSY: a program for calculating supersymmetric spectra", Comput.Phys.Commun. 143:305-331, 2002
- [2] A. Djouadi et al, *Decays of Supersymmetric Particles: the program SUSY-HIT (SUspect-SdecaY-Hdecay-InTerface)* ActaPhys.Polon. B38:635-644, 2007.
- [3] T. Sjostrand et al, "PYTHIA 6.4 Physics and Manual", JHEP 0605:026, 2006
- [4] W. Beenakker et al, "Squark and Gluino Production at Hadron Colliders", Nucl. Phys. B492:51-103, 1997
- [5] The CMS Collaboration, "Study of the Z production in association with jets in proton-proton collisions at  $\sqrt{s} = 10$  TeV with the CMS detector at the CERN LHC", CMS Physics Analysis Summary JME-08-006
- [6] J. Alwall et al, "MadGraph/MadEvent v4: The New Web Generation", JHEP 0709:028, 2007
- [7] M. Cacciari et al, "Updated predictions for the total production cross sections of top and of heavier quark pairs at the Tevatron and at the LHC", JHEP 0809:127, 2008
- [8] S. Frixione, M.L. Mangano, "How accurately can we measure the W cross section?", JHEP 0405:056, 2004
- [9] V+jets wiki, https://twiki.cern.ch/twiki/bin/view/CMS/VplusJets
- [10] M. Mulders et al. "Muon Identification in CMS", CMS Analysis Note 2008/098
- [11] J. Branson et al. "A cut based method for electron identification in CMS", CMS Analysis Note 2008/082
- [12] G. P. Salam, G. Soyez, "A practical Seedless Infrared-Safe Cone jet algorithm", JHEP 0705:086, 2007
- [13] The CMS Collaboration, "Plans for Jet Energy Corrections at CMS", CMS Physics Analysis Summary JME-07-002
- [14] The CMS Collaboration, " $E_T$  performance in CMS", CMS Physics Analysis Summary JME-07-001
- [15] The CMS Collaboration, "Measuring Electron Efficiencies at CMS with Early Data", CMS Physics Analysis Summary EGM-07-001
- [16] Lean Trigger Menu V0.3 wiki, https://twiki.cern.ch/twiki/bin/view/CMS/TSG\_18\_II\_ 09
- [17] G. Karapostoli et al., "Dilepton + Jets + MET channel : Observation and Measurement of  $\tilde{\chi}_2^0 \rightarrow \tilde{\chi}_1^0 ll$ ", CMS Note 2008/038

- [18] W. Verkerke, D. Kirkby, "The RooFit toolkit for data modeling", arXiv:physics/0306116
- [19] RooStats wiki, https://twiki.cern.ch/twiki/bin/view/RooStats/WebHome
- [20] R. D. Cousins et al., *Evaluation of three methods for calculating statistical significance when incorporating a systematic uncertainty into a test of the background-only hypothesis for a Poisson process*, Nuclear Instruments and Methods in Physics Research A595:480-501, 2008

# **A** Datasets

The used dataset from CMSSW\_2\_1\_X are listed in Tab. 9. Each dataset is scaled to the disired luminosity if not illustrated differently. Additionally a k-factor has been applied for some of the datasets as explained in Sec. 5.2.

Table 9: Used CMSSW datasets.

