Data driven background study for new physics searches with same sign dileptons at $\sqrt{s} = 10$ 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

Fermi National Accelerator Laboratory, Batavia, Illinois

Abstract

We discuss data driven methods to estimate the number of same sign Standard Model dilepton events in searches for new physics characterized by large $\not\!\!E_T$ and significant hadronic activity. For these searches, the dominant background is from $t\bar{t}$ events. The study provides a new method to estimate the charge mis-measurement of leptons. The residual flavor enriched background is estimated using the lepton fake rate method. We show sensitivity to several SUSY benchmark points using 100 pb^{-1} of integrated luminosity.

1 Introduction

In this note we describe data driven methods to estimate the backgrounds in searches for new physics in events with two high P_T, isolated, same sign leptons, large $\not\!\!E_T$, and significant hadronic activity. This generic signature is sensitive to several new physics scenarios such as SUSY. 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.

For a reasonable event selection, as shown in Section 4, the dominant background is from $t\bar{t}$ events. We categorize the total background into contributions from charge mis-measurement of leptons and events with fake leptons. The former is estimated using charge fake probability, as discussed in Section 6. The latter, which also includes heavy flavor decays is estimated using the lepton fake rate method described in Section 7.

This note is organized as follows. In Section 2 we list the Monte Carlo data samples, as well as the software tags used in the analysis. In Section 3 we describe the same sign dilepton event selection used in this study. The expected event yields for the dominant Standard Model processes as well as a few SUSY benchmark points are given in Section 4. A brief introduction to classifying the dominant background based on their origin is given in Section 5. In Section 6 we discuss the data driven procedure to estimate the charge mis-measurement followed by lepton fake rate method in Section 7. We apply the above mentioned data driven methods to the Standard Model cocktail in Section 8 and study the effect of applying the background prediction procedure in the presence of new physics. Finally, in Section 9 we summarize the results.

2 Data Samples

This study is based on the 2_2_X re-reco full simulation data samples listed in Table 1. The Standard Model (SM) data sets have been normalized to the cross-sections compiled by the top group [1]; for the SUSY data sets we used the cross-sections from the Summer 2008 production page [2].

```
/TTJets-madgraph/Fall08_IDEAL_V11_redigi_v10/GEN-SIM-RECO
/WJets-madgraph/Summer08_IDEAL_V11_redigi_v1/GEN-SIM-RECO
/ZJets-madgraph/Summer08_IDEAL_V11_redigi_v1/GEN-SIM-RECO
/WW/Summer08_IDEAL_V11_redigi_v1/GEN-SIM-RECO
/WZ_incl/Summer08_IDEAL_V11_redigi_v1/GEN-SIM-RECO
/ZZ/Summer08_IDEAL_V11_redigi_v1/GEN-SIM-RECO
/SingleTop_sChannel/Summer08_IDEAL_V11_redigi_v3/GEN-SIM-RECO
/SingleTop_tWChannel/Summer08_IDEAL_V11_redigi_v3/GEN-SIM-RECO
/SingleTop_tChannel/Summer08_IDEAL_V11_redigi_v3/GEN-SIM-RECO
/SUSY_LM0-sftsht/Summer08_IDEAL_V11_v1/GEN-SIM-RECO
/SUSY_LM*-sftsht/Summer08_IDEAL_V11_redigi_v1/GEN-SIM-RECO
```

Table 1: The data sets used in this study.

Monte Carlo events were analyzed with CMSSW_2_2_10 with the additional tags listed in Table 2

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 recently approved WW[3] and $t\bar{t}$ [4] cross-section analyses. A quick summary of the event selection is:

- The event is required to pass the single e or μ triggers.
- Two isolated, same sign leptons (ee, $e\mu$, and $\mu\mu$).
- Leptons must have $P_T > 10$ GeV and at least one of them must have $P_T > 20$ GeV.
- 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 L2L3 corrected caloJets with $P_T > 30$ GeV and $|\eta| < 2.4$.
- The scalar sum of the P_T of all jets passing the requirements above should be > 200 GeV.

