Search for New Physics in the Same Sign Dilepton final state 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 New Physics in the same sign dilepton final state with at least two b jets and $\not\!\!E_T$. 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 same-sign top production using events with same-sign isolated dileptons, jets, and $\not\!\!E_T$ [1]. In that study, as well as a closely related one [2] the major background is from $t\bar{t}$ production, as shown in Fig. 1.

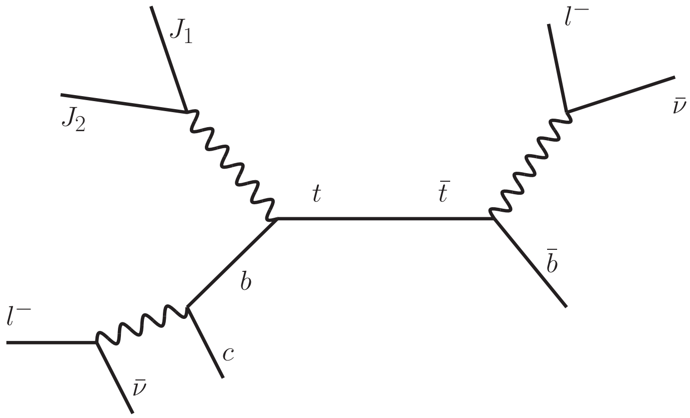

Figure 1: Diagram for $t\bar{t}$ decays giving rise to same-sign dilepton final states

The dominant source of same-sign dileptons in $t\bar{t}$ events are produced via, $t \rightarrow Wb$; where one of the leptons is from $W \rightarrow \ell\nu$ and the other originates from semi-leptonic b decays. We refer to the first as "real lepton" and the second as "fake lepton". 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 signature based search for events with two isolated, same-sign leptons, in association with at least 2 b jets and $\not\!\!E_T$. This generic signature should be sensitive to a wide variety of new physics scenarios leading to one or more same-sign top pairs in the final state, as well as the standard model production of $t\bar{t}W$. In addition to the inclusive search, we thus perform several dedicated searches for a variety of physics scenarios, including the standard model production of $t\bar{t}W$.

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 background estimation for the inclusive search. The description of systematics uncertainties on the acceptance is given in Section 3. In Section 4 we summarize the results of the inclusive search. Section 5 then refines this search for the various physics scenarios considered, and we conclude in Section 6.

2 Search for same-sign dileptons with b jets

This analysis is based on the same-sign dilepton search documented in AN-2011/258 [3] 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)] as summarized in Section 2.1 below. In addition, we require at least 2 b-tagged jets using Track Counting High Efficiency Medium (TCHEM) working point tagger [4]. We refer to TCHEM with the requirement that three of the tracks have IP significance > 3.3. For this tagger the expected b-tagging efficiency is 62% with a roughly 20% systematic uncertainty. The acceptance of light flavor jets is $\sim 2\%$ [4].

2.1 Event Selection

As mentioned previously, this search is a refinement of [3], and we thus discuss here only differences and briefly summarize the basic kinematics and triggers. For more details, we refer to [3].

- We separately look at the dilepton triggerred and dilepton plus H_T triggered samples.

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

- At least two isolated same-sign leptons (ee, $e\mu$, and $\mu\mu$) with $|\eta| < 2.4$.
- For the dilepton triggered samples, we require one lepton with p_T >20 GeV and a second lepton with p_T >10 GeV. We refer to this as the "high-p_T analysis".
- For the dilepton plus H_T triggered sample, we require electrons with p_T >10 GeV and muons with p_T >5 GeV. We refer to this as the "low-p_T analysis".
- At least two particle flow jets tagged using TCHEM tagger with $p_T > 40$ GeV and $|\eta| < 2.4$ corrected with L1FastL2L3 corrections.
- The selected jets must be separated from the lepton by $\Delta R > 0.4$.
- $\not\!\!E_T > 30$ GeV.
- We remove dilepton events with invariant mass $M_{ll} < 5$ GeV.

More details are found in Reference [3].

2.2 Event Yields and Background Estimation

The results of this search in the above-mentioned kinematical region are summerized in Table 1. As mentioned in the introduction, and quantified by this table, SM background is expected to be dominated by $t\bar{t}$ production with one real and one fake lepton. We estimate this background from the data itself using the "Tight-To-Loose ratio" (Fake Rate) method [5]. Electron charge mis-reconstruction is estimated by weighting opposite-sign dilepton events that pass all our cuts by a charge flip rate obtained form single electron Monte Carlo as described in detail in [3]. The probability for muons to be reconstructed with the wrong sign in the relevant momemtum range is negligible. Both of these techniques are described in more detail in [3]. Systematic errors on these two estimates are 50% and 25% respectively.

In addition, we use MC to estimate contributions from the following additional SM production processes. All of these are quite small.

