Available on CMS information server

CMS AN -2010/257

The Compact Muon Solenoid Experiment

Analysis Note

05 September 2010 (v5, 24 December 2010)

Fake Rates for dilepton Analyses

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, J. Yoo

University of California, San Diego

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

Fermi National Accelerator Laboratory, Batavia, Illinois

Abstract

In this note we document the extraction of Fake Rates for dilepton analyses.

Contents

1 Introduction

In this note we document the extraction of Fake Rates (FR) for dilepton analyses. The method has been described and applied previously for 2009 Monte Carlo analyses[1][2][3][4], as well as ICHEP PASs for top[5] and SUSY[6]. Here we provide a self-contained description of the method as applied to the 2010 top dilepton cross-section analysis (TOP-10-001), the same sign analysis (SUS-10-004), the opposite sign analysis (SUS-10-007), and the WW analysis (EWK-10-009).

2 Data Sets, Luminosity

2.1 Data sets for FR for the $t\bar{t}$ analysis with 3.1 pb^{-1} (TOP-10-001)

We use the following data sets:

- `/InclusiveMu15/Spring10-START3X_V26_S09-v1/GEN-SIM-RECO`
- `/QCD_Pt30/Spring10-START3X_V26_S09-v1/GEN-SIM-RECO`
- `/Wenu/Spring10-START3X_V26_S09-v1/GEN-SIM-RECO`
- `/Wmunu/Spring10-START3X_V26_S09-v1/GEN-SIM-RECO`
- `/MinimumBias/Commissioning10-SD_EG-Jun14thSkim_v1/RECO`
- `/MinimumBias/Commissioning10-SD_Mu-Jun14thSkim_v1/RECO`
- `/EG_Run2010A-Jun14thReReco_v1/RECO`
- `/EG_Run2010A-Jul16thReReco_v2/RECO`
- `/EG_Run2010A-PromptReco-v4/RECO`
- `/Mu_Run2010A-Jun14thReReco_v1/RECO`
- `/Mu_Run2010A-Jul16thReReco_v1/RECO`
- `/Mu_Run2010A-PromptReco-v4/RECO`
- `/MinimumBias/Commissioning10-SD_JetMETTau-Jun14thSkim_v1/RECO`
- `/JetMETTau_Run2010A-Jun14thReReco_v2/RECO`
- `/JetMETTau_Run2010A-Jul16ththReReco_v1/RECO`
- `/JetMETTau_Run2010A-PromptReco-v4/RECO`

We use good data taking periods used by the top group, corresponding to an official certified JSON of August 13, as defined in Reference [7]. The integrated luminosity is 0.84 pb^{-1}.

2.2 Data sets for FR for the $t\bar{t}$ analysis with 36.1 pb^{-1} (TOP-10-005)

The following datasets were used to extract the fake rates:

- /EG/Run2010A-Sep17ReReco_v2/RECO
- /Electron/Run2010B-PromptReco-v2/RECO
- /EGMonitor/Run2010B-PromptReco-v2/RECO
- /Mu/Run2010A-Sep17ReReco_v2/RECO
- /Mu/Run2010B-PromptReco-v2/RECO

We use the good run list certified by the Top group [8] on November 5th corresponding to approximately 34.9 pb^{-1} of data in the Mu and EG datasets and 31.8 pb^{-1} for the EGMonitor dataset. Care was taken to avoid double counting between the EG and EGMonitor datasets.

2.3 Data sets for FR for the same sign analysis (SUS-10-004)

We use the following data sets in data:

- /EG/Run2010A-Sep17ReReco_v2/RECO
- /Electron/Run2010B-PromptReco-v2/RECO
- /EGMonitor/Run2010B-PromptReco-v2/RECO
- /Mu/Run2010A-Sep17ReReco_v2/RECO
- /Mu/Run2010B-PromptReco-v2/RECO

We use good data taking periods, corresponding to an official certified JSON covering the run ranges 135821 to 148058, as defined in Reference [7]. The integrated luminosity is 15.21 pb^{-1}.

For Monte Carlo, the fake rates for electrons are obtained from:

- /QCD_Pt_30to50_TuneZ2_7TeV_pythia6/Fall10-START38_V12-v1/GEN-SIM-RECO
- /QCD_Pt_50to80_TuneZ2_7TeV_pythia6/Fall10-START38_V12-v1/GEN-SIM-RECO
- /QCD_Pt_80to120_TuneZ2_7TeV_pythia6/Fall10-START38_V12-v1/GEN-SIM-RECO

These samples are weighted according to their corresponding luminosity equivalents of of $0.056pb^{-1}$, $0.47pb^{-1}$, and $3.8pb^{-1}$ when calculating fake rates.

As these are fairly small luminosity equivalents, we choose the muon enriched samples for our muon fake rates in Monte Carlo instead:

- /QCD_Pt-20_MuEnrichedPt-10_TuneZ2_7TeV-pythia6/Fall10-START38_V12-v1/GEN-SIM-RECO
- /QCD_Pt-20_MuEnrichedPt-15_TuneZ2_7TeV-pythia6/Fall10-START38_V12-v1/GEN-SIM-RECO

Here both samples have the same $\hat{P}_T$ and luminosity weighting is thus not necessary while calculating fake rates. However, when applying the fake rates, e.g. in the jet P_T test in Section 13.2, we use only the first sample in order to avoid an unrealistic overweighting of the $P_T >$ 15 GeV region. The equivalent luminosities of the two samples are $14pb^{-1}$ and $295pb^{-1}$ respectively.

We verified that the muon enriched samples are predominantly heavy flavor. As there is not sufficient statistics to do any electron tests with b-tagging in the regular QCD samples, we use the Pt-10 muon enriched sample as our source of b-tags for tests with electrons in Section 13.2.

For the closure tests in Section 13.3 we use:

- /TT_TuneZ2_7TeV-pythia6-tauola/Fall10-START38_V12-v1/GEN-SIM-RECO
- /WToENu_TuneZ2_7TeV-pythia6/Fall10-START38_V12-v1/GEN-SIM-RECO
- /WToMuNu_TuneZ2_7TeV-pythia6/Fall10-START38_V12-v1/GEN-SIM-RECO
- /WToTauNu_TuneZ2_7TeV-pythia6/Fall10-START38_V12-v1/GEN-SIM-RECO

2.4 Data sets for FR for the WW analysis (EWK-10-009)

We use the following datasets in data:

- /EG/Run2010A-Sep17ReReco_v2/RECO
- /Electron/Run2010B-PromptReco-v2/RECO
- /EGMonitor/Run2010B-PromptReco-v2/RECO

- /Mu/Run2010A-Sep17ReReco_v2/RECO
- /Mu/Run2010B-PromptReco-v2/RECO

We use good data taking periods, utilizing an official certified run JSON file covering runs in the range 135821 to 149442, as defined in Reference [8]. The integrated luminosity is 34.852 pb^{-1}.

The Monte Carlo electron fake rates are derived from:

- /QCD_Pt_30to50_TuneZ2_7TeV_pythia6/Fall10-START38_V12-v1/GEN-SIM-RECO

The Monte Carlo muon fake rates are derived from:

- /QCD_Pt-20_MuEnrichedPt-15_TuneZ2_7TeV-pythia6/Fall10-START38_V12-v1/GEN-SIM-RECO

For the Monte Carlo closure tests in Section 14.3 we use:

- /WToENu_TuneZ2_7TeV-pythia6/Fall10-START38_V12-v1/GEN-SIM-RECO
- /WToMuNu_TuneZ2_7TeV-pythia6/Fall10-START38_V12-v1/GEN-SIM-RECO
- /WToTauNu_TuneZ2_7TeV-pythia6/Fall10-START38_V12-v1/GEN-SIM-RECO

2.5 Data sets for FR for the same sign analysis (SUS-10-004)

We use the following data sets in data:

- /EG/Run2010A-Sep17ReReco_v2/RECO
- /Electron/Run2010B-PromptReco-v2/RECO
- /EGMonitor/Run2010B-PromptReco-v2/RECO
- /Mu/Run2010A-Sep17ReReco_v2/RECO
- /Mu/Run2010B-PromptReco-v2/RECO

We use good data taking periods, corresponding to an official certified JSON covering the run ranges 135821 to 148058, as defined in Reference [7]. The integrated luminosity is 15.21 pb^{-1}.

For Monte Carlo, the fake rates for electrons are obtained from:

- /QCD_Pt_30to50_TuneZ2_7TeV_pythia6/Fall10-START38_V12-v1/GEN-SIM-RECO
- /QCD_Pt_50to80_TuneZ2_7TeV_pythia6/Fall10-START38_V12-v1/GEN-SIM-RECO
- /QCD_Pt_80to120_TuneZ2_7TeV_pythia6/Fall10-START38_V12-v1/GEN-SIM-RECO

These samples are weighted according to their corresponding luminosity equivalents of of $0.056pb^{-1}$, $0.47pb^{-1}$, and $3.8pb^{-1}$ when calculating fake rates.

As these are fairly small luminosity equivalents, we choose the muon enriched samples for our muon fake rates in Monte Carlo instead:

- /QCD_Pt-20_MuEnrichedPt-10_TuneZ2_7TeV-pythia6/Fall10-START38_V12-v1/GEN-SIM-RECO
- /QCD_Pt-20_MuEnrichedPt-15_TuneZ2_7TeV-pythia6/Fall10-START38_V12-v1/GEN-SIM-RECO

Here both samples have the same $\hat{P_T}$ and luminosity weighting is thus not necessary while calculating fake rates. However, when applying the fake rates, e.g. in the jet P_T test in Section 13.2, we use only the first sample in order to avoid an unrealistic overweighting of the $P_T > 15$ GeV region. The equivalent luminosities of the two samples are $14pb^{-1}$ and $295pb^{-1}$ respectively.

We verified that the muon enriched samples are predominantly heavy flavor. As there is not sufficient statistics to do any electron tests with b-tagging in the regular QCD samples, we use the Pt-10 muon enriched sample as our source of b-tags for tests with electrons in Section 13.2.

For the closure tests in Section 13.3 we use:

- `/TT_TuneZ2_7TeV-pythia6-tauola/Fall10-START38_V12-v1/GEN-SIM-RECO`
- `/WToENu_TuneZ2_7TeV-pythia6/Fall10-START38_V12-v1/GEN-SIM-RECO`
- `/WToMuNu_TuneZ2_7TeV-pythia6/Fall10-START38_V12-v1/GEN-SIM-RECO`
- `/WToTauNu_TuneZ2_7TeV-pythia6/Fall10-START38_V12-v1/GEN-SIM-RECO`

3 The Fake Rate (FR) method

3.1 Introduction

The idea of the FR method is that QCD data is used to measure a lepton FR as a function of lepton P_T and $|\eta|$ which is defined as the probability for a lepton candidate passing loose cuts to also pass the analysis cuts. Leptons passing loose cuts are called "Fakeable Objects" (FO).

In a given analysis the FR is then used to estimate the background due to "fake" leptons[1)] as follows:

- Events are selected using all analysis cuts, except for the lepton selection. Dilepton backgrounds (case 1) with one real lepton and one fake lepton are estimated by selecting one lepton passing the full lepton selection and one failing it but passing the FO selection. Backgrounds with two fake leptons (QCD, case 2) are estimated by requiring both lepton candidates to pass the FO selection and fail the full selection.
- in case 1, each event is weighted by a factor FR/(1-FR), where FR is the Fake Rate for the FO.
- in case 2, each event is weighted by the product of two factors of FR/(1-FR), where FR is the Fake Rate for each of the two FO.
- The sum of the weights over the selected events is the background prediction.

Note that the method of case 1 applied to pure QCD backgrounds overestimates the background by a factor of two.

3.2 Measuring the FR on jet data, general considerations

Our initial plan was to measure the FR from lepton candidates in jet-triggered data. However, in CMS the low P_T jet data is so heavily prescaled that this approach is proving to be impractical.

Instead, we now mostly measure the FR using lepton-triggered data. When using lepton triggers, we then apply a jet requirement offline in place of requiring an online jet trigger (see below). We verify that these two methods give consistent results, as described in Section 10.

When using single jet triggers, in order to eliminate a possible trigger bias, we scan the list of HLT objects associated with the given jet trigger. If there is only one such object above threshold, we only consider lepton candidates well separated ($\Delta R > 1$) from it.

When using lepton triggers, we apply a requirement that there be at least a jet above some threshold in the event. Just as in the jet trigger case, if there is only one such jet above threshold, we only consider lepton candidates $\Delta R > 1$ away from it. Furthermore, we require that the lepton under consideration be matched with the corresponding

[1)] Here "fake" leptons refer to truly fake leptons as well as leptons from heavy flavor decays.

HLT lepton trigger object. Here matched means that the offline and HLT objects must be within ΔR of 0.4 from each other.

The lepton sample contains leptons from W and Z decays. One has to take this effect into account, otherwise the QCD FR will be overestimated, To minimize this problem, we reject events with tcMET $>$ 20 GeV or transverse mass $>$ 25 GeV as well as events with two opposite sign FOs that make a mass within 20 GeV of the Z mass.

Even after these requirements, Monte Carlo studies show that the muon FR is biased at high P_T. To eliminate this problem, we only calculate the FR up to $P_T = 35$ GeV and take a constant value above that. For backgrounds from Wjets, this is justified by the fact that the QCD backgrounds that we care about are concentrated at low P_T. These effects are described in Section 9. For backgrounds from top production in the same sign analysis the restricted P_T range leads to an additional systematics which is unavoidable, and is probed as part of the closure test in Section 13.3.

4 Event selection and Lepton ID for the $t\bar{t}$ analysis with 3.1 pb^{-1} (TOP-10-001)

4.1 Event cleanup for the sample used to derive FR for the $t\bar{t}$ analysis with 3.1 pb^{-1} (TOP-10-001)

We apply the standard "goodrun" list. In addition:

- Require BPTX bit 0.
- Veto beam halo bits (36,37,38,39).
- Scraping cut: if there are ≥ 10 tracks, require at least 25% of them to be high purity.
- Require at least one good vertex:
  - not fake
  - ndof > 4
  - $|\rho| < 2$ cm
  - $|z| < 15$ cm. (Yes, this is too tight. To be fixed in the next iteration.)

