Available on CMS information server

CMS AN 2012/XXX

The Compact Muon Solenoid Experiment

Analysis Note

The content of this note is intended for CMS internal use and distribution only

October 4, 2012

Search for Direct Stop Quark Pair Production in the Single Lepton Channel at 8 TeV

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

Fermilab National Accelerator Laboratory, Batavia, USA

D. Barge, C. Campagnari, A. George, F. Golf, J. Gran, D. Kovalskyi, V. Krutelyov

University of California, Santa Barbara, Santa Barbara, USA

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

University of California, San Diego, San Diego, USA

Abstract

This note describes a search for direct stop quark pair production in the single lepton channel using 9.7 fb^{-1} of pp collision data at $\sqrt{s} = 8$ TeV taken with the CMS detector in 2012. A search for an excess of events over the Standard Model prediction is performed in a sample with a single isolated electron or muon, several jets, missing transverse energy and large transverse mass.

Contents

1 Introduction

This note presents a search for the production of supersymmetric (SUSY) stop quark pairs in events with a single isolated lepton, several jets, missing transverse energy, and large transverse mass. We use the full 2012 data sample, corresponding to an integrated luminosity of 9.7 fb^{-1}. This search is of theoretical interest because of the critical role played by the stop quark in solving the hierarchy problem in SUSY models. This solution requires that the stop quark be light, less than a few hundred GeV and hence within reach for direct pair production. We focus on two decay modes $\tilde{t} \to t\chi_1^0$ and $\tilde{t} \to b\chi_1^+$ which are expected to have large branching fractions if they are kinematically accessible, leading to:

- $pp \to \tilde{t}\bar{\tilde{t}} \to t\bar{t}\chi_1^0\chi_1^0$, and
- $pp \to \tilde{t}\bar{\tilde{t}} \to b\bar{b}\chi_1^+\chi_1^- \to b\bar{b}W^+W^-\chi_1^0\chi_1^0$.

Both of these signatures contain high transverse momentum (p_{T}) jets including two b-jets, and missing transverse energy ($E_{\mathrm{T}}^{\mathrm{miss}}$) due to the invisible χ_1^0 lightest SUSY particles (LSP's). In addition, the presence of two W bosons leads to a large branching fraction to the single lepton final state. Hence we require the presence of exactly one isolated, high p_{T} electron or muon, which provides significant suppression of several backgrounds that are present in the all-hadronic channel. The largest backgrounds for this signature are semi-leptonic $\mathrm{t\bar{t}}$ and W+jets. These backgrounds contain a single leptonically-decaying W boson, and the transverse mass (M_{T}) of the lepton-neutrino system has a kinematic endpoint requiring $M_{\mathrm{T}} < M_W$. For signal stop quark events, the presence of additional LSP's in the final states allows the M_{T} to exceed M_W. Hence we search for an excess of events with large M_{T}. The dominant background in this kinematic region is dilepton $\mathrm{t\bar{t}}$ where one of the leptons is not identified, since the presence of two neutrinos from leptonically-decaying W bosons allows the M_{T} to exceed M_W. Backgrounds are estimated from Monte Carlo (MC) simulation, with careful validation and determination of scale factors and corresponding uncertainties based on data control samples.

The expected stop quark pair production cross section (see Fig. 1) varies between O(10) pb for $m_{\tilde{t}} =$ 200 GeV and O(0.01) pb for $m_{\tilde{t}} = 500$ GeV. The critical challenge of this analysis is due to the fact that for light stop quarks ($m_{\tilde{t}} \approx m_t$), the production cross section is large but the kinematic distributions, in particular M_{T}, are very similar to SM $\mathrm{t\bar{t}}$ production. In this regime it becomes very difficult to distinguish the signal and background. For large stop quark mass the kinematic distributions differ from those in SM $\mathrm{t\bar{t}}$ production, but the cross section decreases rapidly, reducing the signal-to-background ratio.

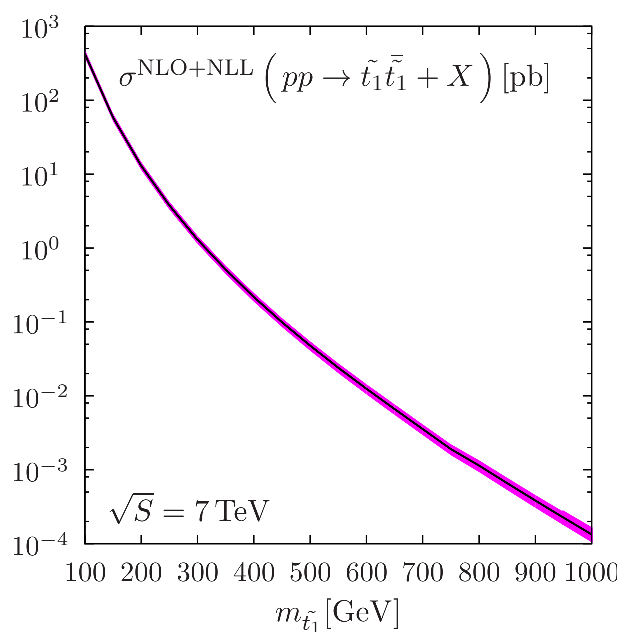

Figure 1: The stop quark pair production cross section in pb, as a function of the stop quark mass. AT SOME POINT WE NEED TO GET A FIGURE FOR 8 TEV CM ENERGY.