V01-08-04 CondFormats/JetMETObjects
V00-06-02-09 DataFormats/METReco
V07-02-12-03 DataFormats/MuonReco
V01-08-02-01 JetMETCorrections/Algorithms
V01-08-15 JetMETCorrections/Configuration
V03-02-06 JetMETCorrections/JetPlusTrack
V02-09-02 JetMETCorrections/Modules
VB04-00-02-04 JetMETCorrections/Type1MET
V01-04-03 RecoJets/JetAssociationAlgorithms
V00-04-02-17 RecoMET/Configuration
V02-05-00-21 RecoMET/METAlgorithms
V02-08-02-17 RecoMET/METProducers
V03-26-04 DataFormats/PatCandidate
V05-05-09 PhysicsTools/PatAlgos
V03-06-03 PhysicsTools/PatUtils
V03-01-16 PhysicsTools/PFCandProducer
V09-30-03 PhysicsTools/HepMCCandAlgos
V05-13-02 DataFormats/HepMCCandidate

Table 2: Additional software tags used in this study.

- We require $\not\!\!E_T > 80$ GeV. Track Corrected MET (tcMET) [5] is used as a measure for $\not\!\!E_T$.

The details of the lepton and trigger selections are given below.

3.1 Electron Selection

- The electron ID is the "e-gamma category based tight", with small modifications to account for changes between CMSSW 1_6_X and 2_2_X; see Reference [3] for details.
- 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 2_2_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 [6] using tracks within cone of 0.3 of the candidate electron:
  - $|\Delta \cot\theta| < 0.02$; the difference between cotangent azimuthal angles of tracks parallel to each other.
  - $|d_{2d}| < 0.02$ cm; the two dimensional distance between points within nearest tracks.

3.2 Muon Selection

- Must be a global muon **and** a tracker muon [7].
- GlobalMuonPromptTight (global χ^2/ndof$<$10) [8].
- At least 11 valid hits for the silicon track [8].
- $|d_0| < 200\ \mu m$ (from silicon track, corrected for beamspot).
- Minimum ionizing: EcalVetoEnergy $<$ 4 GeV and HcalVetoEnergy < 6 GeV [9].
- 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 2_2_X, and are calculated in a cone of 0.3.

3.3 Trigger Selection

We use 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. These triggers are expected to be present in the data taking trigger table.

3.4 Selection due to charge mis-reconstruction

One of the main backgrounds to the same sign isolated dilepton signature consists of SM opposite sign dilepton events, where the charge of one of the leptons is misidentified. Monte Carlo studies show that the muon charge misidentification rate is negligible, up to very high momenta. On the other hand, because of brehmstrahlung in the tracker, electrons have a significant probability of being reconstructed with the wrong sign. In order to reduce the contribution of charge mis-reconstruction, we have studied the charge mis-identification rates in a dedicated "electron gun" sample. The charge mis-identification rate is defined as the ratio of electrons with wrong reconstructed charge compared to the true charge over all reconstructed and truth matched electron candidates. The details of the study can be found elsewhere [10]. Electrons are selected with standard identification and