4.2 Lepton triggers for FR determination for the $t\bar{t}$ analysis with 3.1 pb^{-1} (TOP-10-001)

We use `HLT_Mu9` and `Ele10_LW_L1R`. In principle we could have used more electron triggers, *e.g.*, `Ele15_SW_L1R`, etc. This will be done in the next iteration.

4.3 Muon requirements for the $t\bar{t}$ analysis with 3.1 pb^{-1} (TOP-10-001)

These are the same requirements listed in Reference[9]. We list them here for completeness. Muon candidates are RECO muon objects passing the following requirements:

- $P_T > 10$ GeV.
- $|\eta| < 2.5$.
- Global Muon and Tracker Muon.
- χ^2/ndof of global fit < 10.
- At least 11 hits in the tracker fit.
- Transverse impact parameter with respect to the beamspot < 200 μm.
- $Iso \equiv E_T^{\text{iso}}$/Max(20 GeV, P_T) < 0.15. E_T^{iso} is defined as the sum of transverse energy/momentum deposits in ecal, hcal, and tracker, in a cone of 0.3.
- At least one of the hits from the standalone muon must be used in the global fit.

4.4 Electron requirements for the $t\bar{t}$ analysis with 3.1 pb^{-1} (TOP-10-001)

These are the same requirements listed in Reference[9]. We list them here for completeness. Electron candidates are RECO GSF electrons passing the following requirements:

- $P_T > 10$ GeV. (The $t\bar{t}$ analysis uses 20 GeV but for completeness we calculate FR down to 10 GeV).
- $|\eta| < 2.5$.
- SuperCluster $E_T > 10$ GeV.
- The electron must be ecal seeded.

- VBTF90 identification[11] with alignmnet correction and removing the $\Delta\eta$ requirement.
- Transverse impact parameter with respect to the beamspot < 400 μm.
- $Iso \equiv E_T^{\text{iso}}$/Max(20 GeV, P_T) < 0.15. E_T^{iso} is defined as the sum of transverse energy/momentum deposits in ecal, hcal, and tracker, in a cone of 0.3. A 1 GeV pedestal is subtracted from the ecal energy deposition in the EB, however the ecal energy is never allowed to go negative.
- Electrons with a tracker or global muon within ΔR of 0.1 are vetoed.
- The number of missing expected inner hits must be less than two[13].
- Conversion removal via partner track finding: any electron where an additional GeneralTrack is found with $Dist < 0.02$ cm and $\Delta \cot\theta < 0.02$ is vetoed[13].
- ECAL spike cleaning (aka Swiss-Cross cleaning) has been applied.

4.5 Muon FO requirements for the $t\bar{t}$ analysis with 3.1 pb^{-1} (TOP-10-001)

Fakeable Objects are defined starting from the full lepton selection by relaxing some combination of the identification and isolation requirements. Since most muons in QCD events are from heavy flavor decays, relaxing the identification requirements is not very useful. Thus our muon FO definition consists mainly of relaxing the isolation requirement.

Muon FO definition: relax the following muon requirements from Section 4.3:

- χ^2/ndof of global fit < 50 (was < 10).
- Transverse impact parameter with respect to the beamspot < 2 mm (was < 200 μm).
- $Iso < 0.4$ (was < 0.15).

4.6 Electron FO requirements for the $t\bar{t}$ analysis with 3.1 pb^{-1} (TOP-10-001)

In this case, we define three complementary fakeable objects (V1, V2, V3).

Electron V1 FO definition: relax the following electron requirements from Section 4.4:

- Remove the VBTF90 requirement.
- $Iso < 0.4$ (was < 0.15).
- The impact parameter cut was removed (used to be < 400 μm).

Electron V2 FO definition: relax the following electron requirements from Section 4.4:

- Remove the VBTF90 requirement.
- The impact parameter cut was removed (used to be < 400 μm).

Electron V3 FO definition: relax the following electron requirements from Section 4.4:

- $Iso < 0.4$ (was < 0.15).
- The impact parameter cut was removed (used to be < 400 μm).

To summarize the qualitative differences between the three choices:

- the V1 FR is an extrapolation in both ID and isolation;
- the V2 FR is an extrapolation in ID;
- the V3 FR is an extrapolation in isolation.

5 Event selection and Lepton ID for the $t\bar{t}$ analysis with 36.1 pb^{-1} (TOP-10-005)

5.1 Event cleanup for the sample used to derive FR for the $t\bar{t}$ analysis with 36.1 pb^{-1} (TOP-10-005)

Along with the good run list [8], the following event level cuts are applied:

- Scraping cut: if there are $\geq$ 10 tracks, require at least 25% of them to be high purity.
- Require at least one good vertex:
  - not fake
  - ndof > 4
  - $|\rho| < 2$ cm
  - $|z| < 24$ cm.

Note that we no longer require the BPTX bit 0 or veto on the beam halo bits (see Section 4.1) and have relaxed the requirement on the z-position of the vertex from 15 cm to 24 cm.

5.2 Lepton Triggers used for the FR determination for the $t\bar{t}$ analysis with 36.1 pb^{-1} (TOP-10-005)

The following triggers were used to extract the fake rates:

- Muon triggers: HLT_Mu9. This trigger was prescaled for runs > 145000 so most of the statistics comes from runs < 145000.
- Electron triggers for fake rate versions V1 and V2
  - For the EG dataset and runs < 141956, we use an OR of HLT_Ele10_LW_L1R, HLT_Ele10_SW_L1R and HLT_Ele15_LW_L1R.
  - For the EG dataset and runs < 145761, we use an OR of HLT_Ele15_SW_L1R and HLT_Ele20_SW_L1R.
  - For the EGMonitor dataset we use an OR of HLT_Ele10_SW_L1R, HLT_Ele17_SW_L1R, HLT_Ele10_SW_L1R_v2 and HLT_Ele17_SW_L1R_v2
- For the V3 fake rate, we can add triggers with an identification requirement as the V3 fake rate is not an extrapolation in id. We therefore use:
  - For the EG dataset, we use the same triggers as used for the V1 and V2 fake rates with the addition of HLT_Ele10_SW_EleId_L1R for runs between 141956 and 144114 and the addition of HLT_Ele17_SW_CaloEleId_L1R for runs < 147116.
  - For the EGMonitor dataset, we use the same triggers as used for the V1 and V2 fake rates with the addition of HLT_Ele17_SW_CaloEleId_L1R

5.3 Muon selection requirements for the $t\bar{t}$ analysis with 36.1 pb^{-1} (TOP-10-005)

The muon numerator and FO object selections are similar to those described in Section 4.3, the only difference being that the muon is required to have $|\eta| < 2.4$, instead of $|\eta| < 2.5$ for both the FO and numerator.

5.4 Electron selection requirements for the $t\bar{t}$ analysis with 36.1 pb^{-1} (TOP-10-005)

The electron numerator and FO object selections are almost the same as those described in Section 4.4, the only difference being that the electron's SuperCluster E_T is required to be > 15 GeV/c and the VBTF90 identification requirement, when applied, has the $\Delta\eta$ cut applied. The alignment correction is no longer necessary and is not done.

6 Event selection and Lepton ID for the same sign analysis (SUS-10-004)

6.1 Event cleanup for the sample used to derive FR for the same sign analysis

We apply the standard "goodrun" list. In addition:

- Scraping cut: if there are ≥ 10 tracks, require at least 25% of them to be high purity.
- Require at least one good vertex:
  - not fake
  - ndof > 4
  - $|\rho| < 2$ cm
  - $|z| < 24$ cm.

6.2 Lepton triggers for FR determination for the same sign analysis

The basic idea is to use single lepton triggers that do not bias the FO selection, i.e. are fully efficient for FO, or at least as efficient for FO as for the numerator selection.

For muons this is satisfied by `HLT_Mu9` which is part of the Mu dataset, even after it was prescaled. For electrons, the trigger selection is complicated by the changing trigger conditions, as well as triggers moving from EG to EGMonitor datasets. Moreover, V1 and V2 of our FR differ in the electron Id requirement between numerator and denominator, and thus require us to only use triggers that have no electron Id requirements. V3 is an extrapolation only in isolation, and we can thus use triggers that apply electron Id as well. This leads to the following selection of triggers, run ranges, and datasets for the FRs. This selection is a compromise between getting sufficient statistics and not having too many run range and dataset distinctions. We thus drop the EGMonitor dataset from consideration for Run2010A.

- For V1 and V2 electron FRs
  - For EG dataset and HLT_Ele10_LW_L1R or HLT_Ele10_SW_L1R or HLT_Ele15_LW_L1R use only run 141876 and below.
  - For EG dataset and HLT_Ele15_SW_L1R use all runs in that dataset.
  - For Run2010B use only EGMonitor and the .or. of the four triggers mentioned above.
- For V3 electron FRs
  - Use all the same dataset, run number, trigger combinations as for V1, V2 except for the additions and subtractions below.
  - In addition, for EG dataset in Run2010A allow HLT_Ele10_SW_EleId_L1R for the brief period it was available (Run 141956-144114)
  - In addition, for Electron dataset in Run2010B allow HLT_Ele17_SW_CaloEleId_L1R for the brief period it was available (Run 146428-147116).
  - In addition, for EGMonitor in Run2010B make sure to veto events that pass HLT_Ele17_SW_CaloEleId_L1R as those come in as double counts via some of the prescaled triggers in that dataset.
- Do not use any double lepton triggers.
- Do not use any triggers with isolation.

6.3 Muon requirements for the same sign analysis

These are the same requirements listed in Reference[10]. We list them here for completeness. Muon candidates are RECO muon objects passing the following requirements:

- $P_T > 10$ GeV.
- $|\eta| < 2.4$.

- Global Muon and Tracker Muon.
- χ^2/ndof of global fit < 10.
- At least 11 hits in the tracker fit.
- Transverse impact parameter with respect to the beamspot < 200 μm.
- $Iso \equiv E_T^{\text{iso}}$/Max(20 GeV, P_T) < 0.10. E_T^{iso} is defined as the sum of transverse energy/momentum deposits in ecal, hcal, and tracker, in a cone of 0.3.
- At least one of the hits from the standalone muon must be used in the global fit.
- ECAL veto deposit $\leq$ 4 GeV.
- HCAL veto deposit $\leq$ 6 GeV.

6.4 Electron requirements for the same sign analysis

These are the same requirements listed in Reference[10]. We list them here for completeness. Electron candidates are RECO GSF electrons passing the following requirements:

- $P_T > 10$ GeV.
- $|\eta| < 2.4$.
- SuperCluster $E_T > 10$ GeV.
- The electron must be ecal seeded.
- VBTF80 identification[11] .
- Transverse impact parameter with respect to the beamspot < 200 μm.
- $Iso \equiv E_T^{\text{iso}}$/Max(20 GeV, P_T) < 0.10. E_T^{iso} is defined as the sum of transverse energy/momentum deposits in ecal, hcal, and tracker, in a cone of 0.3. A 1 GeV pedestal is subtracted from the ecal energy deposition in the EB, however the ecal energy is never allowed to go negative.
- Electrons with a tracker or global muon within ΔR of 0.1 are vetoed.
- The number of missing expected inner hits must be zero[13].
- Conversion removal via partner track finding: any electron where an additional GeneralTrack is found with $Dist < 0.02$ cm and $\Delta \cot\theta < 0.02$ is vetoed[13].
- Require that all three charge measurements for a GSF electron agree; one from the charge of the GSF track, one from the charge of the CTF track associated to the GSF track and one, the so-called "supercluster charge", determined from the relative position of the supercluster with respect to the projected track from the pixel seed.

6.5 Muon FO requirements for the same sign analysis

Fakeable Objects are defined starting from the full lepton selection by relaxing some combination of the identification and isolation requirements. Since most muons in QCD events are from heavy flavor decays, relaxing the identification requirements is not very useful. Thus, our muon FO definition consists mainly of relaxing the isolation requirement.

Muon FO definition: relax the following muon requirements from Section 6.3:

- χ^2/ndof of global fit < 50 (was < 10).
- Transverse impact parameter with respect to the beamspot < 2 mm (was < 200 μm).
- $Iso < 0.4$ (was < 0.10).

6.6 Electron FO requirements for the same sign analysis

We define three complementary fakeable objects (V1, V2, V3).

Electron V1 FO definition: relax the following electron requirements from Section 6.4:

- Remove the VBTF80 requirement.
- $Iso < 0.4$ (was $<$ 0.10).
- The impact parameter cut was removed (used to be $<$ 200 μm).

Electron V2 FO definition: relax the following electron requirements from Section 6.4:

- Remove the VBTF80 requirement.
- The impact parameter cut was removed (used to be $<$ 200 μm).

Electron V3 FO definition: relax the following electron requirements from Section 6.4:

- $Iso < 0.4$ (was $<$ 0.10).
- The impact parameter cut was removed (used to be $<$ 200 μm).

To summarize the qualitative differences between the three choices:

- the V1 FR is an extrapolation in both ID and isolation;
- the V2 FR is an extrapolation in ID;
- the V3 FR is an extrapolation in isolation.

7 Event selection and Lepton ID for the WW analysis (EWK-10-009)

7.1 Event cleanup for the samples used to derive FR for the WW analysis

We apply the standard "goodrun" list. In addition:

- Scraping cut: if there are ≥ 10 tracks, require at least 25% of them to be high purity.
- Require at least one good vertex:
  - not fake
  - ndof > 4
  - $|\rho| < 2$ cm
  - $|z| < 24$ cm.

7.2 Lepton triggers for FR determination for the WW analysis

The basic idea is to use single lepton triggers that do not bias the FO selection, i.e. that are fully efficient for the FO, or at least as efficient for the FO as for the numerator selection.

For muons we satisfy this criterion utilizing `HLT_Mu9`, `HLT_Mu11` and `HLT_Mu15_v1` which are part of the Mu dataset.

For electrons the trigger selection is complicated by the changing trigger conditions, as well as triggers moving from the EG dataset to the EGMonitor dataset. Moreover, V1, V2 and V4 of our FR differ in the electron Id requirements between the numerator and denominator selections, and thus require that we only use triggers that have no, or looser, electron Id requirements. V3 is an extrapolation in isolation only, and we can thus use triggers that apply electron Id so long as isolation cuts are not applied. All that being said, for this first data taking period upon which the WW analysis is based, we keep things simple by using only triggers without Id and without isolation requirements for the full run range. This leads to the following selection of triggers, run ranges, and datasets for the FRs. This selection is a compromise between getting sufficient statistics and not having too many run range and dataset distinctions. We thus drop the EGMonitor dataset from consideration for Run2010A.