2 Overview and Strategy for Background Determination

[THIS SECTION IS NOW MORE OR LESS OK. NEED TO FIX THE "XX" IN FORWARD SECTION REFERENCES]

We are searching for a $t\bar{t}\chi^0\chi^0$ or $WbW\bar{b}\chi^0\chi^0$ final state (after top decay in the first mode, the final states are actually the same). So to first order this is "$t\bar{t}+$ extra $E_{\mathrm{T}}^{\mathrm{miss}}$".

We work in the $\ell+$ jets final state, where the main background is $t\bar{t}$. We look for $E_{\mathrm{T}}^{\mathrm{miss}}$ inconsistent with $W \to \ell\nu$. We do this by concentrating on the $\ell\nu$ transverse mass (M_T), since except for resolution effects, $M_T < M_W$ for $W \to \ell\nu$. Thus, the initial analysis is simply a counting experiment in the tail of the M_T distribution.

The event selection is one-and-only-one high p_{T} isolated lepton, four or more jets, and some moderate $E_{\mathrm{T}}^{\mathrm{miss}}$ cut. At least one of the jets has to be btagged to reduce $W+$ jets. The event sample is then dominated by $t\bar{t}$, but there are also contributions from $W+$ jets, single top, dibosons, as well as rare SM processes such as ttW.

The $t\bar{t}$ events in the M_T tail can be broken up into two categories: (i) $t\bar{t} \to \ell+$ jets and (ii) $t\bar{t} \to \ell^+\ell^-$ where one of the two leptons is not found by the second-lepton-veto (here the second lepton can be a hadronically decaying τ). For a reasonable M_T cut, say $M_T > 150$ GeV, the dilepton background is of order 80% of the total. This is because in dileptons there are two neutrinos from W decay, thus M_T is not bounded by M_W. This is a very important point: while it is true that we are looking in the tail of M_T, the bulk of the background events end up there not because of some exotic $E_{\mathrm{T}}^{\mathrm{miss}}$ reconstruction failure, but because of well understood physics processes. This means that the background estimate can be taken from Monte Carlo (MC) after carefully accounting for possible data/MC differences.

In Section XX we will describe the analysis of various Control Regions (CRs) that are used to test the Monte Carlo model and, if necessary, to extract data/MC scale factors. In this section we give a general description of the procedure. The details of how the final background prediction is assembled are given in Section XX.

The search is performed in a number of signal regions defined by minimum requirements on $E_{\mathrm{T}}^{\mathrm{miss}}$ and M_T. These signal regions are defined in Section XX.

One general point is that in order to minimize systematic uncertainties, the MC background predictions are whenever possible normalized to the bulk of the $t\bar{t}$ data, ie, events passing all of the requirements but with $M_T \approx 80$ GeV. This (mostly) removes uncertainties due to $\sigma(t\bar{t})$, lepton ID, trigger efficiency, luminosity, etc.

2.1 $\ell+$ jets background

The $\ell+$ jets background is dominated by $t\bar{t} \to \ell+$ jets, but also includes some $W+$ jets as well as single top. The MC input used in the background estimation is the ratio of the number of events with M_T in the signal region to the number of events with $M_T \approx 80$ GeV. This ratio is (possibly) corrected by a data/MC scale factor obtained from a study of CRs, as outlined below.

Note that the ratio described above is actually different for $t\bar{t}$/single top and $W+$ jets. This is because in W events there is a significant contribution to the M_T tail from very off-shell W. This contribution is much smaller in top events because $M(\ell\nu)$ cannot excees $M_{top} - M_b$.

For $W+$ jets the ability of the Monte Carlo to model this ratio (R_{wjet}) is tested in a sample of $\ell+$ jets enriched in $W+$ jets by the application of a b-veto. The equivalent ratio for top events (R_{top}) is validated in a sample of well identified $Z \to \ell\ell$ with one lepton added to the $E_{\mathrm{T}}^{\mathrm{miss}}$ calculation. This sample is well suited to testing the resolution effects on the M_T tail, since off-shell effects are eliminated by the Z-mass requirement.

Note that the fact that the ratios are different for $t\bar{t}$/single top and $W+$ jets introduces a systematic uncertainty in the background calculation because one needs to know the relative fractions of these two components in $M_T \approx 80$ GeV lepton + jets sample.

