Search for Supersymmetry in same sign dilepton final states with b Jets and Missing Energy at the LHC

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

University of California, Santa Barbara

W. Andrews, G. Cerati, D. Evans, F. Golf, I. MacNeill, S. Padhi, Y. Tu, F. Würthwein, A. Yagil, J. Yoo

University of California, San Diego

L. Bauerdick, I. Bloch, K. Burkett, I. Fisk, Y. Gao, O. Gutsche, B. Hooberman, S. Jindariani, J. Linacre

Fermi National Accelerator Laboratory, Batavia, Illinois

Abstract

We search for supersymmetry in same sign dilepton final states with at least two b jets and large $\not\!\!E_T$. The search is performed in a data sample collected with the CMS detector of pp collisions at a centre-of-mass energy of 7 TeV, corresponding to an integrated luminosity of 349 pb^{-1}. For these searches, the dominant background is from $t\bar{t}$ events. No excess above the standard model background expectation is observed. Upper limits at 95% confidence level are set on the number of observed events.

1 Introduction

The CMS Collaboration recently published a search for new physics in events with same-sign isolated dileptons, jets, and $\not\!\!E_T$ [1]. In that study the major background is from $t\bar{t}$ production, as shown in Fig. 1.

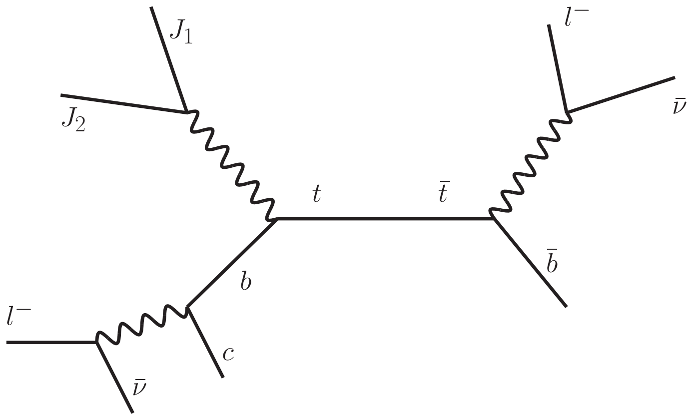

Figure 1: Diagram for $t\bar{t}$ decays giving rise to same-sign dilepton final states

The dominant source of same-sign dileptons in $t\bar{t}$ events are produced via, $t \to Wb$; where one of the leptons is from $W \to \ell\nu$ and the other originates from semi-leptonic b decays. An additional requirement on the number of b jets ≥ 2, is expected to reduce this background significantly as a b-quark can not produce an isolated lepton and at the same time provide a b-tag.

In this note we perform an inclusive search for new physics in events with two isolated, same-sign dileptons, in association with at least 2 b jets and $\not\!\!E_T$. This generic signature should also be sensitive to SUSY involving third generation fermions. For the purpose 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.

This note is organized as follows: in Section 2 we briefly outline the event selection used in this study along with event yields and backgound estimation as well as a discussion on the background. The description of systematics uncertanities on the acceptance is given in Section 3. Finally, in Section 4 we summarize the results followed by the conclusion in Section 5.

2 Search for same-sign dileptons with b jets

This analysis is based on the inclusive same-sign dilepton search documented in AN-2011/258 [2] and corresponds to an integrated luminosity of 349 pb^{-1}. In that study we searched for events with two isolated same-sign leptons in association with 2 additional jets and $\not\!\!E_T$. Here we re-use most of the baseline event selection [1)]. In addition, we require at least 2 b-tagged jets using Track Counting High Efficiency Medium (TCHEM) working point tagger [3]. We refer to TCHEM with the requirement that three of the tracks have IP significance > 3.3. For this tagger the b-tagging efficiency is 0.62 ± 0.01 and the acceptance of light flavor jets is 0.018 ± 0.004 [3].