- For EG dataset and HLT_Ele10_LW_L1R or HLT_Ele10_SW_L1R or HLT_Ele15_LW_L1R use runs below run 141956.
- For EG dataset and HLT_Ele15_SW_L1R or HLT_Ele20_SW_L1R use run 145761 and below.
- For EGMonitor and HLT_Ele10_SW_L1R or HLT_Ele17_SW_L1R use runs below run 149181.
- For EGMonitor and HLT_Ele10_SW_L1R_v2 or HLT_Ele17_SW_L1R_v2 use run 149181 and above.
- Do not use any double lepton triggers
- Do not use any triggers with Id or isolation

7.3 Muon requirements for the WW analysis

Muon candidates are RECO muon objects passing the following requirements:

- $P_T > 20$ GeV.
- $|\eta| < 2.4$.
- Global Muon and Tracker Muon.
- At least 11 hits in the tracker fit.
- χ^2/ndof of global fit < 10.
- At least one of the hits from the standalone muon must be used in the global fit.

- At least one pixel hit.
- At least two muon chambers with hits.
- delta P_T / $P_T < 0.10$.
- Transverse impact parameter with respect to the primary vertex $< 200\ \mu$m.
- Longitudinal impact parameter with respect to the primary vertex < 1 cm.
- $Iso \equiv E_T^{\text{iso}}$/Max(20 GeV, P_T) < 0.15. E_T^{iso} is defined as the sum of transverse energy/momentum deposits in the ecal, the hcal, and the tracker, in a cone of 0.3.

7.4 Electron requirements for the WW analysis

Electron candidates are RECO GSF electrons passing the following requirements:

- $P_T > 20$ GeV.
- $|\eta| < 2.5$.
- VBTF80 identification[11].
- Transverse impact parameter with respect to the beamspot $< 200\ \mu$m.
- Longitudinal impact parameter with respect to the beamspot < 1 cm.
- $Iso \equiv E_T^{\text{iso}}$/Max(20 GeV, P_T) < 0.10. E_T^{iso} is defined as the sum of transverse energy/momentum deposits in the ecal, the hcal, and the tracker, in a cone of 0.3. A 1 GeV pedestal is subtracted from the ecal energy deposition in the barrel, however the ecal energy is never allowed to go negative.
- Electrons that share the same tracker track as a muon are vetoed.
- The number of missing expected inner hits must be zero[13].
- Conversion removal via partner track finding: any electron where an additional GeneralTrack is found with $Dist < 0.02$ cm and $\Delta \cot\theta < 0.02$ is vetoed[13].

7.5 Muon FO requirements for the WW analysis

Fakeable Objects are defined starting from the full lepton selection by relaxing some combination of the identification and isolation requirements. Since most muons in QCD events are from heavy flavor decays, relaxing the identification requirements is not very useful. Our muon FO definition thus consists mainly of relaxing the isolation requirement.

Muon FO_10 definition: relax the following muon requirements from Section 7.3:

- $Iso < 1.0$ (was < 0.15).

Muon FO_10_d0 definition: relax the following muon requirements from Section 7.3:

- Transverse impact parameter with respect to the primary vertex < 2 mm (was $< 200\ \mu$m).
- $Iso < 1.0$ (was < 0.15).

7.6 Electron FO requirements for the WW analysis

We define four complementary fakeable object selections (V1, V2, V3, V4).

Electron V1 definition: relax the following electron requirements from Section 7.4:

- Remove the VBTF80 requirement.
- $Iso < 0.4$ (was < 0.10).
- The transverse impact parameter cut was removed (used to be < 200 μm).

Electron V2 definition: relax the following electron requirements from Section 7.4:

- Remove the VBTF80 requirement.
- The transverse impact parameter cut was removed (used to be < 200 μm).

Electron V3 definition: relax the following electron requirements from Section 7.4:

- $Iso < 1.0$ (was < 0.10).
- The transverse impact parameter cut was removed (used to be < 200 μm).

Electron V4 definition: relax the following electron requirements from Section 7.4:

- Use VBTF90 identification instead of VBTF80.
- $Iso < 1.0$ (was < 0.10).
- The transverse impact parameter cut was removed (used to be < 200 μm).

To summarize the qualitative differences between the four choices:

- the V1 FR is an extrapolation in both ID and isolation;
- the V2 FR is an extrapolation in ID;
- the V3 FR is an extrapolation in isolation;
- the V4 FR is an extrapolation in both partial ID and isolation.

8 Event selection and Lepton ID for the opposite sign analysis (SUS-10-007)

8.1 Event cleanup for the sample used to derive FR for the opposite sign analysis

These are identical to those described in Section 6.1.

8.2 Lepton triggers for FR determination for the opposite sign analysis

These are identical to those described in Section 6.2 except that we do not use the EGMonitor dataset.

8.3 Muon requirements for the opposite sign analysis

These are the same requirements listed in Reference[3]. We list them here for completeness. Muon candidates are RECO muon objects passing the following requirements:

- $P_T > 10$ GeV.
- $|\eta| < 2.4$.
- Global Muon and Tracker Muon.
- χ^2/ndof of global fit < 10.
- At least 11 hits in the tracker fit.
- Transverse impact parameter with respect to the beamspot < 200 μm.
- $Iso \equiv E_T^{\rm iso}$/Max(20 GeV, P_T) < 0.15. $E_T^{\rm iso}$ is defined as the sum of transverse energy/momentum deposits in ecal, hcal, and tracker, in a cone of 0.3.
- At least one of the hits from the standalone muon must be used in the global fit.
- Require tracker $\Delta P_T/P_T < 0.1$.

8.4 Electron requirements for the opposite sign analysis

These are the same requirements listed in Reference[3] We list them here for completeness. Electron candidates are RECO GSF electrons passing the following requirements:

- $P_T > 10$ GeV.
- $|\eta| < 2.5$.
- SuperCluster $E_T > 10$ GeV.
- The electron must be ecal seeded.
- VBTF90 identification[11].
- Transverse impact parameter with respect to the beamspot < 400 μm.

  item $Iso \equiv E_T^{\rm iso}$/Max(20 GeV, P_T) < 0.15. $E_T^{\rm iso}$ is defined as the sum of transverse energy/momentum deposits in ecal, hcal, and tracker, in a cone of 0.3. A 1 GeV pedestal is subtracted from the ecal energy deposition in the EB, however the ecal energy is never allowed to go negative.
- Electrons with a tracker or global muon within ΔR of 0.1 are vetoed.
- The number of missing expected inner hits must be less than two[13].
- Conversion removal via partner track finding: any electron where an additional GeneralTrack is found with $Dist < 0.02$ cm and $\Delta \cot\theta < 0.02$ is vetoed[13].
- Cleaning for ECAL spike (aka Swiss-Cross cleaning) has been applied

8.5 Muon FO requirements for the opposite sign analysis

Fakeable Objects are defined starting from the full lepton selection by relaxing some combination of the identification and isolation requirements. Since most muons in QCD events are from heavy flavor decays, relaxing the identification requirements is not very useful. Thus, our muon FO definition consists mainly of relaxing the isolation requirement.

Muon FO definition: relax the following muon requirements from Section 8.3:

- χ^2/ndof of global fit < 50 (was < 10).
- Transverse impact parameter with respect to the beamspot < 2 mm (was < 200 μm).
- $Iso < 0.4$ (was < 0.10).

8.6 Electron FO requirements for the opposite sign analysis

Unlike the other analyses described in this note, we only use one fakeable object version (The so-called V3).

Electron V3 FO definition: relax the following electron requirements from Section 8.4:

- $Iso < 0.4$ (was < 0.10).
- The impact parameter cut was removed (used to be < 200 μm).

9 Effects of W contamination

The effect of the $W \to \mu$ contamination is shown in Figure 1. Here we show the muon FR for the $t\bar{t}$ analysis in Monte Carlo from QCD only and including the W component. These fake rates require the presence of an offline caloJet with uncorrected $P_T > 15$ GeV.

Clearly the presence of $W \to \mu$ biases the fake rate at high P_T. The application of the tcMET< 20 GeV and transverse mass < 25 GeV cuts reduces the effect of the bias due to W, but for $P_T > 35$ GeV the bias starts to become significant even after these anti-W cuts.

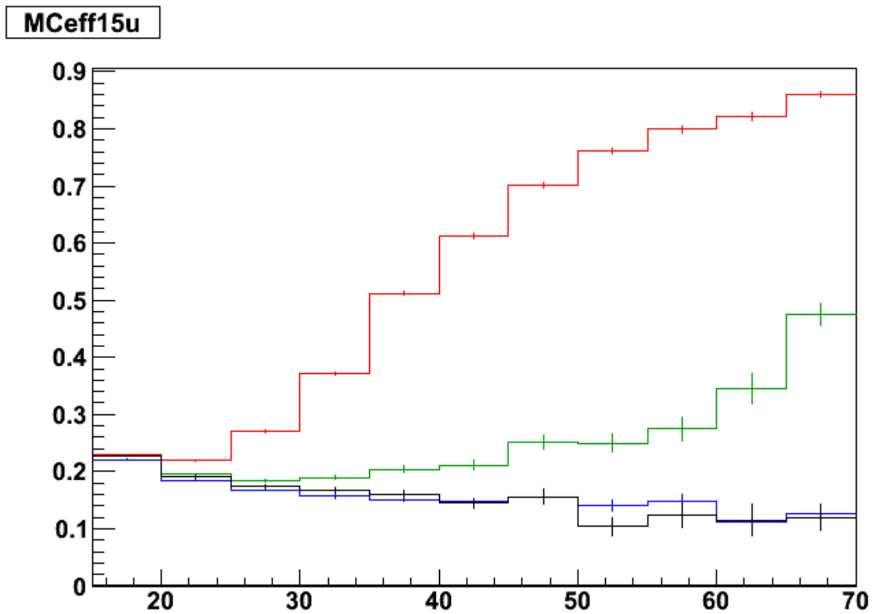

Figure 1: The Monte Carlo muon fake rate as a function of P_T. Blue: QCD. Red: QCD $+W$. Green: QCD $+W$ with the tcMet and transverse mass cuts. Black: QCD with the tcMet and transverse mass cuts.

The effect for the electrons fake rates is much less pronounced, see Figure 2.

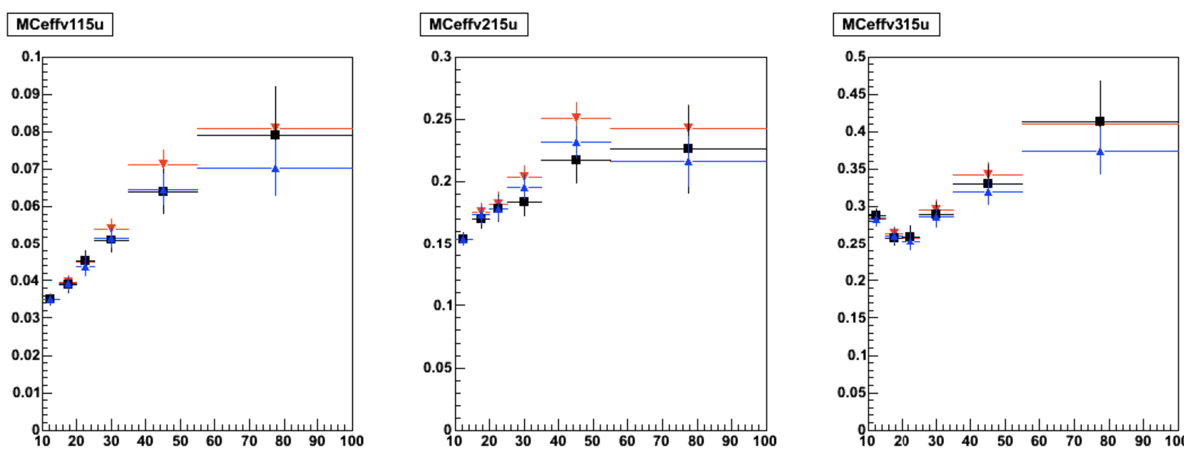

Figure 2: The Monte Carlo electron fake rates as a function of P_T. From left to right: V1, V2, V3. Blue: QCD. Red: QCD $+W$. Black: QCD with the tcMet and transverse mass cuts.

10 Jet triggers vs. lepton triggers

As mentioned in Section 3, we use lepton triggers instead of jet triggers to determine the FR. We have verified that this does not introduce a bias in the FR, as long as no explicit ID or isolation requirements are applied in the trigger. The triggers that we have chosen (see Section 4.2) do not have these requirements.

In Figures 3 and 4 we compare the fake rates for the ttbar analysis from `HLT_Jet15U` and for the lepton triggers of Section 4.2 with an offline caloJet of $P_T > 15$ GeV requirement.

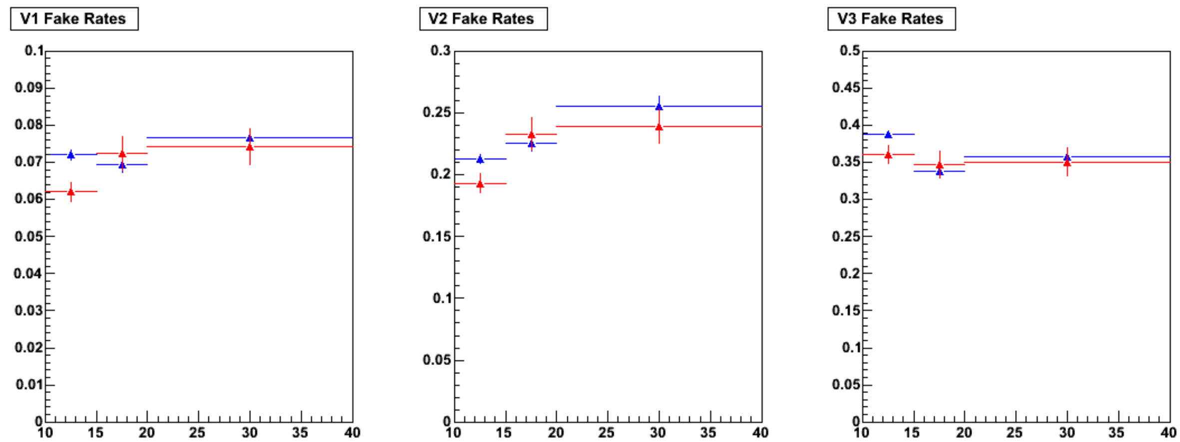

Figure 3: Electron fake rates as a function of P_T. From left to right: V1, V2, V3. Red: from the jet trigger. Blue: from the electron trigger.