2.2 Dilepton background

To suppress dilepton backgrounds, we veto events with an isolated track of p_{T}> 10 GeV. Being the common feature for electron, muon, and one-prong tau decays, this veto is highly efficient for rejecting $t\bar{t}$ to dilepton events. The remaining dilepton background can be classified into the following categories:

- lepton is out of acceptance ($|\eta| > 2.50$)
- lepton has $p_{\mathrm{T}} < 10$ GeV, and is inside the acceptance
- lepton has $p_{\mathrm{T}} > 10$ GeV, is inside the acceptance, but survives the additional isolated track veto

The last category includes 3-prong tau decays as well as electrons and muons from W decay that fail the isolation requirement. Monte Carlo studies indicate that these three components populate the M_T tail in the proportions of roughly 6%, 47%, 47%. We note that at present we do not attempt to veto 3-prong tau decays as they are only 16% of the total dilepton background according to the MC.

The high M_T dilepton backgrounds come from MC, but their rate is normalized to the $M_T \approx 80$ GeV peak. In order to perform this normalization in data, the non-$t\bar{t}$ (eg, $W+$ jets) events in the M_T peak have to be subtracted off. This also introduces a systematic uncertainty.

There are two types of effects that can influence the MC dilepton prediction: physics effects and instrumental effects. We discuss these next, starting from physics.

First of all, many of our $t\bar{t}$ MC samples (eg: MadGraph) have $\mathrm{BR}(W \to \ell\nu) = \frac{1}{9} = 0.1111$. PDG says $\mathrm{BR}(W \to \ell\nu) = 0.1080 \pm 0.0009$. This difference matters, so the $t\bar{t}$ MC must be corrected to account for this.

Second, our selection is $\ell+$ 4 or more jets. A dilepton event passes the selection only if there are two additional jets from ISR, or one jet from ISR and one jet which is reconstructed from the unidentified lepton, *e.g.*, a three-prong tau. Therefore, all MC dilepton $t\bar{t}$ samples used in the analysis must have their jet multiplicity corrected (if necessary) to agree with what is seen in $t\bar{t}$ data. We use a data control sample of well identified dilepton events with $E_{\mathrm{T}}^{\mathrm{miss}}$ and at least two jets as a template to "adjust" the N_{jet} distribution of the $t\bar{t} \to$ dileptons MC samples.

The final physics effect has to do with the modeling of $t\bar{t}$ production and decay. Different MC models could in principle result in different BG predictions. Therefore we use several different $t\bar{t}$ MC samples using different generators and different parameters, to test the stability of the dilepton BG prediction. All these predictions, **after** corrections for branching ratio and N_{jet} dependence, are compared to each other. The spread is a measure of the systematic uncertainty associated with the $t\bar{t}$ generator modeling.

The main instrumental effect is associated with the efficiency of the isolated track veto. We use tag-and-probe to compare the isolated track veto performance in $Z + 4$ jet data and MC, and we extract corrections if necessary. Note that the performance of the isolated track veto is not exactly the same on e/μ and on one prong hadronic tau decays. This is because the pions from one-prong taus are often accompanied by π^0's that can then result in extra tracks due to phton conversions. We let the simulation take care of that. Note that JES uncertainties are effectively "calibrated away" by the N_{jet} rescaling described above.

2.3 Other backgrounds

Other backgrounds are tW, ttV, dibosons, tribosons, Drell Yan. These are small. They are taken from MC with appropriate scale factors for trigger efficiency, etc.

2.4 Future improvements

Finally, there are possible improvements to this basic analysis strategy that can be added in the future:

- Move from counting experiment to shape analysis. But first, we need to get the counting experiment under control.
- Add an explicit three prong tau veto

- Do something to require that three of the jets in the event be consistent with $t \to Wb, W \to q\bar{q}$. This could help reject some of the dilepton BG in the search for $\widetilde{t} \to t\chi^0$, but is not applicable to the $\widetilde{t} \to b\chi^+$ search.
- Consider the $M(\ell b)$ variable, which is not bounded by M_{top} in $\widetilde{t} \to b\chi^+$

3 Data Samples

[UPDATE]

The datasets used for this analysis are summarized in Tables 1 (data) and 2 (MC). The total integrated luminosity is 9.7 fb^{-1}after applying the official good run list. The main Monte Carlo samples are generated with Madgraph, though samples with alternative generators such as Powheg and MC@NLO are also used for the derivation of systematic uncertainties in the $t\bar{t}$ background prediction. The triggers used to select both the signal and control samples are also summarized in Table. 3.

Dataset Name
Single Lepton Samples
/SingleElectron/Run2012A-13Jul2012-v1/AOD
/SingleMu/Run2012A-13Jul2012-v1/AOD
/SingleElectron/Run2012B-13Jul2012-v1/AOD
/SingleMu/Run2012B-13Jul2012-v1/AOD
/SingleElectron/Run2012C-PromptReco-v*/AOD
/SingleMu/Run2012C-PromptReco-v*/AOD
Dilepton Samples (only used for dilepton control region)
/DoubleElectron/Run2012A-13Jul2012-v1/AOD
/DoubleMu/Run2012A-13Jul2012-v1/AOD
/MuEG/Run2012A-13Jul2012-v1/AOD
/DoubleElectron/Run2012B-13Jul2012-v1/AOD
/DoubleMu/Run2012B-13Jul2012-v1/AOD
/MuEG/Run2012B-13Jul2012-v1/AOD
/DoubleElectron/Run2012C-PromptReco-v*/AOD
/DoubleMu/Run2012C-PromptReco-v*/AOD
/MuEG/Run2012C-PromptReco-v*/AOD

Table 1: Summary of data datasets used.