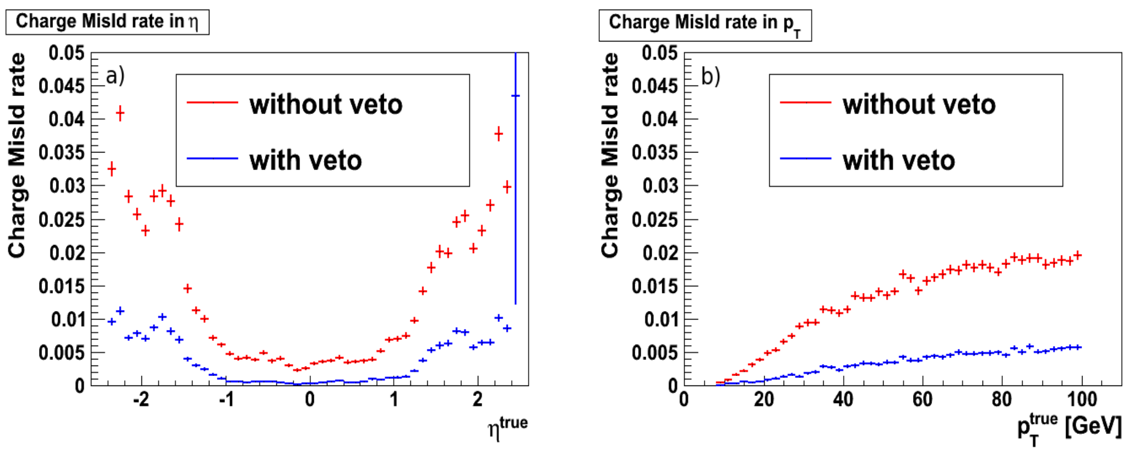

Figure 1: Charge mis-identification rate, as a function of a) η and b) p_T of the generated electron using "electron gun" sample . The solid (blue) distribution shows the rate after the veto; if the reconstructed charge is not same for GSF and CTF tracks.

isolation mentioned earlier. The GSF and GeneralTracks (CTF) are associated based on sharing of the hits in the silicon strip tracking detector. The electrons are vetoed, if the charge of GSF track associated with the electron is not same as the charge of the associated GeneralTrack. The electron is considered valid if it has a matched to a GSF track, but with no associated GeneralTrack. Figure 1 shows the charge mis-identification rate as a function of pseudorapidity (η) and transverse momentum (p_T) of the generated electron. The charge mis-identification rate is reduced by factor of ~ 3.9, with a moderate (2.1%) loss in electron reconstruction efficiency after the veto. The rate after the veto is found to vary between 0.04% to 1% with increase in pseudorapidity. This selection is applied as part of the standard selection for rest of this document.

4 Event Yields

Applying the event selections described in Section 3 to the Monte Carlo data sets described in Section 2, results in the expected event yields in 100 pb^{-1} listed in Table 3 and 4 for SM and SUSY benchmark points, respectively.

This event selection shows sensitivity to several SUSY points. The dominant contribution within SM is from $t\bar{t}$ decays.

5 Discussion of backgrounds

As shown in Section 4, the SM background to this study is dominated by $t\bar{t}$ events. The contribution from single-top, diboson production (WZ, ZZ), are small. The diboson backgrounds will be estimated from Monte Carlo.

Same Sign leptons	Total SM	$t\bar{t}$	tW	WZ	ZZ	WW	DY	Wjets
ee	0.45	0.44	0.00	0.00	0.01	0.00	0.00	0.00
$\mu\mu$	0.17	0.13	0.00	0.03	0.01	0.00	0.00	0.00
$e\mu$	0.48	0.39	0.04	0.04	0.01	0.00	0.00	0.00
total	1.10	0.96	0.04	0.07	0.03	0.00	0.00	0.00

Table 3: Expected number of SM events passing the event selection in 100 pb^{-1} of integrated luminosity.

Same Sign leptons	LM0	LM1	LM2	LM3	LM4	LM5	LM6	LM7	LM8	LM9
ee	9.93	1.95	0.21	1.45	0.49	0.17	0.45	0.22	0.72	0.50
$\mu\mu$	11.99	2.42	0.30	1.91	0.63	0.18	0.45	0.27	0.88	0.64
$e\mu$	22.52	4.55	0.48	3.09	1.18	0.34	0.86	0.42	1.62	1.26
total	44.44	8.92	0.99	6.45	2.3	0.69	1.76	0.91	3.22	2.40

Table 4: Expected number of SUSY benchmark simulated events passing the event selection in 100 pb^{-1} of integrated luminosity.

The Monte Carlo acceptances for these processes will be corrected for differences between data and Monte Carlo lepton identification and trigger efficiencies, as determined using the tag-and-probe method.