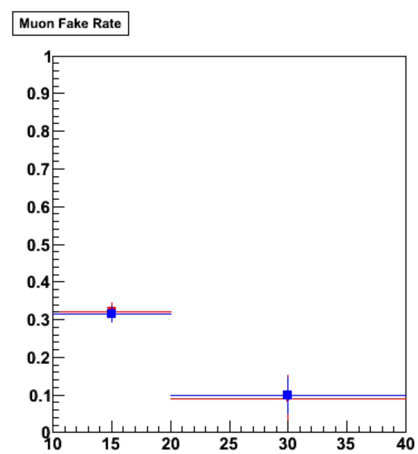

Figure 4: Muon fake rates as a function of P_T for run < 137000. Red: from the muon trigger. Blue: from the jet trigger.

11 FR for $t\bar{t}$ analysis with 3.1 pb^{-1} (TOP-10-001)

11.1 Results relevant to $t\bar{t}$ analysis with 3.1 pb^{-1} (TOP-10-001)

The fake rates used in the $t\bar{t}$ analysis are obtained from the EG and Mu datasets of Section 2.1 using the electron and muon triggers of Section 4.2. Events must have a caloJet of $P_T > 15$ GeV (uncorrected), and in case that there is only one such jet in the event, this jet must be well separated ($\Delta R > 1$) from the lepton candidate.

The choice of a 15 GeV threshold for the jet was made to roughly tailor the P_T distributions of the jets from which the FR is obtained to the P_T distributions of jets in the $W+$ jets background to the $t\bar{t}$ selection. The effects of varying this choice are given in Section 11.2.

The 2D fake rates are summarized in Tables 1 to 4.

Table 1: The muon FR for the top analysis.

p_T / $\lvert\eta\rvert$	$10.000 - 15.000$	$15.000 - 20.000$	$20.000 - 25.000$	$25.000 - 30.000$	$30.000 - 35.000$
$0.000 - 1.000$	0.320 ± 0.004	0.246 ± 0.009	0.241 ± 0.017	0.206 ± 0.026	0.243 ± 0.051
$1.000 - 1.479$	0.357 ± 0.007	0.257 ± 0.014	0.289 ± 0.030	0.218 ± 0.047	0.143 ± 0.059
$1.479 - 2.000$	0.352 ± 0.007	0.306 ± 0.014	0.255 ± 0.026	0.340 ± 0.049	0.343 ± 0.080
$2.000 - 2.500$	0.381 ± 0.012	0.345 ± 0.025	0.253 ± 0.048	0.147 ± 0.061	0.188 ± 0.098

Table 2: The V1 electron FR for the top analysis.

p_T / $\lvert\eta\rvert$	$10.000 - 15.000$	$15.000 - 20.000$	$20.000 - 25.000$	$25.000 - 35.000$
$0.000 - 1.000$	0.056 ± 0.002	0.053 ± 0.003	0.045 ± 0.005	0.049 ± 0.006
$1.000 - 1.479$	0.076 ± 0.003	0.074 ± 0.005	0.077 ± 0.008	0.062 ± 0.009
$1.479 - 2.000$	0.067 ± 0.003	0.081 ± 0.006	0.089 ± 0.010	0.109 ± 0.011
$2.000 - 2.500$	0.093 ± 0.003	0.086 ± 0.005	0.087 ± 0.007	0.111 ± 0.009

Table 3: The V2 electron FR for the top analysis.

p_T / $\lvert\eta\rvert$	$10.000 - 15.000$	$15.000 - 20.000$	$20.000 - 25.000$	$25.000 - 35.000$
$0.000 - 1.000$	0.191 ± 0.006	0.180 ± 0.009	0.160 ± 0.015	0.193 ± 0.022
$1.000 - 1.479$	0.226 ± 0.008	0.257 ± 0.016	0.294 ± 0.028	0.256 ± 0.035
$1.479 - 2.000$	0.205 ± 0.009	0.291 ± 0.019	0.336 ± 0.031	0.396 ± 0.034
$2.000 - 2.500$	0.227 ± 0.007	0.236 ± 0.012	0.243 ± 0.018	0.291 ± 0.021

Table 4: The V3 electron FR for the top analysis.

p_T / $\|\eta\|$	$10.000 - 15.000$	$15.000 - 20.000$	$20.000 - 25.000$	$25.000 - 35.000$
$0.000 - 1.000$	0.386 ± 0.011	0.352 ± 0.016	0.319 ± 0.027	0.342 ± 0.034
$1.000 - 1.479$	0.412 ± 0.012	0.328 ± 0.020	0.333 ± 0.031	0.311 ± 0.040
$1.479 - 2.000$	0.325 ± 0.014	0.302 ± 0.019	0.331 ± 0.030	0.370 ± 0.032
$2.000 - 2.500$	0.408 ± 0.011	0.356 ± 0.016	0.349 ± 0.024	0.419 ± 0.028

11.2 Jet dependence relevant to $t\bar{t}$ analysis with 3.1 pb^{-1} (TOP-10-001)

The dependence of the FR on the minimum jet P_T requirement is shown in Figures 5 and 6.

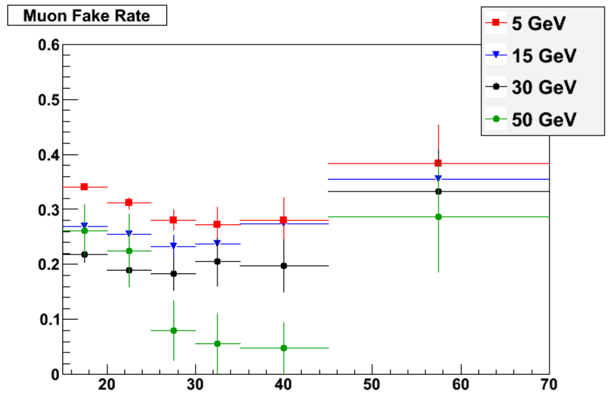

Figure 5: The muon fake rate as a function of P_T for the top analysis obtained from data with different minimum jet P_T requirements. ($P_T > 5, 15, 30, 50$ GeV for uncorrected caloJets).

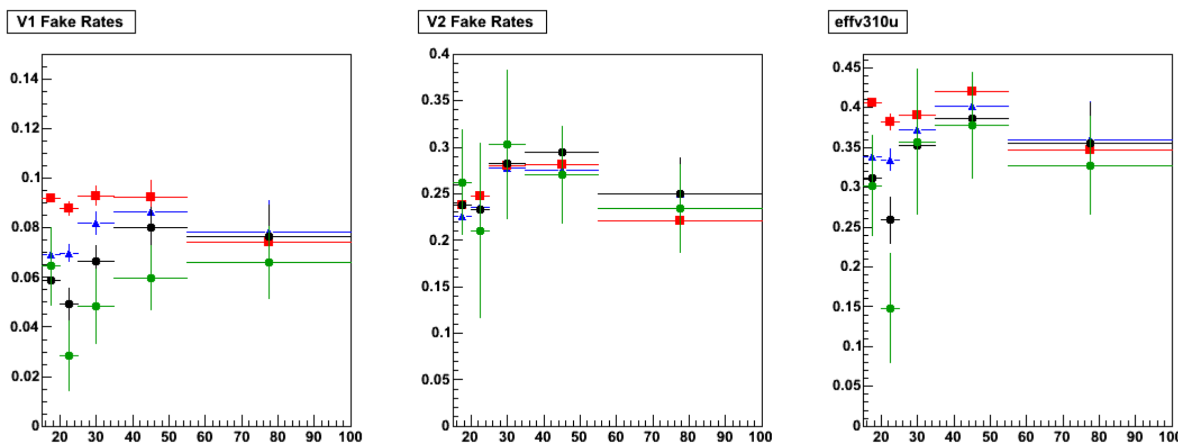

Figure 6: The electron fake rates as a function of P_T for the top analysis obtained from data with different minimum jet P_T requirements. ($P_T > 5, 15, 30, 50$ GeV for uncorrected caloJets). The color coding is the same as Figure 5.

There is a non negligible jet P_T dependence to the FR. A better way to quantify the jet P_T dependence is to use the FR obtained from events with jets of $P_T > 15$ GeV to predict the rate of leptons in events with jets of $P_T > 5, 30, 50$ GeV and compare with the observation. These comparisons are shown in Tables 5 and 6[2)].

Table 5: Observed numbers of muons with $P_T > 20$ in data QCD events as a function of the minimum jet P_T in the event compared with the prediction from the FR method. The FR was obtained from the jet $P_T > 15$ GeV sample.

Jet Requirement	Muons observed	Muons predicted	Observed/predicted
$P_T > 5$ GeV	885	674	1.31
$P_T > 30$ GeV	140	187	0.75
$P_T > 50$ GeV	19	35	0.54

We believe that the jet P_T dependence is due to the isolation extrapolation. Lepton candidates in high P_T jet events tend to have a smaller probability of passing an isolation requirement. This trend is reasonably well reproduced in QCD Monte Carlo, see Table 7, and compare with Table 5. Note that the V2 electron fake rate, which does not extrapolate in isolation (see Section 4.6) gives the most stable results with respect to jet P_T variations.

[2)] The Z-veto, tcMet, and transverse mass requirements from Section 3.2 are included.

Table 6: Observed numbers of electrons with $P_T > 20$ in data QCD events as a function of the minimum jet P_T in the event compared with the prediction from the FR method. The FR was obtained from the jet $P_T > 15$ GeV sample.

Jet Requirement	Ele observed	Ele pred. V1	Ele pred. V2	Ele pred. V3	obs/pred V1	obs/pred V2	obs/pred V3
$P_T > 5$ GeV	1599	1312	1538	1377	1.22	1.04	1.16
$P_T > 30$ GeV	272	324	272	302	0.8	1.0	0.9
$P_T > 50$ GeV	53	76	58	63	0.7	0.9	0.8

Table 7: Same as Table 5 but for Monte Carlo QCD events.

Jet Requirement	Muons observed	Muons predicted	Observed/predicted
$P_T > 5$	37427	31733	1.18
$P_T > 30$	5593	7397	0.76
$P_T > 50$	784	1323	0.60

11.3 Closure test relevant to $t\bar{t}$ analysis with 3.1 pb^{-1} (TOP-10-001)

We apply the FR obtained in QCD Monte Carlo to predict the rate of dileptons with one real and one fake lepton in $W+$ jets and $t\bar{t}$ Monte Carlo. For the W Monte Carlo we require that one of the leptons be truth matched to $W \to \ell$ and we apply the FR method to events where a companion FO is also found. In the top Monte Carlo we must be careful, since rare failures of truth matching can have a big impact on this study due to the presence of real dileptons in the sample. Thus, the test is done only on the subset of events where the $t\bar{t}$ pair decays to $\ell+$ jets; with this precaution, the test then is done exactly as in the W case.

For muons, in the W sample we find 6 fake muons, in good agreement with the FR prediction of 7.2; in the $t\bar{t}$ sample we find 60 fake muons, while the FR prediction is 132. The overprediction of the FR method for muons in the $t\bar{t}$ sample is due to the fact that the typical P_T of jets in the top sample is higher than the P_T of the QCD events from which the FR is derived. This is the same trend that was discussed in the previous Section, see Tables 5 and 7.

For electrons, we do not have closure test results with the latest electron selection and the latest Monte Carlo samples. We did perform this test while the analysis was being designed, using an earlier incarnation of the electron identification requirement (the so called "cand01" operating point of Reference [15]), and with a slightly tighter isolation requirement (0.10 instead of 0.15). Also, the MC samples were from CMSSW 3_1 and the FR was from the QCD30 sample with no explicit jet requirement.

Despite these differences, we think the old results are still interesting and the conclusions that we can draw also apply to the newer selection. In the W case, we found 84 fake electrons compared to FR predictions of 101 (V1), 94 (V2) or 91 (V3). In the $t\bar{t}$ case, we found 168 fake electrons compared to FR predictions of 154 (V1), 194 (V2), and 162 (V3). We find that the FR method reproduces the rate of fake electron to about 20-30%. This is in reasonable agreeemnt with the tests of jet P_T dependence of Table 6.

11.4 Jet flavor dependence relevant to $t\bar{t}$ analysis with 3.1 pb^{-1} (TOP-10-001)

Here we test the stability of the FR in data as a function of jet flavor. The idea is to use b-tagging to enrich the sample in heavy flavor, and to use the FR obtained from generic jets to predict the lepton yield.

For this exercise we use FR obtained with a jet $P_T > 15$ GeV requirement. We apply the FR to a sample with

1. at least one caloJet of uncorrected $P_T > 15$ GeV well separated from the lepton under consideration
2. this caloJet must be btagged by `simpleSecondaryVertexHighEffBJetTag` with medium operating point (cut at 1.74 according to Reference [16]).
3. the Z-veto, tcMET, and transverse mass requirements from Section 3.2.
4. triggered by the lepton triggers of Section 4.2

The 15 GeV requirement is chosen to be the same as the requirement in the FR determination so as to factor out, at least to first order, the known jet P_T dependence (Section 11.2). The results are shown in Table 8. The agreement between the observation and the prediction is quite good.

Table 8: Observed and predicted number of leptons of $P_T > 20$ GeV in the heavy flavor enriched sample.

Observed muons	Predicted Muons	Observed electrons	V1-Predicted electrons	V2-Predicted electrons	V3-Predicted electrons
27	35	19	24	20	25

12 FR for $t\bar{t}$ analysis with 36.1 pb^{-1} (TOP-10-005)

12.1 Results relevant to $t\bar{t}$ analysis with 36.1 pb^{-1} (TOP-10-005)

In Tables 9, 10, 11 and 12 we show the FRs used for the $t\bar{t}$ analysis with 36.1 pb^{-1} of data derived using the EG, EGMonitor and Mu datasets as defined in Section 2.2. Events were required to have a caloJet of $P_T > 15$ GeV (uncorrected). For events with only one such jet, we required the jet to be well separated ($\Delta R > 1$) from the lepton candidate.