With Pileup: Processed dataset name is (S3) Summer12_DR53X-PU_S10_START53_V7A-v*/AODSIM (S3) Summer11-PU_S3_START42_V11-v*/AODSIM		
Description	Primary Dataset Name	cross-section [pb]
$\mathrm{t\bar{t}}$	/TTJets_MassiveBinDECAY_TuneZ2Star_8TeV-madgraph-tauola (S3)	225.2
$\mathrm{W} \to \ell\nu$	/WJetsToLNu_TuneZ2Star_8TeV-madgraph-tauola (S3)	31314.0
WW	/WW_TuneZ2Star_8TeV_pythia6_tauola (S3)	45.6
WZ	/WZ_TuneZ2Star_8TeV_pythia6_tauola (S3)	18.2
ZZ	/ZZ_TuneZ2Star_8TeV_pythia6_tauola (S3)	7.4
t (s-chan)	/T_TuneZ2Star_s-channel_8TeV-powheg-tauola (S3)	3.19
$\bar{t}$ (s-chan)	/Tbar_TuneZ2Star_s-channel_8TeV-powheg-tauola (S3)	1.44
t (t-chan)	/T_TuneZ2Star_t-channel_8TeV-powheg-tauola (S3)	41.92
$\bar{t}$ (t-chan)	/Tbar_TuneZ2Star_t-channel_8TeV-powheg-tauola (S3)	22.65
tW	/T_TuneZ2Star_tW-channel-DR_8TeV-powheg-tauola (S3)	7.87
$\bar{t}W$	/Tbar_TuneZ2Star_tW-channel-DR_8TeV-powheg-tauola (S3)	7.87
$Z/\gamma^* \to \ell\ell$	/DYJetsToLL_TuneZ2Star_M-50_8TeV-madgraph-tarball (S3)	3532.8
$\mathrm{t\bar{t}}W$	/TTW_TuneZ2Star_8TeV-madgraph (S3)	0.1633
$\mathrm{t\bar{t}}Z$	/TTZ_TuneZ2Star_8TeV-madgraph (S3)	0.139
$\mathrm{t\bar{t}}\gamma$	/TTPhoton_TuneZ2Star_8TeV-madgraph (S3)	0.6545
$\mathrm{WW}\gamma$	/WWPhoton_TuneZ2Star_8TeV-madgraph (S3)	0.177
WWZ	/WWZNoGstar_TuneZ2Star_8TeV-madgraph (S3)	0.0268
WWW	/WWW_TuneZ2Star_8TeV-madgraph (S3)	0.038
WZZ	/WZZNoGstar_TuneZ2Star_8TeV-madgraph (S3)	0.0088
ZZZ	/ZZZNoGstar_TuneZ2Star_8TeV-madgraph (S3)	0.00288
$\tilde{t}\bar{\tilde{t}} \to t\bar{t}\chi_1^0\chi_1^0$	/SMS-T2tt_Mstop-225to1200_mLSP-50to1025_8TeV-Pythia6Z (S2)	scan
$\tilde{t}\bar{\tilde{t}} \to b\bar{b}\chi_1^+\chi_1^-$	/SMS-T2bw_x-0p25to0p75_mStop-50to850_mLSP-50to800_8TeV-Pythia6Z (S2)	scan
$\mathrm{t\bar{t}}$ ($Q^2 \times 2$)	/TTjets_TuneZ2Star_scaleup_8TeV-madgraph-tauola (S3)	225.2
$\mathrm{t\bar{t}}$ ($Q^2 \times 0.5$)	/TTjets_TuneZ2Star_scaledown_8TeV-madgraph-tauola (S3)	225.2
$\mathrm{t\bar{t}}$ ($x_q > 40$ GeV)	/TTjets_TuneZ2Star_matchingup_8TeV-madgraph-tauola (S3)	225.2
$\mathrm{t\bar{t}}$ ($x_q > 10$ GeV)	/TTjets_TuneZ2Star_matchingdown_8TeV-madgraph-tauola (S3)	225.2
$\mathrm{t\bar{t}}$ ($m_{\mathrm{top}} = 178.5$ GeV)	/TTJets_TuneZ2Star_mass178_5_8TeV-madgraph-tauola (S3)	225.2
$\mathrm{t\bar{t}}$ ($m_{\mathrm{top}} = 166.5$ GeV)	/TTJets_TuneZ2Star_mass166_5_8TeV-madgraph-tauola (S3)	225.2
$\mathrm{t\bar{t}}$	/TT_TuneZ2Star_8TeV-powheg-tauola (S3)	225.2

Table 2: Summary of Monte Carlo datasets used. TO BE UPDATED.