We study the same sign dileptons for the dominant $t\bar{t}$ background based on their origin. Typically, dileptons are produced via top decays, $t \to Wb$; where $W \to \ell\nu$, based on the truth matched to their "parents" we classify them as follows:

- Type-I: both leptons originate from real W (including $W \to \tau \to \ell$) bosons, one with mis-reconstructed charge.
- Type-II a): one of the lepton is from real W and the other originate from heavy flavor sources (b, c).
- Type-II b): one of the lepton is from W and the other is a fake lepton.
- Type-III: both leptons are not from W and are fakes or from heavy flavor sources.

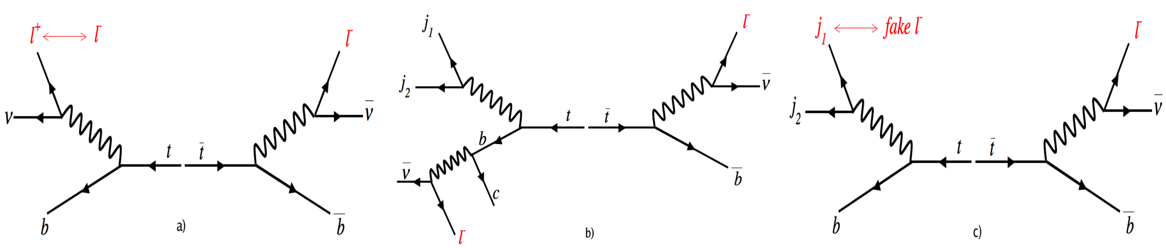

Figure 2: Classification of same sign dileptons from $t\bar{t}$ decays. a) lepton from W with mis-reconstructed charge; b) one of the lepton is from W, the other originating from heavy flavor sources; c) one of the lepton is from W, and the other is a fake lepton.

Figure 2 illustrates various contribution of background types in $t\bar{t}$ decays. The MC expectations to these contributions are given in Table 5.

Same Sign leptons	Total	Type-I	Type-II	Type-II a)	Type-II b)	Type-III
ee	0.44	0.09	0.35	0.17	0.17	0.00
$\mu\mu$	0.13	0.00	0.13	0.13	0.00	0.00
$e\mu$	0.39	0.13	0.26	0.26	0.00	0.00
Total	0.96	0.22	0.74	0.57	0.17	0.00

Table 5: Expected number of $t\bar{t}$ events for various types in 100 pb^{-1} of integrated luminosity.

It is interesting to note the following:

- About $\sim 23\%$ of the contribution is from Type-I (Charge mis-identification).
- Almost all of the charge mis-identification is from electrons, in ee and $e\mu$ channel.
- The bulk of the $t\bar{t}$ contribution in our event selection is from Type-II. The dominant among them is due to heavy flavor sources ($\sim 60\%$)
- No events with both leptons being fake (Type-III) are found.

In the following sections, we will briefly describe two different data driven approaches to estimate these contributions. The Type-I, contribution will be estimated using the "charge-flip rate" described below, where as Type-II part will be estimated used Lepton Fake rate method [11].

6 Data driven method to estimate Charge Mis-Identification (Type-I)

The goal of the method is to predict, in a data driven way, the charge mis-identification of the electrons. This is accomplished by measuring "Charge-flip Rate", $P_{ChargeFlip}$ using the Drell-Yan sample within the Z mass region ($76 < m_{ee} < 106$ GeV).

6.1 Description of Charge-flip rate

The method uses the probability of an electron to be reconstructed with a wrong charge as function of its p_T and η.

$$P_{ChargeFlip} = \frac{N_{Wrong}(P_T, \eta)}{N_{Total}(P_T, \eta)} \tag{1}$$

where $N_{Wrong}(P_T, \eta)$ is the number of wrongly charged electrons and $N_{Total}(P_T, \eta)$ is the total number of electrons in the sample. We select events with same sign (SS) and opposite sign (OS) dielectrons within the Z mass range, with both electrons passing the selection described in sections 3.1 and 3.4.