The choice of a 15 GeV threshold for the jet was made to roughly tailor the P_T distributions of the jets from which the FR is obtained to the P_T distributions of jets in the $W+$ jets background to the $t\bar{t}$ selection. The effects of varying this choice are given in Section 12.2.

Table 9: Muon FR for the top analysis with the full 2010 dataset

p_T / $\lvert\eta\rvert$	10.000 – 15.000	15.000 – 20.000	20.000 – 25.000	25.000 – 30.000	30.000 – 35.000
0.000 – 1.000	0.3203 $\pm$ 0.0012	0.2461 $\pm$ 0.0026	0.2364 $\pm$ 0.0051	0.2253 $\pm$ 0.0085	0.2243 $\pm$ 0.0139
1.000 – 1.479	0.3531 $\pm$ 0.0019	0.2840 $\pm$ 0.0042	0.2571 $\pm$ 0.0081	0.2566 $\pm$ 0.0136	0.2407 $\pm$ 0.0213
1.479 – 2.000	0.3630 $\pm$ 0.0019	0.2987 $\pm$ 0.0042	0.2794 $\pm$ 0.0081	0.2798 $\pm$ 0.0134	0.3277 $\pm$ 0.0231
2.000 – 2.400	0.3643 $\pm$ 0.0034	0.3072 $\pm$ 0.0072	0.2710 $\pm$ 0.0137	0.3070 $\pm$ 0.0254	0.2773 $\pm$ 0.0410

Table 10: Electron V1 FR for the top analysis with the full 2010 dataset

p_T / $\lvert\eta\rvert$	15.000 – 20.000	20.000 – 25.000	25.000 – 30.000	30.000 – 35.000
0.000 – 1.000	0.0579 $\pm$ 0.0008	0.0508 $\pm$ 0.0009	0.0485 $\pm$ 0.0013	0.0490 $\pm$ 0.0017
1.000 – 1.479	0.0746 $\pm$ 0.0011	0.0705 $\pm$ 0.0015	0.0698 $\pm$ 0.0020	0.0686 $\pm$ 0.0027
1.479 – 2.000	0.0601 $\pm$ 0.0010	0.0607 $\pm$ 0.0013	0.0677 $\pm$ 0.0018	0.0622 $\pm$ 0.0024
2.000 – 2.500	0.0531 $\pm$ 0.0007	0.0523 $\pm$ 0.0009	0.0538 $\pm$ 0.0012	0.0554 $\pm$ 0.0017

12.2 Jet dependence relevant to $t\bar{t}$ analysis with 36.1 pb^{-1} (TOP-10-005)

The dependence of the FR on the minimum jet P_T requirement is shown in Figures 7 and 8. The variation is similar to that seen in the 3.1 pb^{-1} analysis 11.2.

Table 11: Electron V2 FR for the top analysis with the full 2010 dataset

$\lvert\eta\rvert$ \ p_T	15.000 – 20.000	20.000 – 25.000	25.000 – 30.000	30.000 – 35.000
0.000 – 1.000	0.2198 $\pm$ 0.0027	0.2015 $\pm$ 0.0035	0.1994 $\pm$ 0.0048	0.2057 $\pm$ 0.0067
1.000 – 1.479	0.2581 $\pm$ 0.0035	0.2650 $\pm$ 0.0050	0.2760 $\pm$ 0.0070	0.2765 $\pm$ 0.0097
1.479 – 2.000	0.2133 $\pm$ 0.0033	0.2262 $\pm$ 0.0044	0.2481 $\pm$ 0.0061	0.2243 $\pm$ 0.0077
2.000 – 2.500	0.1457 $\pm$ 0.0018	0.1455 $\pm$ 0.0023	0.1440 $\pm$ 0.0031	0.1445 $\pm$ 0.0042

Table 12: Electron V3 FR for the top analysis with the full 2010 dataset

$\lvert\eta\rvert$ \ p_T	15.000 – 20.000	20.000 – 25.000	25.000 – 30.000	30.000 – 35.000
0.000 – 1.000	0.3312 $\pm$ 0.0032	0.3108 $\pm$ 0.0042	0.3111 $\pm$ 0.0060	0.3090 $\pm$ 0.0082
1.000 – 1.479	0.3467 $\pm$ 0.0039	0.3251 $\pm$ 0.0050	0.3188 $\pm$ 0.0070	0.3396 $\pm$ 0.0100
1.479 – 2.000	0.3429 $\pm$ 0.0042	0.3234 $\pm$ 0.0050	0.3270 $\pm$ 0.0068	0.3217 $\pm$ 0.0091
2.000 – 2.500	0.3801 $\pm$ 0.0036	0.3654 $\pm$ 0.0046	0.3793 $\pm$ 0.0065	0.3991 $\pm$ 0.0093

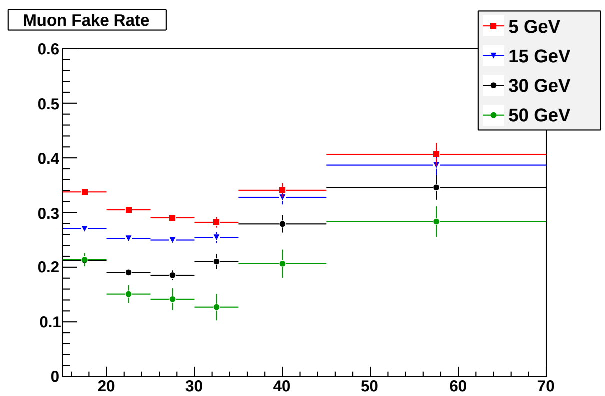

Figure 7: The muon fake rate as a function of P_T for the top analysis obtained from data with different minimum jet P_T requirements. ($P_T > 5, 15, 30, 50$ GeV for uncorrected caloJets).

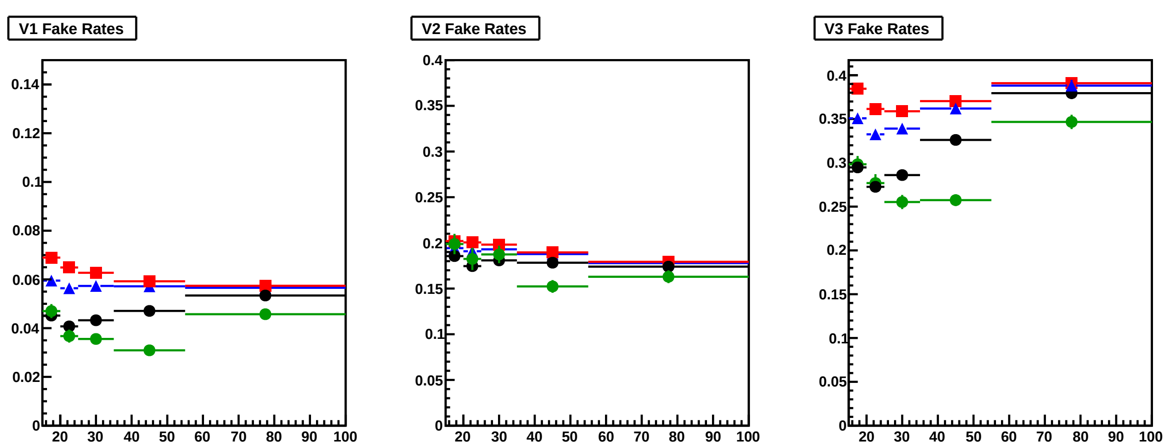

Figure 8: The electron fake rates as a function of P_T for the top analysis obtained from data with different minimum jet P_T requirements. ($P_T > 5, 15, 30, 50$ GeV for uncorrected caloJets). The color coding is the same as Figure 7.

13 FR for same sign analysis

13.1 Results relevant to same sign analysis

The fake rates used in the same sign analysis are obtained from the EG, EGMonitor, and Mu datasets of Section 2.3 using the electron and muon triggers of Section 6.2. Events must have a pfJet of $P_T > 40$ GeV (corrected), and in case that there is only one such jet in the event, this jet must be well separated ($\Delta R > 1$) from the lepton candidate. Results for varying jet thresholds are presented in Section 13.2.

The choice of a 40 GeV threshold for the jet was made to roughly tailor the P_T distributions of the jets from which the FR is obtained to the P_T distributions of b-quarks that produce FOs in the top pair Monte Carlo. Production of $t\bar{t}$ followed by decay into the one lepton plus jets final state is the main background to the same sign analysis. One lepton from a W and the second lepton from semi-leptonic decay of the b-quark from the other top decay typically leads to two same sign leptons, MET, and three additional jets in the event.

Figure 9 shows the P_T distribution of b-quarks that produce FOs as well as b-quarks that produce numerators for muons. Figure 10 shows the same for electrons. It is quite apparent that there is a significant difference in the distributions for FOs and numerators. The high P_T tail produces FOs that essentially never pass the numerator requirements. The 40 GeV threshold, as well as its variations, attempts to model FRs that approximate the relevant range of b-quark P_T to minimize this effect. Section 13.3 summarizes residual systematics as determined by applying the FR from QCD Monte Carlo to top pair backgrounds in the same sign analysis in Monte Carlo.

The 2D fake rates are summarized in Tables 13 to 20 for Monte Carlo and Data.

Table 13: The muon FR for the same sign analysis in Monte Carlo.

$\lvert\eta\rvert$ \ p_T	10.000 – 15.000	15.000 – 20.000	20.000 – 25.000	25.000 – 35.000
0.000 – 1.000	0.1519 ± 0.0011	0.0845 ± 0.0005	0.0602 ± 0.0007	0.0466 ± 0.0007
1.000 – 1.479	0.1683 ± 0.0017	0.1063 ± 0.0008	0.0814 ± 0.0011	0.0671 ± 0.0012
1.479 – 2.000	0.1792 ± 0.0017	0.1171 ± 0.0008	0.0888 ± 0.0011	0.0795 ± 0.0013
2.000 – 2.500	0.1755 ± 0.0029	0.1198 ± 0.0014	0.0928 ± 0.0020	0.0792 ± 0.0023

Table 14: The muon FR for the same sign analysis in Data.

$\lvert\eta\rvert$ \ p_T	10.000 – 15.000	15.000 – 20.000	20.000 – 25.000	25.000 – 35.000
0.000 – 1.000	0.1558 ± 0.0026	0.1062 ± 0.0043	0.1030 ± 0.0070	0.0987 ± 0.0084
1.000 – 1.479	0.1904 ± 0.0044	0.1148 ± 0.0069	0.1245 ± 0.0120	0.1262 ± 0.0145
1.479 – 2.000	0.1917 ± 0.0043	0.1482 ± 0.0074	0.1100 ± 0.0111	0.1076 ± 0.0133
2.000 – 2.500	0.1877 ± 0.0074	0.1324 ± 0.0123	0.0963 ± 0.0180	0.1159 ± 0.0250

Table 15: The V1 electron FR for the same sign analysis in data.

$\lvert\eta\rvert$ \ p_T	$10.000 - 15.000$	$15.000 - 20.000$	$20.000 - 25.000$	$25.000 - 35.000$
0.000 – 1.000	0.0271 ± 0.0029	0.0177 ± 0.0010	0.0168 ± 0.0012	0.0143 ± 0.0011
1.000 – 1.479	0.0284 ± 0.0033	0.0203 ± 0.0014	0.0188 ± 0.0018	0.0195 ± 0.0018
1.479 – 2.000	0.0141 ± 0.0027	0.0141 ± 0.0013	0.0134 ± 0.0016	0.0182 ± 0.0017
2.000 – 2.500	0.0157 ± 0.0030	0.0151 ± 0.0013	0.0171 ± 0.0018	0.0151 ± 0.0015

Table 16: The V2 electron FR for the same sign analysis in data.

$\lvert\eta\rvert$ \ p_T	$10.000-15.000$	$15.000-20.000$	$20.000-25.000$	$25.000-35.000$
0.000 – 1.000	0.1913 ± 0.0183	0.1407 ± 0.0072	0.1368 ± 0.0094	0.1234 ± 0.0089
1.000 – 1.479	0.1728 ± 0.0188	0.1538 ± 0.0095	0.1582 ± 0.0137	0.1646 ± 0.0138
1.479 – 2.000	0.1111 ± 0.0205	0.1279 ± 0.0109	0.1245 ± 0.0141	0.1425 ± 0.0127
2.000 – 2.500	0.0841 ± 0.0155	0.0997 ± 0.0081	0.1129 ± 0.0110	0.0947 ± 0.0092

Table 17: The V3 electron FR for the same sign analysis in data.

$\lvert\eta\rvert$ \ p_T	$10.000-15.000$	$15.000-20.000$	$20.000-25.000$	$25.000-35.000$
0.000 – 1.000	0.2307 ± 0.0054	0.1810 ± 0.0062	0.1725 ± 0.0082	0.1906 ± 0.0091
1.000 – 1.479	0.2586 ± 0.0073	0.1917 ± 0.0085	0.1816 ± 0.0114	0.2070 ± 0.0120
1.479 – 2.000	0.2491 ± 0.0104	0.1843 ± 0.0107	0.1907 ± 0.0138	0.2404 ± 0.0143
2.000 – 2.500	0.3400 ± 0.0134	0.2466 ± 0.0145	0.2633 ± 0.0177	0.3296 ± 0.0187

Table 18: The V1 electron FR for the same sign analysis in Monte Carlo.

$\lvert\eta\rvert$ \ p_T	$10.000-15.000$	$15.000-20.000$	$20.000-25.000$	$25.000-35.000$
0.000 – 1.000	0.0224 ± 0.0036	0.0168 ± 0.0043	0.0036 ± 0.0024	0.0140 ± 0.0054
1.000 – 1.479	0.0164 ± 0.0038	0.0184 ± 0.0068	0.0135 ± 0.0075	0.0123 ± 0.0076
1.479 – 2.000	0.0146 ± 0.0039	0.0027 ± 0.0010	0.0171 ± 0.0091	0.0098 ± 0.0061
2.000 – 2.500	0.0139 ± 0.0040	0.0052 ± 0.0029	0.0115 ± 0.0071	0.0092 ± 0.0057

Table 19: The V2 electron FR for the same sign analysis in Monte Carlo.