Triggers
Single Muon Sample
`HLT_IsoMu17_v*`
`HLT_IsoMu24_v*`
`HLT_IsoMu30_eta2p1_v*`
Single Electron Sample
`HLT_Ele25_CaloIdVT_TrkIdT_CentralTriJet30_v*`
`HLT_Ele25_CaloIdVT_TrkIdT_TriCentralJet30_v*`
`HLT_Ele25_CaloIdVT_CaloIsoT_TrkIdT_TrkIsoT_TriCentralJet30_v*`
`HLT_Ele25_CaloIdVT_CaloIsoT_TrkIdT_TrkIsoT_TriCentralPFJet30_v*`
Dimuon Sample (only used for dilepton control regions)
`HLT_DoubleMu7_v*`
`HLT_Mu13_Mu7_v*`
`HLT_Mu13_Mu8_v*`
`HLT_Mu17_Mu8_v*`
Electron-Muon Sample (only used for dilepton control regions)
`HLT_Mu17_Ele8_CaloIdL_v*`
`HLT_Mu8_Ele17_CaloIdL_v*`
`HLT_Mu17_Ele8_CaloIdT_CaloIsoVL_v*`
`HLT_Mu8_Ele17_CaloIdT_CaloIsoVL_v*`
Dielectron Sample (only used for dilepton control regions)
`HLT_Ele17_CaloIdL_CaloIsoVL_Ele8_CaloIdL_CaloIsoVL_v*`
`HLT_Ele17_CaloIdT_TrkIdVL_CaloIsoVL_TrkIsoVL_Ele8_CaloIdT_TrkIdVL_CaloIsoVL_TrkIsoVL_v*`
`HLT_Ele17_CaloIdT_CaloIsoVL_TrkIdVL_TrkIsoVL_Ele8_CaloIdT_CaloIsoVL_TrkIdVL_TrkIsoVL_v*`

Table 3: Summary of triggers used. TO BE UPDATED

4 Event Selection

This analysis uses several different control regions in addition to the signal regions. All of these different regions are defined in this section.

4.1 Single Lepton Selection

[UPDATE SELECTION]

The single lepton preselection sample is based on the following criteria

- satisfy the trigger requirement (see Table. 1). Note that the analysis triggers are inclusive single lepton triggers. Dilepton triggers are used only for the dilepton control region.
- select events with one high p_T electron or muon, requiring
  - $p_T > 30$ GeV/c and $|\eta| < 2.1$
  - satisfy the identification and isolation requirements detailed in the same-sign SUSY analysis (SUS-11-010) for electrons and the opposite-sign SUSY analysis (SUS-11-011) for muons
- require at least 4 PF jets in the event with $p_T > 30$ GeV within $|\eta| < 2.5$ out of which at least 1 satisfies the CSV medium working point b-tagging requirement
- require moderate $E_T^{miss} > 50$ GeV

Table 4 shows the yields in data and MC without any corrections for this preselection region.

Table 4: Raw Data and MC predictions without any corrections are shown after preselection.

4.2 Signal Region Selection

[MOTIVATIONAL BLURB ON MET AND MT,
CAN ADD SIGNAL VS. TTBAR MC PLOT
ADD SIGNAL YIELDS FOR AVAILABLE POINTS,
DISCUSS CHOICE SIG REGIONS]

The signal regions (SRs) are selected to improve the sensitivity for the single lepton requirements and cover a range of scalar top scenarios. The M_T and E_T^{miss} variables are used to define the signal regions and the requirements are listed in Table 5.

Signal Region	Minimum M_T [GeV]	Minimum E_T^{miss} [GeV]
SRA	150	100
SRB	120	150
SRC	120	200
SRD	120	250
SRE	120	300

Table 5: Signal region definitions based on M_T and E_T^{miss} requirements. These requirements are applied in addition to the baseline single lepton selection.

Table 6 shows the expected number of SM background yields for the SRs. A few stop signal yields for four values of the parameters are also shown for comparison. The signal regions with looser requirements are sensitive to lower stop masses M($\tilde{t}$), while those with tighter requirements are more sensitive to higher M($\tilde{t}$).

Sample	SRA	SRB	SRC	SRD
$t\bar{t} \to \ell\ell$	700 ± 15	408 ± 12	134 ± 7	43 ± 4
$t\bar{t} \to \ell +$ jets & single top (1ℓ)	111 ± 6	71 ± 5	15 ± 2	4 ± 1
W+jets	58 ± 35	57 ± 35	29 ± 26	26 ± 26
Rare	63 ± 3	40 ± 3	17 ± 2	7 ± 1
Total	932 ± 39	576 ± 38	195 ± 27	80 ± 26

Table 6: Expected SM background contributions, including both muon and electron channels. This is "dead reckoning" MC with no correction. It is meant only as a general guide. The uncertainties are statistical only. ADD SIGNAL POINTS.