The number of wrongly charged electrons in the barrel part of the detector can be obtained:

$$N_{Wrong}(P_T, \eta) = SS(P_T, \eta) - k(P_T, \eta) * OS(P_T, \eta) \tag{2}$$

The normalization $k(P_T, \eta)$ is given by the ratio of doubly charged $SS_{++/--}$ electrons to the admixture of a distribution of correctly charge-identified electrons and incorrectly-charge identified OS electrons. Once we get the $P_{ChargeFlip}$ in the barrel, we use $P^{barrel}_{ChargeFlip}$ to obtain the rate for the endcap. The details of the procedure can be found in Ref. [12].

The Charge-flip rate obtained using the Z sample has limited coverage in p_T and η with respect to the phase space covered by the $t\bar{t}$ events. In order to extrapolate to the p_T and η range covered by the $t\bar{t}$ decays, we use a large "electron gun" MC. The Charge-flip rate is determined for the MC sample. The distribution is then validated using the data driven method.

Figure 3 shows the charge-flip rate as a function of $|\eta|$. The distribution from the "electron gun" sample agrees well with the expectation from data driven method using Z as well as MC truth matched Z events. It should be noted that the charge-flip rate roughly reproduces the Charge mis-identification rate we observed earlier in Figure 1 a).

6.2 Application of Charge-flip rate to our analysis

We now perform a rate test on $t\bar{t}$ MC events by relaxing cuts on $\not{E}_T$and jets. The test is meant to demonstrate that the Charge-flip rate as determined from the "electron gun" sample can be applied to $t\bar{t}$ events. Dilepton events are selected without any $\not{E}_T$and jet cuts using an integrated luminosity of 100 pb^{-1}. The observed event yield is obtained by selecting same sign electrons events. In order to get the estimation we use the following procedure:

- Select opposite sign dielectrons using the standard selection.
- Obtain the $P^1_{ChargeFlip}$ and $P^2_{ChargeFlip}$ for each electron for a given p_T and η.

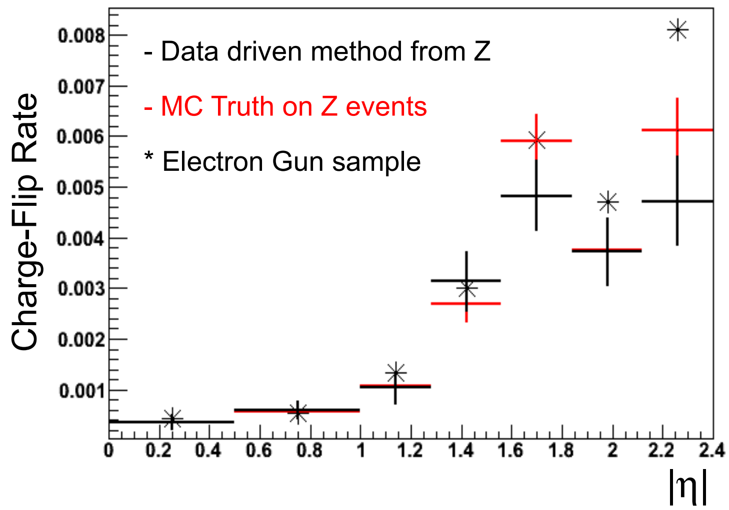

Figure 3: Charge-flip rate as a function $|\eta|$ for the "electron gun" sample (star) compared with distributions using data driven method from Z (black histograms) as well as truth matched Z events (red histogram).

- Assuming either of the electrons can flip signs, the flip probability is given by $F = P_{ChargeFlip}/(1 - P_{ChargeFlip})$.
- Weight each event by $weight * (F^1 + F^2)$.
- Add up all of the weights

Results of the Monte Carlo tests for event yields is given in Tables 6. From this study we can conclude that the Charge-flip rate does a very good job of reproducing the rate of charge mis-identification of electrons in $t\bar{t}$ events.