- $qqW^{\pm}W^{\pm}, WWW, t\bar{t}W$, and double parton $W^{\pm}W^{\pm}$ with two real leptons in the final state.
- WZ and ZZ with two real leptons in the final state.
- $W\gamma$ with one real lepton and a photon conversion. This background is a priori not estimated by the fake rate method because the photon is generally isolated. In practice, this background is completely negligible.

Need to update the tables to include the remaining MC bkg as the MC becomes available !!!

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.

We have visually scanned all the events in data and provide details in Section 4.

2.3 Discussion of Background Expectation From MC

Three MC studies are presented. First, we show the origin of fake leptons in MC. Second, we show explicitly the degree to which the fake rate method accurately predicts the fake lepton background in $t\bar{t}$ MC, and third, we present evidence for our assertion that a b-quark can not simultaneously provide a b-tag and produce a fake lepton.

We use a large $t\bar{t}$ sample [2)] normalized to 1 fb^{-1} for these studies. The fake rates were obtained from QCD MC as described in detail in [3].

We classify $t\bar{t}$ background events based on truth matching to their "parent parton" as either "Heavy Flavor" or "Light Flavor". Charm quarks from W decay are classified as heavy flavor. It is thus possible to have three heavy quarks in the same event, one of which provides the isolated lepton, the two others the b-tags.

As is shown in Table 3 about 60% (40%) of the fake leptons are from heavy (light) flavor.

[2)] The POWHEG sample TTToLNu2Q2B_7TeV-powheg-pythia6_Spring11-PU_S1_START311_V1G1-v1

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.

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.

Table 4 shows the same breakdown for the fake rate prediction. Here we use an MC sample where one of the W's is forced to decay leptonically while the other is not allowed to decay leptonically. We use this sample as it has x10 the luminosity equivalent than the standard $t\bar{t}$ sample, thus providing higher statistics for this test. We observe an overprediction of 30% which appears to be primarily due to overpredicting the fakes from heavy flavor.

Finally, Figure 2 shows the minimum ΔR between one of the two b-tagged jets and the fake lepton in the MC after relaxing the jet veto cone around the lepton. The black line indicates the jet veto cone size. Clearly, a b-quark can not simultaneously provide a b-tag and an isolated lepton.

3 Acceptance Systematics

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 [3], our treatment of efficiency systematics parallels the one in that analysis. In this section, we briefly summarize those results, and describe the uncertainties due to the b-tagging requirement.

For the inclusive 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

Same Sign Leptons	Total	Heavy Flavor	Light Flavor
ee	0.31±0.07	0.11±0.04	0.21±0.05
$\mu\mu$	0.26±0.06	0.22±0.05	0.04±0.04
$e\mu$	0.57±0.09	0.37±0.07	0.21±0.05
total	1.15±0.13	0.70±0.10	0.45± 0.08

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

Same Sign Leptons	Total	Heavy Flavor	Light Flavor
ee	0.39±0.03	0.20±0.02	0.19±0.03
$\mu\mu$	0.36±0.03	0.30±0.03	0.06±0.01
$e\mu$	0.76±0.05	0.54±0.04	0.22±0.03
total	1.51±0.06	1.04±0.05	0.47±0.04

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

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 (μ) > 0. This produces heavy squarks with light gluinos leading to several heavy flavor final states.

For the b-tagging efficiency as well as the systematic uncertainties we refer to the work by the BTV POG group for 2011 data [4]. In that study they provide the uncertainties 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 [4]. We thus use the measurement of scale factors (SFs) for $t\bar{t}$ for the inclusive search.

Once we have all the various signal MCs we mention in Section 5, we will replace Figure 3 with one that compares all of these models, and revisit the statement about b-tagging systematics. It is likely that we will adopt different systematics for different models, given the differences in jet p_T for the different models.

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

4 Results on Inlcusive Signature Search

This section is missing two things. First, we will add a $\not{E}_T$ vs H_T plot that shows the data and one of the models overlayed. Second, we will add a summary of the events we see. We expect to add this information once we update the results for the full data sample up to the technical stop on June 29th 2011.

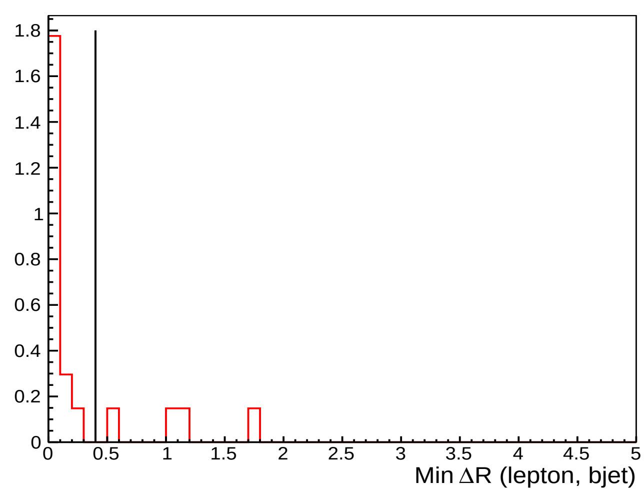

Figure 2: Minimum ΔR between the lepton and the b-tag jet in $t\bar{t}$ decays.