2.1 Event Selection

The event selection for high p$_T$ and low p$_T$ baseline regions are briefly summerized as follows:

- At least two isolated same-sign leptons (ee, $e\mu$, and $\mu\mu$) with $|\eta| < 2.4$
  - For the high-p_T study, both leptons are required to be $p_T > 10$ GeV, with one of them $p_T > 20$ GeV.
  - For the low-p_T study, the electrons are required to be $p_T > 10$ GeV and muons $p_T > 5$ GeV.
- At least two particle flow jets tagged using TCHEM tagger with $p_T > 40$ GeV and $|\eta| < 2.4$ corrected with L1FastL2L3 corrections.

1) The additional Z veto is not applied in this study

- The selected jets must be separated from the lepton by $\Delta R > 0.4$.
- The scalar sum of momenta $H_T > 80$ ($H_T > 200$) GeV is required for high-p_T (low-p_T) analysis.
- $\not\!\!E_T > 30$ GeV.
- We remove dilepton events with invariant mass $M_{ll} < 5$ GeV.

More details are found in Reference [2].

2.2 Event Yields and Background Estimation

The results of this search in the above-mentioned kinematical region are summerized in Table 1, 2. Data-driven background predictions are used to estimate SM backgrounds. This is based on a combination of estimating "Tight-To-Loose ratio" (Fake Rate) and electron charge mis-reconstruction rate (Charge Flip rate). The probability for muons to be reconstructed with the wrong sign in the relevant momemtum range is negligible.

The event yields have the following characteristics:

- The contributions from rare processes such as $qqW^{\pm}W^{\pm}, WWW, t\bar{t}W$, and double parton $W^{\pm}W^{\pm}$ are negligibly small.
- The diboson backgrounds WW, WZ, and ZZ are found to be negligible.
- The prediction from fake rates includes the systematic error of 50%.
- The prediction from flip rates incldues the systematic error of 25%.
- The systematic errors are added when propagting the fake/flip rates into the total prediction.
- We do not consider any MC driven estimation for the final prediction.

Sample	$e^{\pm}e^{\pm}$	$\mu^{\pm}\mu^{\pm}$	$e^{\pm}\mu^{\pm}$	total
$t\bar{t}$	0.15 ± 0.09	0.25 ± 0.11	0.25 ± 0.11	0.65 ± 0.18
Single top	0.02 ± 0.02	0.02 ± 0.02	0.02 ± 0.01	0.06 ± 0.03
wjets	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00
DY	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00
VV	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00
Total MC	0.17 ± 0.09	0.27 ± 0.11	0.27 ± 0.11	0.71 ± 0.18
data (349 pb^{-1})	1	1	0	2
fake rate prediction				
single fake	0.38 ± 0.38	0.52 ± 0.37	0.00 ± 0.00	0.90 ± 0.52 (3 evts)
double fake	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.31 (0 evts)
fake prediction	0.38 ± 0.38	0.52 ± 0.37	0.00 ± 0.00	0.90 ± 0.60
flip rate prediction	0.05 ± 0.01	0	0.06 ± 0.02	0.11 ± 0.03
total fake rate prediction	0.38 ± 0.43	0.52 ± 0.45	0.00 ± 0.00	0.90 ± 0.69
total bkg prediction	0.43 ± 0.43	0.52 ± 0.45	0.06 ± 0.2	1.01 ± 0.75

Table 1: Data and Monte Carlo yields for the same-sign high-p_T dileptons with $H_T > 80$ GeV and $\not\!\!E_T > 30$ GeV. Uncertainties in the lower three rows also include the systematic uncertainties on the method used.

The dominant SM contribution in both low- and high-p_T selections are found to be from $t\bar{t}$ decays. The estimation is in a good agreement with the observation. We also note that the backgrounds with respect to the inclusive same sign dilepton search is suppressed by an order of magnitude due to the b-tag requirements.