4.3 Control Region Selection

[1 PARAGRAPH BLURB RELATING BACKGROUNDS (IN TABLE FROM PREVIOUS SECTION) TO INTRODUCE CONTROL REGIONS]

Control regions (CRs) are used to validate the background estimation procedure and derive systematic uncertainties for some contributions. The CRs are selected to have similar kinematics to the SRs, but have a different requirement in terms of number of b-tags and number of leptons, thus enhancing them in different SM contributions. The four CRs used in this analysis are summarized in Table 7.

4.4 MC Corrections

[UPDATE SECTION]

4.4.1 Corrections to Jets and $E_{\mathrm{T}}^{\mathrm{miss}}$

[UPDATE, ADD FEW MORE DETAILS ON WHAT IS DONE HERE]

The official recommendations from the Jet/MET group are used for the data and MC samples. In particular, the jet energy corrections (JEC) are updated using the official recipe. L1FastL2L3Residual (L1FastL2L3) corrections are applied for data (MC), based on the global tags GR_R_42_V23 (DESIGN42_V17) for data (MC). In addition, these jet energy corrections are propagated to the $E_{\mathrm{T}}^{\mathrm{miss}}$ calculation, following the official prescription for deriving the Type I corrections.

Events with anomalous "rho" pile-up corrections are excluded from the sample since these correspond to events with unphysically large $E_{\mathrm{T}}^{\mathrm{miss}}$ and M_{T} tail signal region. In addition, the recommended MET filters are applied.

4.4.2 Branching Fraction Correction

The leptonic branching fraction used in some of the $t\bar{t}$ MC samples differs from the value listed in the PDG $(10.80 \pm 0.09)\%$. Table. 8 summarizes the branching fractions used in the generation of the various

Selection Criteria	exactly 1 lepton	exactly 2 leptons	1 lepton + isolated track
0 b-tags	CR1) W+Jets dominated: Validate W+Jets M_{T} tail	CR2) apply Z-mass constraint $\to$ Z+Jets dominated: Validate $t\bar{t} \to \ell +$ jets M_{T} tail comparing data vs. MC "pseudo-M_{T}"	CR3) not used
≥ 1 b-tags	SIGNAL REGION	CR4) Apply Z-mass veto $\to t\bar{t} \to \ell\ell$ dominated: Validate "physics" modelling of $t\bar{t} \to \ell\ell$	CR5) $t\bar{t} \to \ell\ell$, $t\bar{t} \to \ell\tau$ and $t\bar{t} \to \ell$fake dominated: Validate τ and fake lepton modeling/ detector effects in $t\bar{t} \to \ell\ell$

Table 7: Summary of signal and control regions.

$t\bar{t}$ MC samples. For $t\bar{t}$ samples with the incorrect leptonic branching fraction, event weights are applied based on the number of true leptons and the ratio of the corrected and incorrect branching fractions.

$t\bar{t}$ Sample - Event Generator	Leptonic Branching Fraction
Madgraph	0.111
MC@NLO	0.111
Pythia	0.108
Powheg	0.108

Table 8: Leptonic branching fractions for the various $t\bar{t}$ samples used in the analysis. The primary $t\bar{t}$ MC sample produced with Madgraph has a branching fraction that is almost 3% higher than the PDG value.

4.4.3 Efficiency Corrections

[TO BE UDPATED WITH T&P STUDIES ON ID, TRIGGER ETC]

5 Control Region Studies

5.1 W+Jets MC Modelling Validation from CR1

The estimate of the uncertainty on this background is based on CR1, defined by applying the full signal selection, but requiring 0 b-tags. SPECIFY THE EXACT REQUIREMENT FOR THE BVETO AND SAY WHETHER THERE IS AN ISOLATED TRACK VETO. The sample is dominanted by W+jets and is thus used to validate the MC modelling of this background.

In Table 9 we show the amount that we need to scale the Wjets MC by in order to have agreement between data and Monte Carlo in the M_T peak region, defined as $XX < M_T < YY$ GeV. These scale factors are not terribly important, but it is reassuring that they are not too different from 1. (ARE THESE SCALED FOR TRIGGER EFFICIENCY???)

Sample	CR1PRESEL	CR1A	CR1B	CR1C	CR1D
Muon M_{T}-SF	0.91 ± 0.03	0.96 ± 0.07	0.88 ± 0.11	1.05 ± 0.21	1.27 ± 0.41
Electron M_{T}-SF	0.82 ± 0.03	0.86 ± 0.06	1.09 ± 0.15	1.24 ± 0.30	1.11 ± 0.40

Table 9: M_{T} peak Data/MC scale factors applied to the single lepton samples and $\mathrm{t\bar{t}} \to \ell\ell$. The raw MC is used for backgrounds from rare processes. CR1PRESEL refers to a sample with $E_{\mathrm{T}}^{\mathrm{miss}} > 50$ GeV. The uncertainties are statistical only.