DBS datasetpath	No. events	$\sigma_{LO}$ [pb]	Name
/SUSY_LM0-sftsht/Summer08_IDEAL_V11_v1/GEN-SIM-RECO	88872	110.00	SUSY LM0
/SUSY_LM1-sftsht/Summer08_IDEAL_V9_v1/GEN-SIM-RECO	200000	16.06	SUSY LM1
/SUSY_LM9t175-sftsht/Summer08_IDEAL_V9_v2/GEN-SIM-RECO	176672	11.12	SUSY LM9
/TTJets-madgraph/Fall08_IDEAL_V9_v1/GEN-SIM-RECO	1005253	319.00	tt+jets
/ZJets-madgraph/Fall08_IDEAL_V9_v1/GEN-SIM-RECO	1221378	3700.00	Z+jets
/AstarJets-madgraph/Fall08_IDEAL_V9_v2/GEN-SIM-RECO	117390	1260.00	Z+jets
/WJets-madgraph/Fall08_IDEAL_V9_v1/GEN-SIM-RECO	6230138	40000.00	W+jets
/VVJets-madgraph/Fall08_IDEAL_V9_v2/GEN-SIM-RECO	101778	11.80	VV+jets
/Wgamma/Summer08_IDEAL_V9_v1/GEN-SIM-RECO	102012	36.60	$W\gamma$
/Zgamma/Summer08_IDEAL_V9_v1/GEN-SIM-RECO	102300	11.00	$Z\gamma$
/JPsi/Summer08_IDEAL_V9_v1/GEN-SIM-RECO	1847135	941.28	$J/\Psi$
/Upsilon1S/Summer08_IDEAL_V9_v10/GEN-SIM-RECO	10203	1390.00	Υ
/Upsilon2S/Summer08_IDEAL_V9_v10/GEN-SIM-RECO	244341	47.15	Υ
/QCDpt80/Summer08_IDEAL_V9_v2/GEN-SIM-RECO	3327876	1934639.57	QCD
/QCDpt170/Summer08_IDEAL_V9_v3/GEN-SIM-RECO	2997540	62562.88	QCD
/QCDpt300/Summer08_IDEAL_V9_v1/GEN-SIM-RECO	3106580	3664.61	QCD
/QCDpt470/Summer08_IDEAL_V9_v1/GEN-SIM-RECO	2808000	315.51	QCD
/QCDpt800/Summer08_IDEAL_V9_v3/GEN-SIM-RECO	1027080	11.94	QCD

## **B** Different background model



Figure 13: The combined fit at LM0 for 200 pb<sup>-1</sup> is shown in (a). The red curve represents the SUSY signal model, the light green curve is the background function (Landau) and the dark green dashed line the Z contribution. The black points represent the MC events. In (b) the fit of the background Landau function to the  $e\mu$  invariant mass distribution is shown.

# C Toy MC study



Figure 14: Monte Carlo toy study of the fit in Fig. 12(a). The toy events are generated according to the invariant mass distribution at LM9 obtained from the full detector simulation.



Figure 15: Monte Carlo toy study of the fit in Fig. 12(b). The toy events are generated according to the invariant mass distribution at LM1 obtained from the full detector simulation.

To check if the bias is introduced by the probability density function used to perform the combined fit a Monte Carlo toy study has been carried out. The mass edge, the error and the pull distribution for 1000 pseudo experiments generated according to the combined propability density function. In the pull distribution a mean compatible with zero and a width of one is found, which indicates that the bias is not introduced by the fit itself.



Figure 16: Monte Carlo toy study of the combined PDF used in the fit in Fig. 10(a).



Figure 17: Monte Carlo toy study of the combined PDF used in the fit in Fig. 12(a).



Figure 18: Monte Carlo toy study of the combined PDF used in the fit in Fig. 12(b).

## **D** Calibration

To reduce the bias of the fit of the invariant mass distribution in case of a 3-body decay two calibrations have been used at LM0 and LM9. In each of the parameter regions five mSUGRA points have been simulated using the CMS Fast Simulation. We varied the mSUGRA parameter  $m_{1/2}$  to obtain different mass differences for the samples. At each of the simulated points the fit is performed and we plot the fitted invariant mass against the reconstructed invariant mass.



Figure 19: Calibration used at LM0.

Figure 20: Calibration used at LM9.

This method implies the knowledge of the region in parameter space from different processes, which will not be available at the discovery of SUSY. Later on the method provides a solution to reduce the bias in the fit.

The calibrated fit of the invariant mass distribution at LM0 (Fig. 10(a)) yields a value of

$$m_{ll,max} = (51.1 \pm 2.0) \text{ GeV},$$
 (23)

which agrees within the statistical error the theoretical value of 52.7 GeV.

The calibrated fit of the invariant mass distribution at LM9 (Fig. 12(a)) yields a value of

$$m_{ll,max} = (62.5 \pm 0.7) \text{ GeV},$$
 (24)

which agrees within the statistical error the theoretical value of 62.8 GeV.