Sample	$e^{\pm}e^{\pm}$	$\mu^{\pm}\mu^{\pm}$	$e^{\pm}\mu^{\pm}$	total
$t\bar{t}$	0.15 ± 0.09	0.25 ± 0.11	0.30 ± 0.19	0.70 ± 0.19
Single top	0.02 ± 0.02	0.02 ± 0.02	0.02 ± 0.01	0.06 ± 0.03
wjets	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00
DY	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00
VV	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00
Total MC	0.17 ± 0.09	0.27 ± 0.11	0.32 ± 0.19	0.76 ± 0.19
data (349 pb^{-1})	0	1	0	1
fake rate prediction				
single fake	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.56 (0 evts)
double fake	0.00 ± 0.00	0.51 ± 0.30	0.00 ± 0.00	0.51 ± 0.30 (3 evts)
fake prediction	0.00 ± 0.00	0.51 ± 0.30	0.00 ± 0.00	0.51 ± 0.63
flip rate prediction	0.01 ± 0.004	0	0.02 ± 0.006	0.03 ± 0.01
total fake rate prediction	0.00 ± 0.00	0.51 ± 0.68	0.00 ± 0.00	0.51 ± 0.68
total bkg prediction	0.01 ± 0.004	0.51 ± 0.68	0.02 ± 0.006	0.54 ± 0.68

Table 2: Data and Monte Carlo yields for the same-sign low-p_T dileptons with $H_T > 200$ GeV and $\not\!\!E_T > 30$ GeV. Uncertainties in the lower three rows also include the systematic uncertainties on the method used.

2.3 Discussion of Backgrounds

As shown earlier, the primary source of background events are from $t\bar{t}$ decays, which are estimated using the fake rate method. Here we further investigate by performing a closure test on a large $t\bar{t}$ sample [2)] corresponding to a luminosity normalization of 1fb^{-1}. The fake rates for electrons and muons are determined from the QCD sample [2]. This test is meant to check if the fake rate, as determined from the QCD events can be applied to $t\bar{t}$ with at least 2 b-tagged jets.

We classify them as follows, based on truth matched to their "parents":

- Type-I: both leptons originate from real W (including $W \to \tau \to l$) bosons, one with mis-reconstructed charge.
- Type-II a): one of the leptons is from a real W and the other originates from heavy flavor sources (b, c).
- Type-II b): one of the leptons is from a W and the other is a fake lepton from light flavor sources.
- Type-III: both leptons are fakes.

Same Sign Leptons	Total	Type-I	Type-II	Type-II a)	Type-II b)	Type-III
ee	0.31±0.07	0.00±0.00	0.31±0.07	0.11±0.04	0.21±0.05	0.00±0.00
$\mu\mu$	0.26±0.06	0.00±0.00	0.26±0.06	0.22±0.05	0.04±0.04	0.00±0.00
$e\mu$	0.57±0.09	0.00±0.00	0.57±0.09	0.37±0.07	0.21±0.05	0.00±0.00
total	1.15±0.13	0.00±0.00	1.15±0.13	0.70±0.10	0.45±0.08	0.00±0.00

Table 3: Expected number of $t\bar{t}$ events, of various types in 1 fb^{-1} of integrated luminosity. Uncertainties are from MC statistics.

Same Sign Leptons	Total	Type-I	Type-II	Type-II a)	Type-II b)	Type-III
ee	0.39±0.03	0.00±0.00	0.39±0.03	0.20±0.02	0.19±0.03	0.00±0.00
$\mu\mu$	0.36±0.03	0.00±0.00	0.36±0.03	0.30±0.03	0.06±0.01	0.00±0.00
$e\mu$	0.76±0.05	0.00±0.00	0.76±0.05	0.54±0.04	0.22±0.03	0.00±0.00
total	1.51±0.06	0.00±0.00	1.51±0.06	1.04±0.05	0.47±0.04	0.00±0.00

Table 4: Predicted number of $t\bar{t}$ events, of various types in 1 fb^{-1} of integrated luminosity. Uncertainties are from MC statistics.