Sample	Event yield
$t\bar{t}$ (Observed)	2.4 $\pm$ 0.3
$t\bar{t}$ (Predicted)	2.1

Table 6: Monte Carlo test of the electron Charge-flip rate. Rates are normalized to 100 pb^{-1} of integrated luminosity.

We are now ready to apply the Charge-flip rate to our $t\bar{t}$ sample. The same sign dilepton sample will have three different contributions from Type-I, Type-II and Type-III events. The results of the application of the procedure outlined above is summarized in Table 7.

Same Sign leptons	Total	Type-I	Type-II	Type-II a)	Type-II b)	Type-III
ee (predicted)	0.05	0.05	0.00	0.00	0.00	0.00
$\mu\mu$ (predicted)	0.00	0.00	0.00	0.00	0.00	0.00
$e\mu$ (predicted)	0.07	0.07	0.00	0.00	0.00	0.00
total (predicted)	0.12	0.12	0.00	0.00	0.00	0.00

Table 7: The number of events predicted using Charge-flip rate in $t\bar{t}$ events for various types. Rates are normalized to 100 pb^{-1}.

Using Table 5 as observed and Table 7 as the prediction, we find that the Charge-flip method predicts the bulk of the Type-I events. As expected, the method does not predict any contribution from Type-II events. We consider the agreement to be satisfactory.

7 Data driven Method for Fake Lepton backgrounds (Type-II)

In this analysis, the primary source of fake leptons are from Type-II category. As shown in Table 5, roughly 3/4 of these events are expected to be from heavy flavor sources, while the remainder is from fake leptons (Type-II b)). This means that one needs to be careful in defining the lepton fake rates to make sure they predict both sources of fakes with sufficient accuracy.

In Reference [11] we described a data-driven method to predict the fake background in a dilepton analysis. This method is being applied by us also in the WW [3] as well as $t\bar{t}$ [4] analyses. Here we briefly summarize the method, and we apply it to the same sign dilepton study.

7.1 The fake rate definition

The method starts by defining a fake rate (FR) measured in QCD events. We use the Pythia QCD samples with $\hat{P}_T > 30, 80$ GeV. The fake rate is defined as the probability for a lepton passing loose cuts (Fakeable Object, FO) to pass the analysis cuts as a function of p_T and η. The idea is to then apply the FR to dilepton candidates passing loose cuts to obtain a prediction to the fake lepton contribution. The details of the applications of the FR are given in Section 7.2.

Fakeable Objects are defined as follows:

- Electron Fakeable Object, eFO:
  - GSFElectron with $p_T > 10$ GeV;
  - $|\eta| < 2.4$;
  - No reconstructed muon with $\Delta R < 0.1$;
  - Electron ID and Conversion rejection defined in Section 3.1 as well as veto used in Section 3.4;
  - Iso $<$ 0.4, where Iso=Sum/Max(20 GeV, P_T), and Sum = tkIso + hcalIso + Max(0 GeV, ecalIso - 2 GeV).
- Muon Fakeable Object, μFO:
  - Global and Tracker Muon with $p_T > 10$ GeV;
  - $|\eta| < 2.4$;
  - Global fit χ^2/ndof < 20;
  - $|d_0| < 200\ \mu m$ (from silicon track, corrected for beamspot);
  - Iso $<$ 0.4, where Iso=Sum/Max(20 GeV, P_T), and Sum = tkIso + hcalIso + ecalIso.

The FR for electrons and muons are determined from the QCD sample. We note that the absolute value of the FR is meaningless. It is the probability for a fake lepton passing loose identification requirements as well as isolation to also pass a tighter selection.