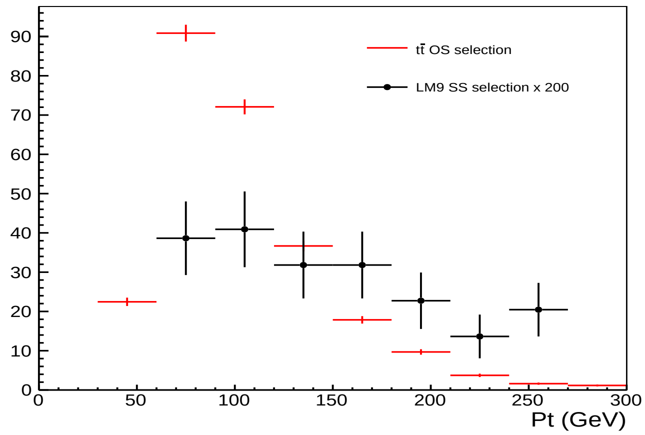

Figure 3: Differential distributions of leading b-tag jet p_T for the LM9 benchmark point and $t\bar{t}$ simulations using 349 pb^{-1} of luminosity normalization.

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)%

We see 2 (1) events in the high-p_T (low-p_T) analysis with a predicted background of 1.0± 0.8 (0.5± 0.7). 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.

The first method used to compute the upper limit is based on the Bayesian method [7]. A posterior probability $p(r)$ is used as a function of the signal strength $r = \sigma/\sigma_{SM}$ assuming a uniform prior for the signal strength 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 [6] 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 Searches for Specific Models

Our signature, two isolated same-sign leptons plus, at least two b-tagged jets, and $\not{E}_T$, is common to many different new physics scenarios, as well as standard model $t\bar{t}W$ production.

Here we refine our analysis to define dedicated signal regions for a few of these scenarios, and provide 95% C.L. upper limits on their respective model parameter space.

We expect to explore the following models for the final paper:

- Top pair production via t-channel Z' exchange as proposed by [10] and searched for by CMS with 2010 data [1]. This is in some ways a minimal model for our purposes, as it produces nothing other than same sign dileptons, two b-jets, and moderate $\not\!\!E_T$ typical for two leptonic W decays.
- A simplified model for same-sing top pair production as suggested in [12]. The production mechanisms here range from t-channel uu-scattering to tt, as in the previous example, to $tt\bar{u}\bar{u}$. The main difference to the previous model is different kinematics of the top quarks due to an intermeditate resonance $\eta^0 \rightarrow t\bar{u}$ decay. The mass of the η^0 is the main parameter in the simplified model.
- Standard model $t\bar{t}W$ production. Here the final state is same sign dileptons, two neutrinos, i.e. moderate $\not\!\!E_T$, plus one top quark that decays generically.
- $T_{5/3}$ fermion from Little Higgs models as suggested in [11] lead to a $t\bar{t}W^+W^-$ final state, thus providing one additional W over the previous example.
- SUSY gluino pair production with the gluino decaying to top and stop as suggested in [8][9]. Depending on stop mass, this leads to a final state of either $t\bar{t}t\bar{t}$ plus $\not\!\!E_T$, or $tt\bar{Q}\bar{Q}$ plus $\not\!\!E_T$ with Q being charm or beauty, and charge conjugates, as sources for the same sign dilepton plus two btags plus $\not\!\!E_T$. Independent of stop mass, final states with same sign dileptons plus four heavy flavor quarks plus $\not\!\!E_T$ are produced.

We will fill this in as we generate the samples, and figure out what minor changes to our baseline cuts each model requires to be sort of moderately optimized. In general, we expect changes in lepton p_T, $\not\!\!E_T$, H_T, and possibly jet p_T requirements, but no object selection changes.

6 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] “ ” ,this is our same sign top paper from 2010

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

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

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

[5] “Fake Rates for Dilepton Analyses”, AN2010/257

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

[7] whatever, some bayesian reference.

[8] http://arxiv.org/abs/hep-ph/0512284,SUSY Gluinos to light stop

[9] http://arxiv.org/abs/1004.2256,more SUSY Gluinos to light stop

[10] The same sign top, E. Berger et al.

[11] http://arxiv.org/abs/0801.1679, $T_{5/3}$ fermion pair production leading to $t\bar{t}W^+W^-$ final state.

[12] http://lhcnewphysics.org/l.006.00.r000, Simplified model by Felix Yu, UCI