[2)] The POWHEG sample TTToLNu2Q2B_7TeV-powheg-pythia6_Spring11-PU_S1_START311_V1G1-v1 is only used for the closure test. In this sample the $W \to l\nu$ is already forced at the generator level.

Tables 3 and 4 show the estimated and predicted event yields. We predict within 50% the Type-II contributions, where one of the leptons is from a real W and the other is a fake lepton. Within the Type-II the largest source of uncertanity is due to overprediction of Type-II a) component of the background, where we expected one of the leptons to be originated from heavy flavor sources (b, c).

Another possible source of $t\bar{t}$ background could be due to the residual jet from the semi-leptonic $b \to c$ decays. We expect this residual b-jet to be vetoed using the overlap removal between the lepton and the jet. In order to study this effect, we relax the ΔR requirement between the lepton and the b-tag jet.

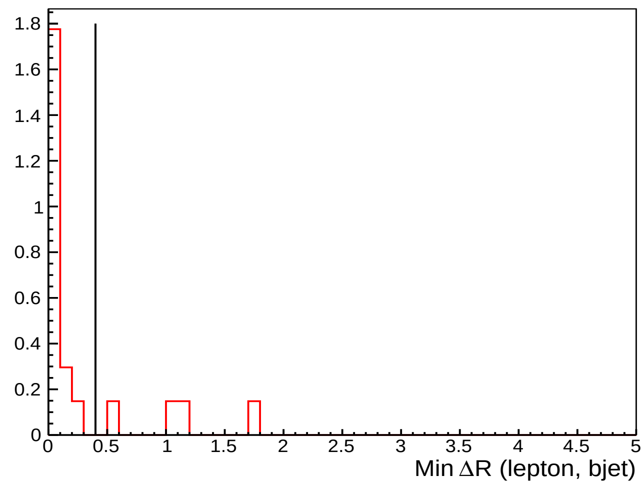

Figure 2: Minimum ΔR between the lepton and the b-tag jet in $t\bar{t}$ decays.

Fig. 2 shows that the bulk of such events in $t\bar{t}$ decays are well within the $\Delta R < 0.4$ between lepton and the b-jet.

3 Systematic Uncertainties

Systematic uncertainties arise from uncertainties on event selections expected in simulation compared to the actual performance of the detector. As this search is in many ways similar to the inclusive same-sign dilepton search, we also expect most of the systematic uncertainties to be applicable in this case. In this section, we mainly focus on the uncertainties due to the b-tagging requirement.

For this search without a well-defined signal, we evaluate the systematics with reference to the SUSY benchmark point LM9, as well as opposite sign $t\bar{t}$ simulation, both of which have b-enriched event topologies. The CMS benchmark point LM9 defines the common scalar mass (m0) $= 1.45$ TeV, the common gaugino mass (m1/2) = 175 GeV, the ratio of the Higgs expectation values (tanβ) $= 10$, tri-linear coupling (A0) = 0 and the sign of the Higgsino mass parameter $(\mu) > 0$. This produces heavy squarks with light gluinos leading to several heavy flavor final states.

The b-tagging efficiency as well as the systematic uncertainties are recently being approved by the BTV POG group for 2011 data [3]. In that study they provided the uncertanities on b-tagging efficiency, as well as scale factors (SF) for $t\bar{t}$ events. In Fig. 3 we compare the leading jet p_T normalized distribution for LM9 using the same-sign dilepton selection. For the $t\bar{t}$ events, we explicitely use opposite-sign dileptons in order to gain statistics. Given that the bulk of the p_T range accessable by LM9 is also covered by the $t\bar{t}$ decays, we consider 8.0% as the systematic uncertainty on the efficiency (per b-jet) from PtRel measurements [3]. Thus, the measurement of scale factors (SFs) for $t\bar{t}$ is applicable to LM9.

A complete summary of systematic uncertainties is given in Table 5.

4 Results

In absence of any significant deviation from the predicted background, we set 95% CL. on the number of observed events. Two statistical methods have been used for the upper limit. Both methods assume the uncertainties on signal and background are un-correlated and use a log-normal distribution for error pdfs.