7.2 Application of lepton fake rate to our analysis

We evaluate the FR in $t\bar{t}$ Monte Carlo events. This test is meant to demonstrate that the FR as determined from the QCD events can be applied to $t\bar{t}$. In order to perform this test we define the following four event selections:

- $t\bar{t} \rightarrow WbWb \rightarrow \mu + e\ \nu\nu bb$:
  - Require a global muon with $p_T > 10$ GeV, truth matched to $W \rightarrow \mu$.
  - Require a same sign electron that passes all the standard identification and isolation requirements.
- $t\bar{t} \rightarrow WbWb \rightarrow \mu + (eFO \times FR)\ \nu\nu bb$:
  - Require a global muon with $p_T > 10$ GeV, truth matched to $W \rightarrow \mu$.
  - Require a same sign eFO; weight each event by the FR for the corresponding eFO.

- $t\bar{t} \to WbWb \to e + \mu\ \nu\nu bb$:
  - Require an electron with $p_T > 10$ GeV, truth matched to $W \to e$.
  - Require a same sign muon that passes all the standard identification and isolation requirements.
- $t\bar{t} \to WbWb \to e + (\mu FO \times FR)\ \nu\nu bb$:
  - Require an electron with $p_T > 10$ GeV, truth matched to $W \to e$.
  - Require a same sign μFO; weight each event by the FR for the corresponding μFO.

Sample	Event yield
$t\bar{t}$ with $\mu + e$ (Observed)	2.1 $\pm$ 0.3
$t\bar{t}$ with $\mu + (eFO \times FR)$ (Predicted)	2.9

Table 8: Monte Carlo test of the electron fake rate using 100 pb^{-1} of integrated luminosity.

Sample	Event yield
$t\bar{t}$ with $e + \mu$ (Observed)	2.4 $\pm$ 0.3
$t\bar{t}$ with $e + (\mu DO \times FR)$ (Predicted)	2.7

Table 9: Monte Carlo test of the muon fake rate using 100 pb^{-1} of integrated luminosity.

The Monte Carlo test for the electron FR consists of comparing event yields and distributions for $\mu + e$ and $\mu + (eFO \times FR)$. Similarly, for muon FR it consists of comparing event yields and distributions for $e + \mu$ and $e + (\mu DO \times FR)$. Results of the Monte Carlo tests for event yields are given in Tables 8 and 9. The uncertainties are from the MC statistics. From these studies we conclude that the QCD FR parametrization does a good job of reproducing the rate of fake electrons and muons in $t\bar{t}$ events.

Our procedure to estimate the fake contribution to our analysis is the following:

- Select lepton $+FO$ events where
  - one of the leptons passes all the standard identification and isolation requirements.
  - the other lepton is a FO but fails the standard identification and isolation requirements.
- The event passes all the standard kinematical cuts as outlined in Section 3
- Weigh each event by $FR/(1 - FR)$, where FR is the fake rate for the FO under consideration.
- Add up all the weights.

Same Sign leptons	Total	Type-I	Type-II	Type-II a)	Type-II b)	Type-III
ee (predicted)	0.21	0.01	0.20	0.15	0.05	0.00
$\mu\mu$ (predicted)	0.10	0.00	0.10	0.09	0.01	0.00
$e\mu$ (predicted)	0.31	0.01	0.30	0.26	0.04	0.00
total (predicted)	0.62	0.02	0.60	0.50	0.10	0.00

Table 10: The number of events predicted using lepton fake rate method in $t\bar{t}$ events for various types. Rates are normalized to 100 pb^{-1}.

The results of the application of the procedure outlined above is summarized in Table 10. We conclude the following in comparison with the observed events from Table 5:

- We predict within $\sim 20\%$ of the observed Type-II contributions.
- Within Type-II, the contributions with events from heavy flavor sources are largely predicted ($\sim 88\%$).
- The method introduces an overestimate for the true leptons in Type-I, which is at $\sim 2\%$ level, this is negligible compared to the associated statistical as well as systematic uncertainties.

8 Application of the data driven methods to the SM and SUSY benchmark points

We apply the two data driven procedures to predict the backgrounds in the $t\bar{t}$ -dominated SM sample. Table 11 shows the contribution of all SM background. The prediction and observation agree to within $\sim 30\%$.