$\lvert\eta\rvert$ \ p_T	$10.000-15.000$	$15.000-20.000$	$20.000-25.000$	$25.000-35.000$
0.000 – 1.000	0.1636 ± 0.0242	0.1290 ± 0.0307	0.0386 ± 0.0260	0.0993 ± 0.0362
1.000 – 1.479	0.1050 ± 0.0230	0.1457 ± 0.0502	0.0955 ± 0.0511	0.1114 ± 0.0650
1.479 – 2.000	0.1114 ± 0.0280	0.0286 ± 0.0111	0.1204 ± 0.0603	0.0969 ± 0.0573
2.000 – 2.500	0.0935 ± 0.0254	0.0309 ± 0.0171	0.0756 ± 0.0455	0.0510 ± 0.0306

Table 20: The V3 electron FR for the same sign analysis in Monte Carlo.

$\lvert\eta\rvert$ \ p_T	$10.000-15.000$	$15.000-20.000$	$20.000-25.000$	$25.000-35.000$
0.000 – 1.000	0.2209 ± 0.0315	0.2348 ± 0.0518	0.0492 ± 0.0330	0.2197 ± 0.0722
1.000 – 1.479	0.2235 ± 0.0455	0.2053 ± 0.0675	0.1556 ± 0.0796	0.1601 ± 0.0891
1.479 – 2.000	0.2330 ± 0.0548	0.0440 ± 0.0176	0.2698 ± 0.1198	0.1706 ± 0.0951
2.000 – 2.500	0.3566 ± 0.0805	0.1573 ± 0.0832	0.2767 ± 0.1466	0.2042 ± 0.1103

13.2 Jet dependence relevant to same sign analysis

To quantify the jet P_T dependence, we use the FR obtained from events with pfjets of $P_T > 40$ GeV to predict the rate of leptons in events with jets of $P_T > 20, 40, 60$ GeV and compare with the observation. We make these comparisons for the sample the FR was determined from, as well as a b-quark enriched sample. We do so for both data and Monte Carlo.

These comparisons are shown in Tables 21 to 30[3)]. The statistics for V1,V2 and V3 in data differs in these tables due to the differences in triggers used, as described in Section 6.2.

We believe that the jet P_T dependence is due to the isolation extrapolation. Lepton candidates in high P_T jet events tend to have a smaller probability of passing an isolation requirement. The tables show that this trend is well reproduced in data and Monte Carlo, with and without b-tagging. We make the following disclaimers and observations:

- The Monte Carlo studies for electrons are limited by statistics. The fake rates themselves have errors ranging from 20-70% per bin as shown in Tables 18 to 20. This is driven by a few events entering with very large weights given that the three QCD samples span two orders of magnitude in equivalent luminosity.
- The $P_T > 40$ GeV test is not completely statistically independent from the fake rates themselves as follows.
  - For the muon fake rates in Monte Carlo, both the mu15 and the mu10 enriched sample were used in the fake rates, while only the mu10 sample was used for the application of the fake rate. There is thus significant but not perfect overlap between test and fake rate determination.
  - The test on data without b-tagging for both electrons and muons, there is a small non-overlap between the events that were used to produce the fake rate and the events used to test the fake rate. This is due to numerators and FOs with $P_T > 35$ GeV which were not used in the fake rate determination to avoid bias from W contamination (See Section 9), but are used during the test of the fake rate. The prediction being slightly less than the observed is thus probably attributable to the small W contamination in the sample after MET and transverse mass cuts.
  - The test on electrons without b-tagging uses very close to the same sample as the fake rate determination, except that $P_T > 35$ GeV FOs and numerators are ignored in the fake rate determination but treated as if they were in the 25-35 GeV bin when the fake rate is applied.
  - All of the tests on b-enriched samples for data and Monte Carlo are essentially statistically independent from the fake rate because the b-tagged sample is such a small subsample of the total sample, or as in the case of electrons in Monte Carlo, a completely different sample dominates.
- The $P_T > 20$ GeV and 60 GeV tests are sufficiently independent of the fake rate sample as can be seen by the statistics in the tables.
- The V2 electron fake rate, which does not extrapolate in isolation (see Section 6.6) tends to give more stable results with respect to jet P_T variations in data. This is as expected.
- Neither the V1 nor V2 electron fake rates, both of which extrapolate in Id (see Section 6.6), do very well in predicting the observation in the b-tagged sample. This is true for both data and Monte Carlo.
- The V3 electron fake rate, which extrapolates only in isolation (see Section 6.6), does equally well for the b-enriched as for the generic QCD samples, for data as well as Monte Carlo. It appears as if a jet is a jet no matter what it's made by, as far as isolation is concerned.

[3)] The MET, and transverse mass requirements from Section 3.2 are included.

Table 21: Observed numbers of muons with $P_T > 10$ in Monte Carlo QCD events as a function of the minimum jet P_T in the event compared with the prediction from the FR method. The FR was obtained from the jet $P_T > 40$ GeV sample. The fake rate is applied only to the Pt10 muon enriched dataset as described in Section 2.3.

Jet Requirement	Muons observed	Muons predicted	Predicted/observed
$p_T > 20$ GeV	303973	233252	0.77
$p_T > 40$ GeV	37201	36181.8	0.97
$p_T > 60$ GeV	5676	5971.09	1.05

Table 22: Observed numbers of muons with $P_T > 10$ in data QCD events as a function of the minimum jet P_T in the event compared with the prediction from the FR method. The FR was obtained from the jet $P_T > 40$ GeV sample.

Jet Requirement	Muons observed	Muons predicted	Predicted/observed
$p_T > 20$ GeV	88051	57149.2	0.65
$p_T > 40$ GeV	8945	8801	0.98
$p_T > 60$ GeV	1482	1547.2	1.04

Table 23: Observed numbers of muons with $P_T > 10$ in Monte Carlo QCD events after b-tagging as a function of the minimum jet P_T in the event compared with the prediction from the FR method. The FR was obtained from the jet $P_T > 40$ GeV sample. The fake rate is applied only to the Pt10 muon enriched dataset as described in Section 2.3.

Jet Requirement	Muons observed	Muons predicted	Predicted/observed
$p_T > 20$ GeV	28699	20256.1	0.71
$p_T > 40$ GeV	6909	5571.32	0.81
$p_T > 60$ GeV	1480	1270.21	0.86

Table 24: Observed numbers of muons with $P_T > 10$ in data QCD events after b-tagging as a function of the minimum jet P_T in the event compared with the prediction from the FR method. The FR was obtained from the jet $P_T > 40$ GeV sample.

Jet Requirement	Muons observed	Muons predicted	Predicted/observed
$p_T > 20$ GeV	5162	3944.05	0.76
$p_T > 40$ GeV	976	1008.19	1.03
$p_T > 60$ GeV	232	232.09	1.00

Table 25: Observed numbers of electrons with $P_T > 10$ in Monte Carlo QCD events as a function of the minimum jet P_T in the event compared with the prediction from the FR method. The FR was obtained from the jet $P_T > 40$ GeV sample.

Jet Requirement	Ele observed	v1 pred	v1 pred/obs	v2 pred	v2 pred/obs	v3 pred	v3 pred/obs
$p_T > 20$ GeV	5.974	5.620	0.94	5.425	0.91	5.972	1.00
$p_T > 40$ GeV	1.966	1.918	0.98	1.949	0.99	1.904	0.97
$p_T > 60$ GeV	0.413	0.397	0.96	0.386	0.94	0.393	0.95

Table 26: Observed numbers of electrons with $P_T > 10$ in b-tagged Monte Carlo QCD events as a function of the minimum jet P_T in the event compared with the prediction from the FR method. The FR was obtained from the jet $P_T > 40$ GeV sample. To gain statistics, we apply the fake rates to the Pt10 muon-enriched samples, and require the electron candidate to be in the away jet from the muon.

Jet Requirement	Ele observed	v1 pred	v1 pred/obs	v2 pred	v2 pred/obs	v3 pred	v3 pred/obs
$p_T > 20$ GeV	518	133.53	0.26	172.50	0.33	416.62	0.80
$p_T > 40$ GeV	230	62.20	0.27	70.04	0.30	172.61	0.75
$p_T > 60$ GeV	50	21.18	0.42	22.90	0.46	54.86	1.10

Table 27: Observed numbers of electrons with $P_T > 10$ in data QCD events as a function of the minimum jet P_T in the event compared with the prediction from the FR method. The FR was obtained from the jet $P_T > 40$ GeV sample. Shown here are only V1 and V2.

Jet Requirement	observed	v1 pred	v1 pred/obs	v2 pred	v2 pred/obs
$p_T > 20$ GeV	14688	8744.14	0.60	12891.16	0.88
$p_T > 40$ GeV	2401	2329.47	0.97	2379.94	0.99
$p_T > 60$ GeV	472	579.80	1.23	494.25	1.05

Table 28: Observed numbers of electrons with $P_T > 10$ in data QCD events as a function of the minimum jet P_T in the event compared with the prediction from the FR method. The FR was obtained from the jet $P_T > 40$ GeV sample. Shown here is only V3.

Jet Requirement	observed	v3 pred	v3 pred/obs
$p_T > 20$ GeV	61666	41100.40	0.67
$p_T > 40$ GeV	7578	7234.83	0.95
$p_T > 60$ GeV	1400	1447.60	1.03

Table 29: Observed numbers of electrons with $P_T > 10$ in b-tagged data QCD events as a function of the minimum jet P_T in the event compared with the prediction from the FR method. The FR was obtained from the jet $P_T > 40$ GeV sample. Shown here are only V1 and V2.

Jet Requirement	observed	v1 pred	v1 pred/obs	v2 pred	v2 pred/obs
$p_T > 20$ GeV	502	232.67	0.46	273.43	0.54
$p_T > 40$ GeV	142	97.28	0.69	85.59	0.60
$p_T > 60$ GeV	33	32.94	1.00	28.99	0.88

Table 30: Observed numbers of electrons with $P_T > 10$ in b-tagged data QCD events as a function of the minimum jet P_T in the event compared with the prediction from the FR method. The FR was obtained from the jet $P_T > 40$ GeV sample. Shown here is only V3.

Jet Requirement	observed	v3 pred	v3 pred/obs
$p_T > 20$ GeV	2215	1965.66	0.89
$p_T > 40$ GeV	506	590.52	1.17
$p_T > 60$ GeV	119	152.54	1.28

13.3 Closure test relevant to same sign analysis

We apply the FR obtained in QCD Monte Carlo to predict the rate of dileptons with one real and one fake lepton in $t\bar{t}$ Monte Carlo. We must be careful, since rare failures of truth matching can have a big impact on this study due to the presence of real dileptons in the sample. Thus, the test is done only in the subset of events where the $t\bar{t}$ pair decays to $\ell+$ jets.

In the $t\bar{t}$ sample we find 54 fake muons, while the FR prediction is 94.75, for a ratio of predicted/observed = 1.75. For electrons, in the $t\bar{t}$ sample we find 47 fake electrons, while the FR prediction is 89 ± 11, for a ratio of predicted/observed = 1.89 ± 0.36. This is for V3 of the electron fake rates. The V1 and V2 fake rates predict 30 ± 4 and 21 ± 3 respectively. The errors on the prediction for electrons in this case include statistical errors on fake rates as well as top sample statistics.

We hypothesize that the overprediction of the FR method for both muons and electrons in the $t\bar{t}$ sample is partly due to the fact that the high jet P_T tail visible in Figure 9 for FOs is larger in the $t\bar{t}$ Monte Carlo than in the QCD Monte Carlo the FR was derived from. The discrepancy in the two distributions overlayed is supporting this notion. In addition, there may be an effect due to the 3-4 jet environment that forms the top background to the same sign analysis. To verify the latter effect, we select a sub-sample of the top Monte Carlo for which there is no jet within an annulus in ΔR from 0.4 to 1.0. This sample provides an event environment for the FO that is more like the one from which the fake rate was derived. In this test we observe 35 (33) muons (electrons) and predict 57.2 ± 2.8 (43.2 ± 3.0), for a ratio predicted over observed of 1.6 (1.3) as compared to 1.75 (1.59) in the full sample. Errors here do not include the statistics on the electron fake rates. While neither of these two tests by itself is awfully statistically significant, the two combined may be indicative of a trend that supports the hypothesis.

For electrons we also considered closure test in a Wjets sample. However, the statistics is so poor that no firm conclusions can be drawn. We observe 6 ± 2.45 and predict 7.5 ± 0.3 (V1), 6.6 ± 0.9 (V2), and 4.9 ± 1.1 (V3).

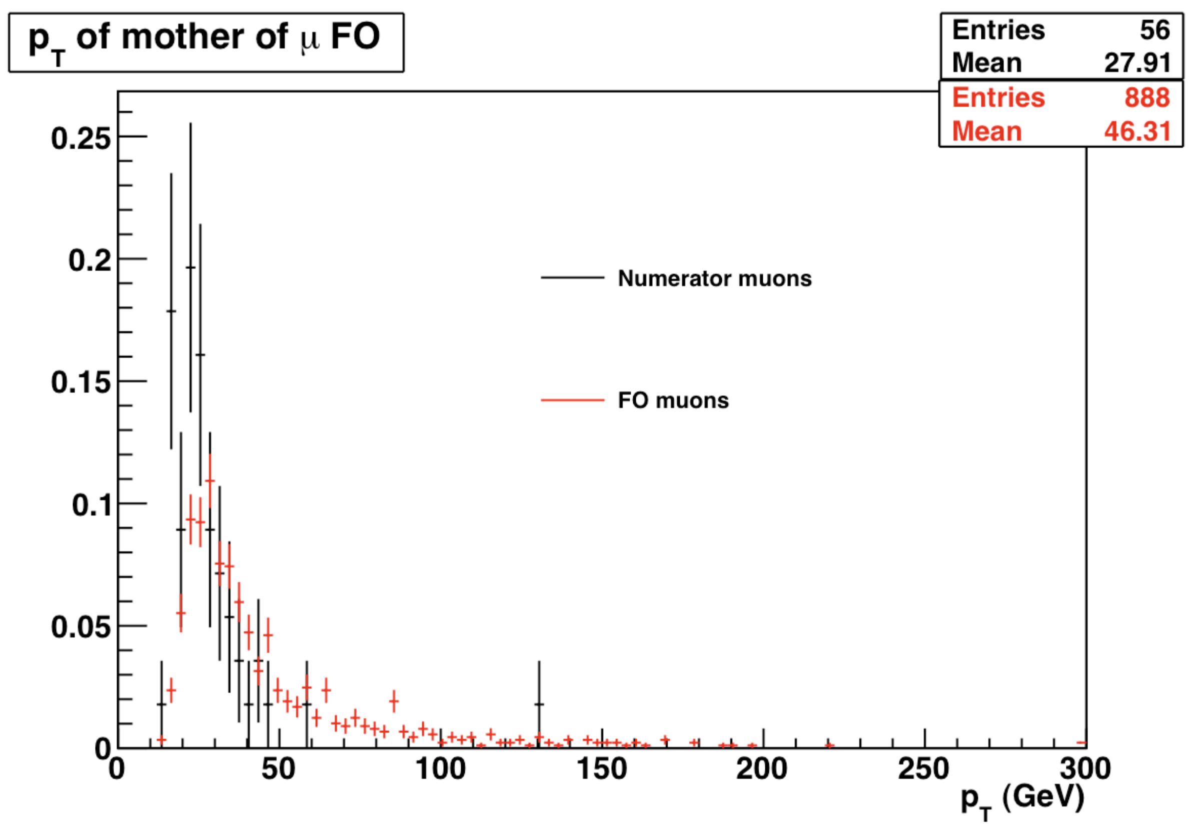

Figure 9: The b-quark pT at generator level for b-quarks that result in muon FOs, overlayed with the pT for those b-quarks that result in numerator muons. Both curves are normalized to the same area.