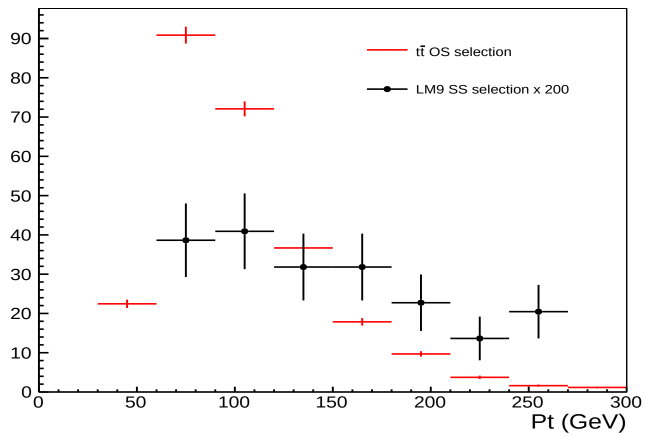

Figure 3: Differential distributions of leading b-tag jet p_T for the LM9 benchmark point and $t\bar{t}$ simulations using 349 pb^{-1} of luminosity normalization.

Table 5: Summary of systematic uncertainties on the signal selection and expectation. Reported values are fractional, relative to the total cross section. The values in parentheses are for electrons with p_T below 20 GeV and muons with p_T below 15 GeV for muons.

Source	ee	$\mu\mu$	$e\mu$	all
Lepton selection	11(15)%	11(15)%	11(15)%	11(15)%
Energy scale	5%	5%	5%	5%
ISR/FSR and PDF	2%	2%	2%	2%
b-tag selection	16%	16%	16%	16%
Total without luminosity	20(23)%	20(23)%	20(23)	20(23)%
Integrated luminosity	6%	6%	6%	6%
Total	21(24)%	21(24)%	21(24)%	21(24)%

The first method used to compute the upper limit is based on Bayesian statistics [**?**]. A posterior probability $p(r)$ is used as a function of the signal strength $r = \sigma/\sigma_{SM}$ assuming a uniform prior for r integrating the nuisance parameters associated with the uncertainties. The upper limit at 95% confidence level is then determined by integrating $p(r)$ to determine r', which satisfies $\int_{r'}^{\inf} p(r)dr = 0.05$.

We use the hybrid frequentist-bayesian CLs approach [4] as the second method. Although the two statistical approaches are not equivalent, in this case we get similar results.

- Upper limit using high-p_T analysis at 95% CL. with 24% signal systematic error using Bayesian approach = 6.1
- Upper limit using high-p_T analysis at 95% CL. with 24% signal systematic error using CLs = 5.8
- Upper limit using low-p_T analysis at 95% CL. with 24% signal systematic error using Bayesian approach = 4.8
- Upper limit using low-p_T analysis at 95% CL. with 24% signal systematic error using CLs = 4.6

We use 6.1 and 4.8 events as the upper limit for the rest of this document for high- and low-p_T analyses.

5 Conclusion

In conclusion, the first search using same-sign dileptons with b-jets and $\not{E}_T$has been presented. In the proton-proton collision data sample corresponding to an integrated luminosity of 349 pb^{-1} at $\sqrt{s}$ = 7 TeV, no significant deviations from the Standard Model expectations are observed. We use this data to set 95% CL. on the number of observed events.

References

[1] CMS Collaboration, “Search for new physics with same-sign di-leptons at the LHC”, arXiv:1104.3168, Submitted to JHEP.

[2] “Search for New Physics with Same-Sign Dileptons using the 2011 dataset of CMS”, CMS AN-2011/258

[3] “Measurement from data of efficiency and mistag rate of b-tagging algorithms using 2010 data”, PAS BTV-11-001 (Submitted)

[4] A.L. Read, CERN Report 2000-005 p. 81 (2000).