In Table 10 we compare the data and MC yields in the four M_T signal regions and in a looser control region. We also derive the data/MC scale factors SFR^{e}_{wjet} and SFR^{μ}_{wjet}. The underlying $E_{\mathrm{T}}^{\mathrm{miss}}$ and M_T distributions are shown in Fig. 2 and 3

Sample	CR1PRESEL	CR1A	CR1B	CR1C	CR1D
Muon MC	456 ± 73	174 ± 44	51 ± 7	18 ± 2	10 ± 2
Muon Data	657	246	142	43	12
Muon Data/MC SF: (SFR^{μ}_{wjet})	1.44 ± 0.24	1.41 ± 0.37	2.80 ± 0.47	2.37 ± 0.46	1.23 ± 0.42
Electron MC	396 ± 64	147 ± 36	54 ± 4	19 ± 2	8 ± 2
Electron Data	702	223	144	50	23
Electron Data/MC SF: (SFR^{e}_{wjet})	1.77 ± 0.29	1.52 ± 0.39	2.68 ± 0.30	2.57 ± 0.49	2.73 ± 0.76

Table 10: Yields in M_{T} tail comparing the MC prediction (after applying SFs) to data. CR1PRESEL refers to a sample with $E_{\mathrm{T}}^{\mathrm{miss}} > 50$ GeV and $M_{\mathrm{T}} > 150$ GeV. The uncertainties are statistical only. VERENA MAKE SURE YOU ADD SCALE FACTOR SYMBOLS TO THIS TABLE.

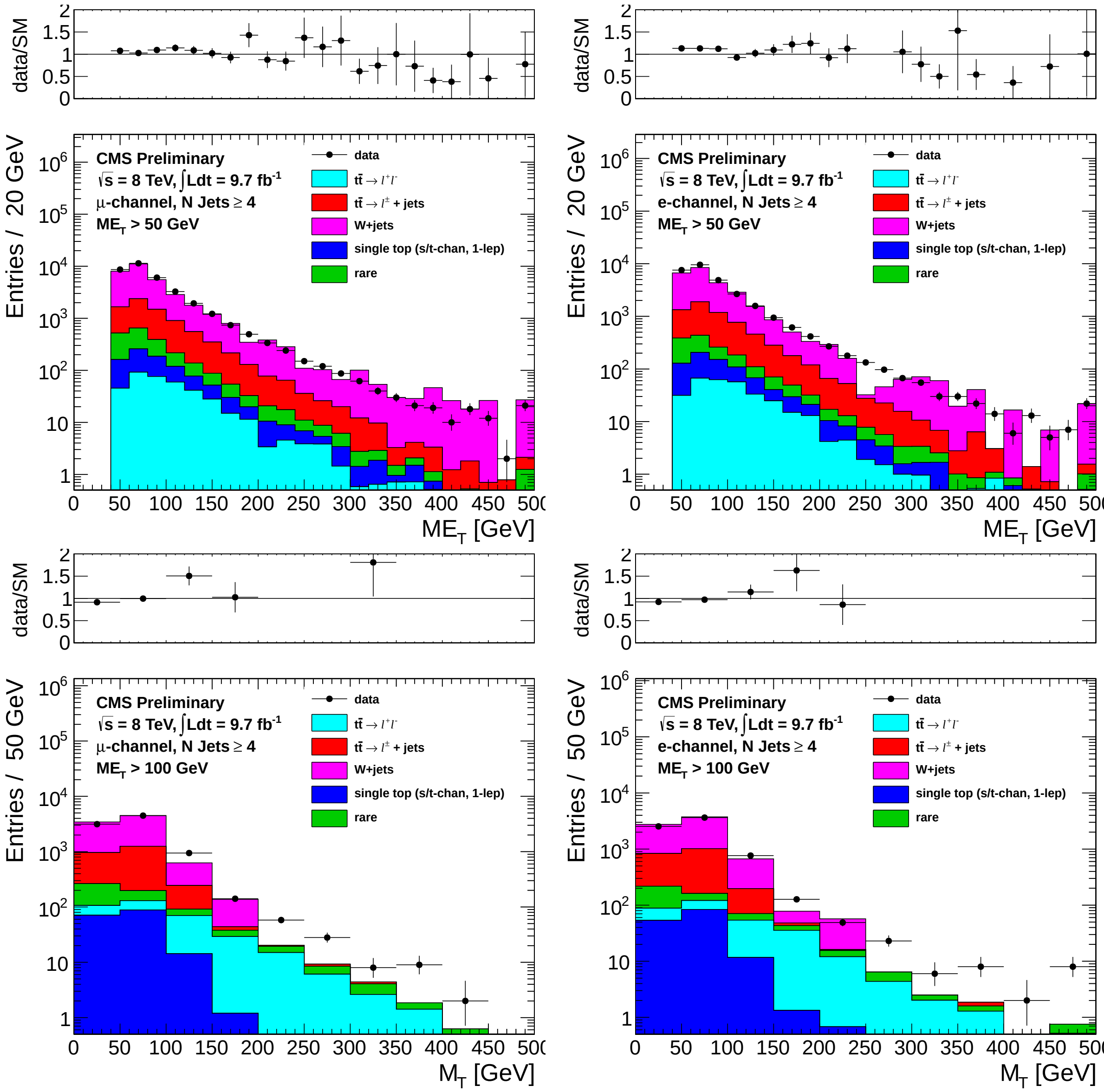

Figure 2: Comparison of the $E_T^{\rm miss}$ (top) and M_T for $E_T^{\rm miss} > 100$ (bottom) distributions in data vs. MC for events with a leading muon (left) and leading electron (right) satisfying the requirements of CR1.