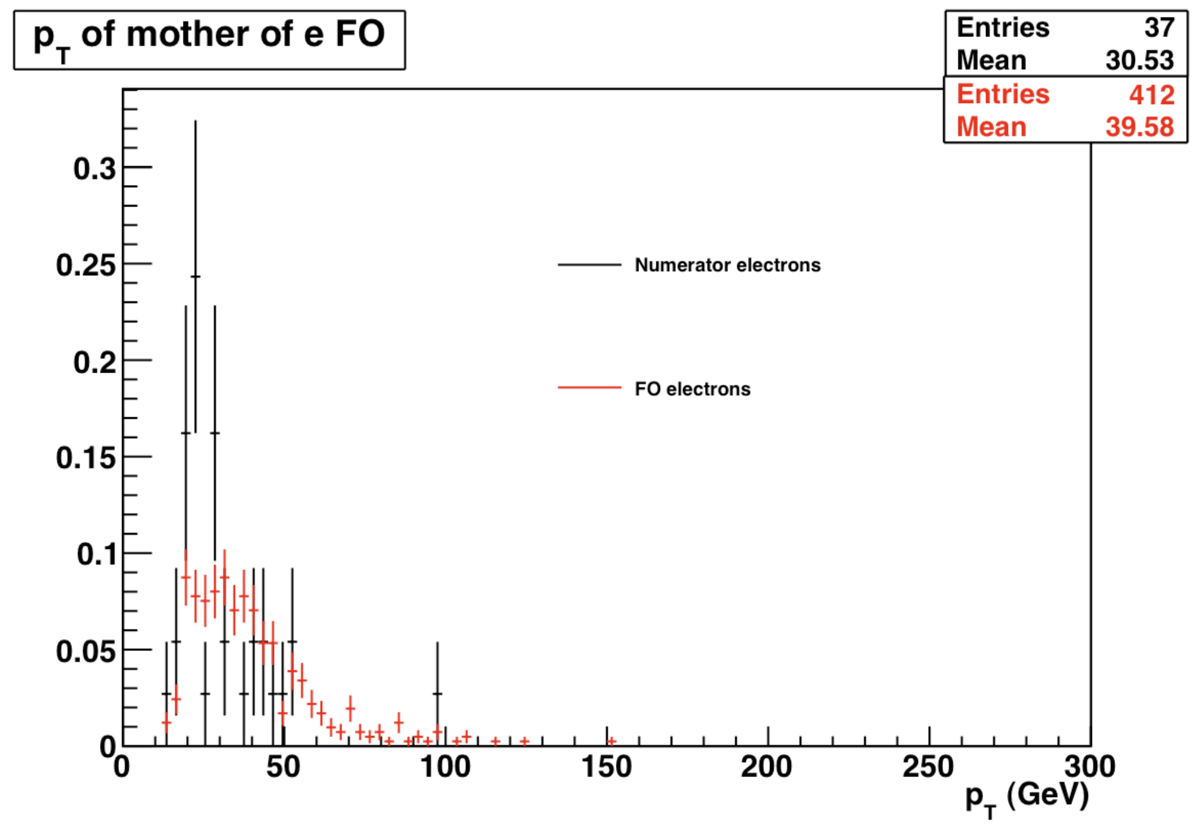

Figure 10: The b-quark pT at generator level for b-quarks that result in electron FOs, overlayed with the pT for those b-quarks that result in numerator electrons. Both curves are normalized to the same area.

13.4 Summary and Conclusions relevant to same sign analysis

In summary, we find that our fake rates derived with a pfjet $P_T > 40$ GeV should successfully predict the muon and electron fakes to within a $\pm 50\%$ systematic error. For electrons we will use V3 of the fake rates, the one that extrapolates in isolation only, as it is much more stable with regard to the composition of the background, and stable enough with regard to P_T variations. Furthermore, V3 is much less problematic in its application to the same sign analysis given the electron triggers available. Electron triggers require Id for a significant fraction of the 2010 luminosity. However, it is possible to make a trigger selection that does not require isolation without sacrificing much if any trigger efficiency. Using V3 we thus do not have to deal with trigger biases in the fake rate application to the final signal sample.

The systematics thus covers anything between a x2 overprediction (covered by -50% error) and a 1/3 underprediction (covered by $+50\%$ error).

14 FR for the WW analysis

14.1 Results relevant to the WW analysis

The fake rates used in the WW analysis are obtained from the EG, EGMonitor, and Mu datasets of Section 2.4 using the electron and muon triggers of Section 7.2. Events must have a pfjet of $P_T > 15$ GeV (uncorrected) that is well separated ($\Delta R > 1$) from the lepton candidate. Results for varying jet thresholds are presented in Section 14.2.

The choice of a 15 GeV threshold for the jet was made to roughly tailor the P_T distributions of the jets from which the FR is obtained to the P_T distributions of jets in the $W+$ jets background to the WW selection.

The 2D fake rates are summarized in Tables 31 to 43 for Monte Carlo and data.

Table 31: The FO_10 muon FR for the WW analysis in Monte Carlo.

$\lvert\eta\rvert$ \ p_T	20.000 – 25.000	25.000 – 30.000	30.000 – 35.000
0.000 – 1.000	0.0957 ± 0.0003	0.0818 ± 0.0005	0.0725 ± 0.0008
1.000 – 1.479	0.1279 ± 0.0005	0.1129 ± 0.0009	0.0971 ± 0.0014
1.479 – 2.000	0.1547 ± 0.0006	0.1405 ± 0.0011	0.1266 ± 0.0017
2.000 – 2.500	0.1672 ± 0.0008	0.1533 ± 0.0015	0.1439 ± 0.0026

Table 32: The FO_10_d0 muon FR for the WW analysis in Monte Carlo.

Table 33:

$\lvert\eta\rvert$ \ p_T	20.000 – 25.000	25.000 – 30.000	30.000 – 35.000
0.000 – 1.000	0.0824 ± 0.0003	0.0706 ± 0.0004	0.0627 ± 0.0007
1.000 – 1.479	0.1098 ± 0.0005	0.0973 ± 0.0008	0.0842 ± 0.0012
1.479 – 2.000	0.1322 ± 0.0005	0.1208 ± 0.0009	0.1093 ± 0.0015
2.000 – 2.500	0.1412 ± 0.0007	0.1309 ± 0.0013	0.1237 ± 0.0023

Table 34: The FO_10 muon FR for the WW analysis in data.

$\lvert\eta\rvert$ \ p_T	20.000 – 25.000	25.000 – 30.000	30.000 – 35.000
0.000 – 1.000	0.1364 ± 0.0012	0.1316 ± 0.0022	0.1265 ± 0.0036
1.000 – 1.479	0.1662 ± 0.0021	0.1688 ± 0.0038	0.1525 ± 0.0061
1.479 – 2.000	0.2007 ± 0.0024	0.1975 ± 0.0043	0.1986 ± 0.0074
2.000 – 2.500	0.2126 ± 0.0043	0.2109 ± 0.0081	0.2320 ± 0.0145

Table 35: The FO_10_d0 muon FR for the WW analysis in data.

$\lvert\eta\rvert$ \ p_T	20.000 – 25.000	25.000 – 30.000	30.000 – 35.000
0.000 – 1.000	0.1195 ± 0.0011	0.1152 ± 0.0019	0.1110 ± 0.0032
1.000 – 1.479	0.1443 ± 0.0018	0.1474 ± 0.0034	0.1342 ± 0.0054
1.479 – 2.000	0.1747 ± 0.0021	0.1719 ± 0.0038	0.1757 ± 0.0067
2.000 – 2.500	0.1848 ± 0.0038	0.1845 ± 0.0072	0.2085 ± 0.0133

Table 36: The V1 electron FR for the WW analysis in Monte Carlo.

$\lvert\eta\rvert$ \ p_T	20.000 – 25.000	25.000 – 30.000	30.000 – 35.000
0.000 – 1.000	0.0089 ± 0.0024	0.0096 ± 0.0036	0.0151 ± 0.0075
1.000 – 1.479	0.0143 ± 0.0041	0.0248 ± 0.0077	0.0296 ± 0.0146
1.479 – 2.000	0.0080 ± 0.0028	0.0060 ± 0.0034	0.0100 ± 0.0070
2.000 – 2.500	0.0081 ± 0.0021	0.0059 ± 0.0026	0.0155 ± 0.0069

Table 37: The V2 electron FR for the WW analysis in Monte Carlo.

$\lvert\eta\rvert$ \ p_T	20.000 – 25.000	25.000 – 30.000	30.000 – 35.000
0.000 – 1.000	0.0667 ± 0.0172	0.0579 ± 0.0212	0.0678 ± 0.0327
1.000 – 1.479	0.1008 ± 0.0276	0.1515 ± 0.0441	0.1739 ± 0.0790
1.479 – 2.000	0.0635 ± 0.0217	0.0370 ± 0.0210	0.0625 ± 0.0428
2.000 – 2.500	0.0387 ± 0.0098	0.0219 ± 0.0097	0.0413 ± 0.0181

Table 38: The V3 electron FR for the WW analysis in Monte Carlo.

$\lvert\eta\rvert$ \ p_T	20.000 – 25.000	25.000 – 30.000	30.000 – 35.000
0.000 – 1.000	0.0809 ± 0.0207	0.1029 ± 0.0369	0.2857 ± 0.1207
1.000 – 1.479	0.1237 ± 0.0334	0.2564 ± 0.0699	0.2105 ± 0.0935
1.479 – 2.000	0.0964 ± 0.0324	0.0968 ± 0.0531	0.2500 ± 0.1531
2.000 – 2.500	0.1648 ± 0.0389	0.1316 ± 0.0548	0.2632 ± 0.1010

Table 39: The V4 electron FR for the WW analysis in Monte Carlo.

$\lvert\eta\rvert$ \ p_T	20.000 – 25.000	25.000 – 30.000	30.000 – 35.000
0.000 – 1.000	0.0301 ± 0.0079	0.0407 ± 0.0151	0.0678 ± 0.0327
1.000 – 1.479	0.0405 ± 0.0115	0.0952 ± 0.0286	0.0816 ± 0.0391
1.479 – 2.000	0.0290 ± 0.0101	0.0210 ± 0.0120	0.0444 ± 0.0307
2.000 – 2.500	0.0517 ± 0.0130	0.0368 ± 0.0161	0.0862 ± 0.0369

Table 40: The V1 electron FR for the WW analysis in data.

$\lvert\eta\rvert$ \ p_T	20.000 – 25.000	25.000 – 30.000	30.000 – 35.000
0.000 – 1.000	0.0201 ± 0.0005	0.0157 ± 0.0006	0.0154 ± 0.0009
1.000 – 1.479	0.0262 ± 0.0008	0.0219 ± 0.0010	0.0194 ± 0.0014
1.479 – 2.000	0.0190 ± 0.0006	0.0187 ± 0.0009	0.0162 ± 0.0012
2.000 – 2.500	0.0136 ± 0.0004	0.0148 ± 0.0005	0.0133 ± 0.0008

Table 41: The V2 electron FR for the WW analysis in data.

$\lvert\eta\rvert$ \ p_T	20.000 – 25.000	25.000 – 30.000	30.000 – 35.000
0.000 – 1.000	0.1168 $\pm$ 0.0027	0.1014 $\pm$ 0.0040	0.1084 $\pm$ 0.0062
1.000 – 1.479	0.1439 $\pm$ 0.0039	0.1373 $\pm$ 0.0060	0.1271 $\pm$ 0.0085
1.479 – 2.000	0.1140 $\pm$ 0.0035	0.1165 $\pm$ 0.0053	0.1028 $\pm$ 0.0071
2.000 – 2.500	0.0512 $\pm$ 0.0013	0.0553 $\pm$ 0.0020	0.0496 $\pm$ 0.0028

Table 42: The V3 electron FR for the WW analysis in data.

$\lvert\eta\rvert$ \ p_T	20.000 – 25.000	25.000 – 30.000	30.000 – 35.000
0.000 – 1.000	0.1423 $\pm$ 0.0032	0.1431 $\pm$ 0.0055	0.1528 $\pm$ 0.0085
1.000 – 1.479	0.1681 $\pm$ 0.0045	0.1631 $\pm$ 0.0071	0.1743 $\pm$ 0.0114
1.479 – 2.000	0.1874 $\pm$ 0.0055	0.2160 $\pm$ 0.0092	0.2289 $\pm$ 0.0146
2.000 – 2.500	0.2282 $\pm$ 0.0052	0.2686 $\pm$ 0.0086	0.2540 $\pm$ 0.0126

Table 43: The V4 electron FR for the WW analysis in data.

$\lvert\eta\rvert$ \ p_T	20.000 – 25.000	25.000 – 30.000	30.000 – 35.000
0.000 – 1.000	0.0605 $\pm$ 0.0014	0.0520 $\pm$ 0.0021	0.0532 $\pm$ 0.0031
1.000 – 1.479	0.0628 $\pm$ 0.0018	0.0555 $\pm$ 0.0026	0.0545 $\pm$ 0.0038
1.479 – 2.000	0.0610 $\pm$ 0.0019	0.0631 $\pm$ 0.0029	0.0607 $\pm$ 0.0043
2.000 – 2.500	0.0754 $\pm$ 0.0019	0.0890 $\pm$ 0.0032	0.0852 $\pm$ 0.0047

14.2 Jet dependence relevant to the WW analysis

To quantify the jet P_T dependence we use the FR obtained from events with pfjets of $P_T > 15$ GeV to predict the rate of leptons in events with jets of $P_T > 5$, 30, 50 GeV and compare with the observation. We make these comparisons for both data and Monte Carlo, for the samples from which the FR were derived.