Same Sign leptons	Total SM	$t\bar{t}$	tW	WZ	ZZ	WW	DY	Wjets
ee (observed)	0.45	0.44	0.00	0.00	0.01	0.00	0.00	0.00
ee (predicted)	0.27	0.26	0.01	0.00	0.00	0.00	0.00	0.00
$\mu\mu$ (observed)	0.17	0.13	0.00	0.03	0.01	0.00	0.00	0.00
$\mu\mu$ (predicted)	0.11	0.10	0.01	0.00	0.00	0.00	0.00	0.00
$e\mu$ (observed)	0.48	0.39	0.04	0.04	0.01	0.00	0.00	0.00
$e\mu$ (predicted)	0.39	0.38	0.01	0.00	0.00	0.00	0.00	0.00
total (observed)	1.10	0.96	0.04	0.07	0.03	0.00	0.00	0.00
total (predicted)	0.77	0.74	0.03	0.00	0.00	0.00	0.00	0.00

Table 11: Observed and predicted number of SM events passing the event selection in 100 pb^{-1} of integrated luminosity.

We also apply both of these methods to combination of SM and SUSY samples to derive the prediction. Table 12, show the contribution of observed and expected prediction to the sample. Typically one would compare observed with the prediction to look for excess in “signal” over the total background. The associated systematic uncertainties

Same Sign	SM+LM0	SM+LM1	SM+LM2	SM+LM3	SM+LM4	SM+LM5	SM+LM6	SM+LM7	SM+LM8	SM+LM9
Observed	45.54	10.02	2.09	7.55	3.4	1.79	2.86	2.01	4.32	3.50
Predicted	4.10	1.26	0.83	1.24	0.91	0.81	0.84	0.83	1.03	1.00

Table 12: Observed and predicted number of SM and SUSY events passing the event selection in 100 pb^{-1} of integrated luminosity.

are not discussed in this document. We plan to measure them in data. Overall, the sources of systematics can be categorized as detector effects, effects of modeling of the contributing processes, uncertainties of the data-driven background prediction methods. Systematic uncertainties from the lepton selection, ID, and reconstruction efficiencies will be estimated based on the corresponding systematics of the tag-and-probe method used to determine these efficiencies $Z \to \ell\ell$ in data. We will assess the uncertainty arising from the jet energy scale using γ/Z balance with jets. The uncertainties of the data-driven background estimate will be studied using a measure of “bias per lepton” as a function of the either charge fake candidate in charge-flip rate or FO in lepton fake rate selections. The current document focuses on reducing and measuring SM backgrounds using the aforesaid data driven methods.

9 Conclusion

We have studied two different data driven methods to predict the backgrounds for searches beyond standard model, in events with high P_T same sign dileptons (e or μ), significant hadronic activity, and high $\not\!\!E_T$. For these searches, the dominant background is from $t\bar{t}$. We characterized the background into different types based on charge mis-identification, leptons from heavy flavor sources, as well as fake leptons.

We used the charge-flip rate method to predict the number of wrongly charged leptons. The fakes including leptons from heavy flavor sources are predicted using lepton fake rates. The prediction using the methods is shown to work for the considered SM samples. Using both methods on the ensemble of the SM and SUSY samples, we show sensitivity towards several SUSY benchmark points using 100 pb^{-1} of integrated luminosity.

References

[1] http://www.ge.infn.it/t̃osi/cms/topMC.html.

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

[3] CMS AN-2009/042 and PAS EWK-09-002.

[4] CMS AN-2009/050 and PAS TOP-09-002.

[5] CMS AN-2009/022.

[6] http://indico.cern.ch/materialDisplay.py?contribId=3&materialId=slides&confId=58309.

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

[8] CMS AN-2008/098.

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

[10] http://indico.cern.ch/materialDisplay.py?contribId=3&materialId=slides&confId=64436.

[11] CMS AN-2009/041.

[12] P. Kalavase, et. al, Analysis Note in preparation.