These comparisons are shown in Tables 44 to 47.

We believe that the jet P_T dependence is due to the isolation extrapolation. Lepton candidates in high P_T jet events tend to have a smaller probability of passing an isolation requirement. The tables show that this trend is well reproduced in data and Monte Carlo. Note that the V2 electron fake rate, which does not extrapolate in isolation (see 7.6) gives the most stable results with respect to jet P_T variations.

Table 44: Observed numbers of muons with $P_T > 20$ GeV in Monte Carlo QCD events as a function of the minimum jet P_T in the event compared with the prediction from the FR method. The FR was obtained from the jet $P_T > 15$ GeV sample. The fake rate is applied to the Pt15 muon enriched dataset as described in Section 2.4.

Jet Requirement	Mu observed	FO_10 pred	FO_10 obs/pred	FO_10_d0 pred	FO_10_d0 obs/pred
$p_T > 5$ GeV	361302	338607	1.07	338346	1.07
$p_T > 30$ GeV	141150	212101	0.67	213127	0.66
$p_T > 50$ GeV	26962	72729.3	0.37	73779	0.37

Table 45: Observed numbers of muons with $P_T > 20$ GeV in data events as a function of the minimum jet P_T in the event compared with the prediction from the FR method. The FR was obtained from the jet $P_T > 15$ GeV sample.

Jet Requirement	Mu observed	FO_10 pred	FO_10 obs/pred	FO_10_d0 pred	FO_10_d0 obs/pred
$p_T > 5$ GeV	40207	35846.1	1.12	35820.6	1.12
$p_T > 30$ GeV	16600	24513.1	0.68	24671.6	0.67
$p_T > 50$ GeV	3869	9466.14	0.41	9651.21	0.40

Table 46: Observed numbers of electrons with $P_T > 20$ GeV in data events as a function of the minimum jet P_T in the event compared with the prediction from the FR method. The FR was obtained from the jet $P_T > 15$ GeV sample. Shown here are only V1 and V2.

Jet Requirement	Ele observed	V1 pred	V1 obs/pred	V2 pred	V2 obs/pred
$p_T > 5$ GeV	11538	10499.6	1.10	11069.1	1.04
$p_T > 30$ GeV	4128	5762.82	0.72	4594.61	0.90
$p_T > 50$ GeV	940	1858.89	0.51	1173.11	0.80

Table 47: Observed numbers of electrons with $P_T > 20$ GeV in data events as a function of the minimum jet P_T in the event compared with the prediction from the FR method. The FR was obtained from the jet $P_T > 15$ GeV sample. Shown here are only V3 and V4.

Jet Requirement	Ele observed	V3 pred	V3 obs/pred	V4 pred	V4 obs/pred
$p_T > 5$ GeV	11538	10093.9	1.14	10228.4	1.13
$p_T > 30$ GeV	4128	5564.76	0.74	5747.24	0.72
$p_T > 50$ GeV	940	1737.72	0.54	1899.14	0.50

14.3 Closure test relevant to the WW analysis

We apply the full WW analysis selection (EWK-10-009) to $W+$ jets Monte Carlo, and use the FR obtained from QCD Monte Carlo to predict the rate of dileptons with one real and one fake lepton. We require that one of the leptons passes the nominal lepton selection, and apply the FR method to events where a second lepton passes the FO selection but not the nominal selection. We looked at the generator level truth information and removed events originating from $W\gamma$ as the FR method is not meant to predict its contribution to the final yield, removed the number of fake electrons in the electron-muon channel when reporting the number of muon fakes, and removed the number of fake muons in the electron-muon channel when reporting the number of electron fakes. The yields listed below are scaled to an integrated luminosity of 34.9 pb^{-1}. The errors include statistical errors on both FO yields and fake rates.

For muons we find 0.26 ± 0.15 fake muon events while the FR prediction is 0.39 ± 0.07. This is for the FO_10 muon fake rates. The FO_10_d0 fake rate prediction is 0.37 ± 0.06.

For electrons we find 1.64 ± 0.41 fake electron events while the FR prediction is 2.45 ± 0.41. This is for the V2 electron fake rates. The V1, V3 and V4 fake rate predictions are 2.35 ± 0.39, 1.69 ± 0.41 and 1.89 ± 0.33, respectively.

All predictions are good to within the statistical and $\pm 50\%$ systematic errors we will assign.

15 FR for the opposite sign analysis

15.1 Results relevant to the OS analysis analysis

The fake rates used in the opposite analysis are obtained from the datasets of Section 2.5 using the electron and muon triggers of Section 8.2. Events must have a caloJet of $P_T > 15$ GeV (uncorrected), and in case that there is only one such jet in the event, this jet must be well separated ($\Delta R > 1$) from the lepton candidate.

The choice of a 15 GeV threshold for the jet was made to roughly tailor the P_T distributions of the jets from which the FR is obtained to the P_T distributions of jets in the $W+$ jets $t\bar{t}$ samples. The effects of varying this choice are given in Section 15.2.

The 2D fake rates are summarized in Tables 48 and 49.

Table 48: The muon FR for the opposite sign analysis.

$|\eta|$ \ p_T	10.000 – 15.000	15.000 – 20.000	20.000 – 25.000	25.000 – 30.000	30.000 – 35.000
0.000 – 1.000	0.3203 $\pm$ 0.0012	0.2461 $\pm$ 0.0026	0.2364 $\pm$ 0.0051	0.2253 $\pm$ 0.0085	0.2243 $\pm$ 0.0139
1.000 – 1.479	0.3531 $\pm$ 0.0019	0.2840 $\pm$ 0.0042	0.2571 $\pm$ 0.0081	0.2566 $\pm$ 0.0136	0.2407 $\pm$ 0.0213
1.479 – 2.000	0.3630 $\pm$ 0.0019	0.2987 $\pm$ 0.0042	0.2794 $\pm$ 0.0081	0.2798 $\pm$ 0.0134	0.3277 $\pm$ 0.0231
2.000 – 2.500	0.3603 $\pm$ 0.0033	0.3029 $\pm$ 0.0071	0.2685 $\pm$ 0.0136	0.3003 $\pm$ 0.0251	0.2773 $\pm$ 0.0410

Table 49: The V3 electron FR for the opposite sign analysis.

$|\eta|$ \ p_T	10.000 – 15.000	15.000 – 20.000	20.000 – 25.000	25.000 – 30.000	30.000 – 35.000
0.000 – 1.000	0.4102 $\pm$ 0.0024	0.3344 $\pm$ 0.0025	0.3081 $\pm$ 0.0033	0.3081 $\pm$ 0.0046	0.3095 $\pm$ 0.0064
1.000 – 1.479	0.4235 $\pm$ 0.0027	0.3450 $\pm$ 0.0030	0.3218 $\pm$ 0.0039	0.3145 $\pm$ 0.0054	0.3367 $\pm$ 0.0075
1.479 – 2.000	0.4112 $\pm$ 0.0034	0.3367 $\pm$ 0.0033	0.3204 $\pm$ 0.0039	0.3218 $\pm$ 0.0052	0.3170 $\pm$ 0.0069
2.000 – 2.500	0.4521 $\pm$ 0.0030	0.3745 $\pm$ 0.0027	0.3612 $\pm$ 0.0034	0.3761 $\pm$ 0.0047	0.3996 $\pm$ 0.0066

15.2 Jet dependence relevant to opposite sign analysis

The dependence of the FR on the minimum jet P_T requirement is shown in Figures 11 and 12.

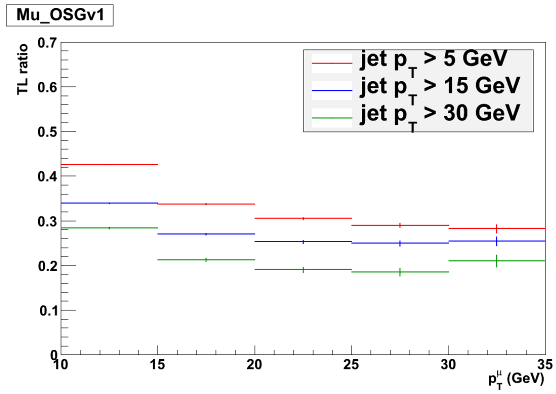

Figure 11: The muon fake rate as a function of P_T for the opposite sign analysis obtained from data with different minimum jet P_T requirements. ($P_T > 5, 15, 30$ GeV for uncorrected caloJets).

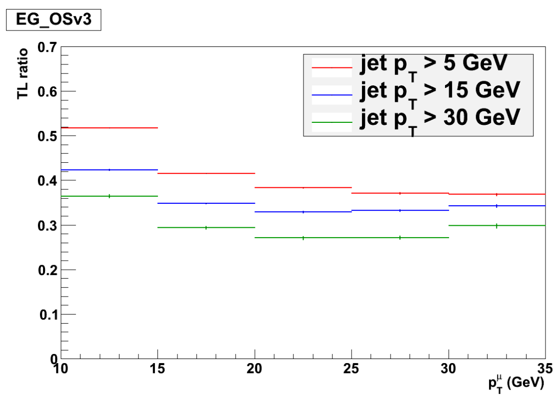

Figure 12: The electron fake rates as a function of P_T for the top analysis obtained from data with different minimum jet P_T requirements. ($P_T > 5, 15, 30$ GeV for uncorrected caloJets).

There is a non negligible jet P_T dependence to the FR. A better way to quantify the jet P_T dependence is to use the FR obtained from events with jets of $P_T > 15$ GeV to predict the rate of leptons in events with jets of $P_T > 5, 30, 50$ GeV and compare with the observation. These comparisons are shown in Tables 50 and 51[4].

Table 50: Observed numbers of muons with $P_T > 10$ in data QCD events as a function of the minimum jet P_T in the event compared with the prediction from the FR method. The FR was obtained from the jet $P_T > 15$ GeV sample.

Jet Requirement	Muons observed	Muons predicted	Observed/predicted
$P_T > 5$ GeV	498,202	119,258	1.44
$P_T > 30$ GeV	14,399	19,072	0.75
$P_T > 50$ GeV	1,659	2,329	0.71

We believe that the jet P_T dependence is due to the isolation extrapolation. Lepton candidates in high P_T jet events tend to have a smaller probability of passing an isolation requirement.

15.3 Closure test relevant to OS analysis

We apply the FR obtained in QCD Monte Carlo to predict the rate of dileptons with one real and one fake lepton in $W+$ jets and $t\bar{t}$ Monte Carlo passing the preselection of Reference [3]. The results (normalized to 35 pb^{-1}) are

[4] The Z-veto, tcMet, and transverse mass requirements from Section 3.2 are included.

Table 51: Observed numbers of electrons with $P_T > 10$ in data QCD events as a function of the minimum jet P_T in the event compared with the prediction from the FR method (V3 version). The FR was obtained from the jet $P_T > 15$ GeV sample.

Jet Requirement	Ele observed	Ele predicted	Observed/predicted
$P_T > 5$ GeV	462,143	336,948	1.37
$P_T > 30$ GeV	20,417	26,204	0.78
$P_T > 50$ GeV	2,529	3,394	0.75

the following:

- in $t\bar{t}$ MC we observe 1.7 ± 0.1 (stat.) events and we predict 2.1
- in $W+$ jets MC we observe 0.3 ± 0.2 events and we predict 0.8

We conclude that we get closure within the 50% systematic uncertainty associated with the FR method.

15.4 Jet flavor dependence relevant to opposite sign analysis

Here we test the stability of the FR in data as a function of jet flavor. The idea is to use b-tagging to enrich the sample in heavy flavor, and to use the FR obtained from generic jets to predict the lepton yield.

For this exercise we use FR obtained with a jet $P_T > 15$ GeV requirement. We apply the FR to a sample with

1. at least one caloJet of uncorrected $P_T > 15$ GeV well separated from the lepton under consideration
2. this caloJet must be btagged by `simpleSecondaryVertexHighEffBJetTag` with medium operating point (cut at 1.74 according to Reference [16]).
3. the Z-veto, tcMET, and transverse mass requirements from Section 3.2.
4. triggered by the lepton triggers of Section 8.2

The 15 GeV requirement is chosen to be the same as the requirement in the FR determination so as to factor out, at least to first order, the known jet P_T dependence (Section 15.2). The results are shown in Table 52. The agreement between the observation and the prediction is within about 20% for both muons and electrons.

Table 52: Observed and predicted number of leptons of $P_T > 10$ GeV in the heavy flavor enriched sample.

Observed muons	Predicted Muons	Observed electrons	V3-Predicted electrons
4407	5427	2721	3825

References

[1] C. Campagnari *at al.*, AN-CMS2009/050.

[2] C. Campagnari *at al.*, AN-CMS2009/042.

[3] D. Barge *at al.*, AN-CMS2010/370.

[4] D. Barge *at al.*, AN-CMS2009/138.

[5] TOP-10-004

[6] D. Barge *et al.*, AN-CMS2010/64 and SUS-10-001.

[7] https://twiki.cern.ch/twiki/bin/viewauth/CMS/Top4ICHEP2010CommonDataSet#JSON_Files_version_13_August_201

[8] https://twiki.cern.ch/twiki/bin/view/CMS/TopCommonDataSet#Nov_5_Certification

[9] http://arxiv.org/abs/1010.5994; D. Barge *et al.* AN-CMS2010/258.

[10] D. Barge *et al.* AN-CMS2010/247.

[11] https://twiki.cern.ch/twiki/bin/viewauth/CMS/SimpleCutBasedEleID

[12] https://hypernews.cern.ch/HyperNews/CMS/get/ewk/378.html

[13] D. Barge *at al.*, AN-CMS2009/159.

[14] W. Andrews *at al.*, AN-CMS2009/041.

[15] http://indico.cern.ch/getFile.py/access?contribId=6&resId=0&materialId=slides&confId=93821

[16] https://twiki.cern.ch/twiki/bin/view/CMS/BTagPerformanceOP