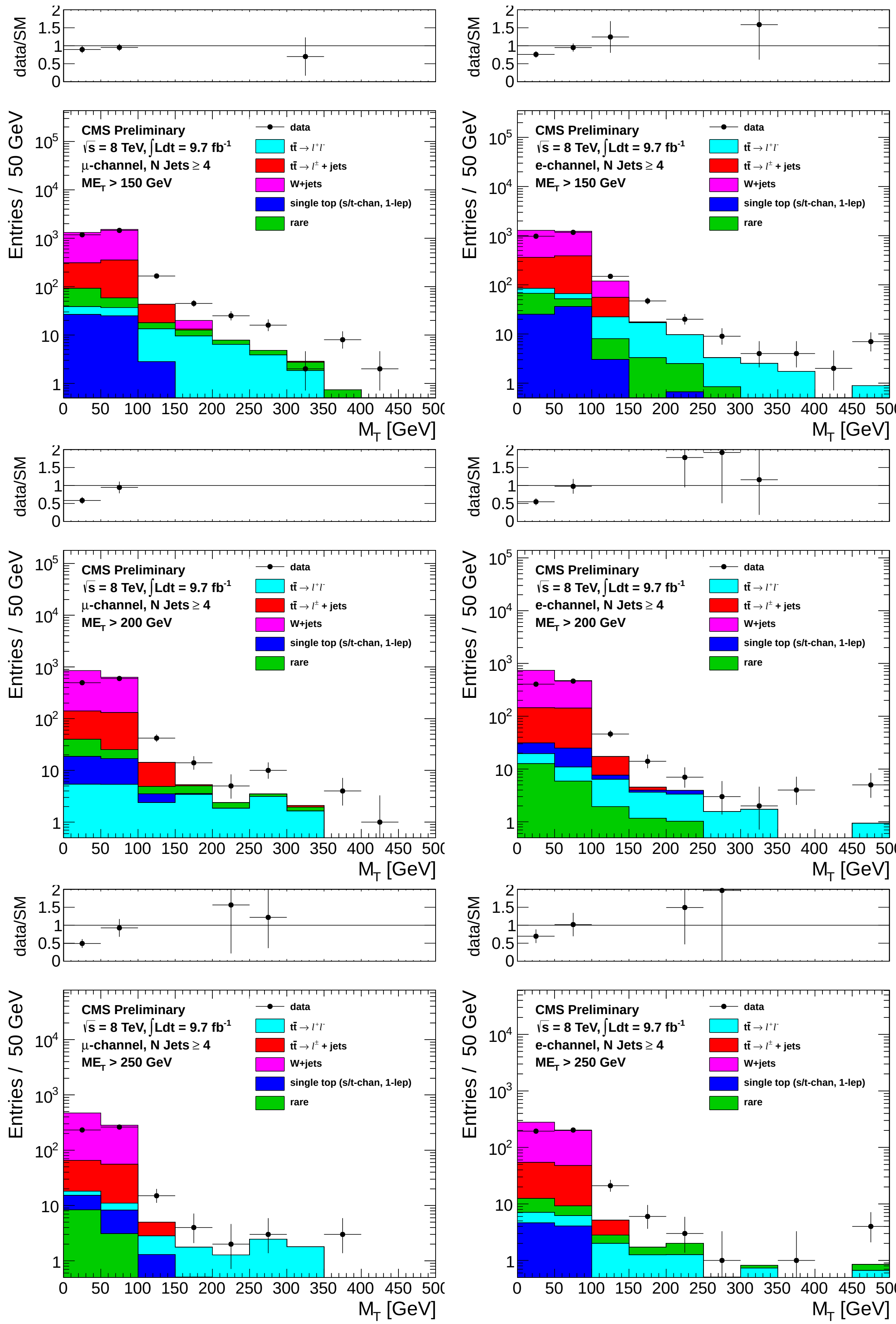

Figure 3: Comparison of the M_T distribution in data vs. MC for events with a leading muon (left) and leading electron (right) satisfying the requirements of CR1. The $E_T^{\rm miss}$ requirements used are 150 GeV (top), 200 GeV (middle) and 250 GeV (bottom).

5.2 Single Lepton Top MC Modelling Validation from CR2

IS THIS GOING TO BE DONE WITH A BVETO OR NOT. IF SO, IS IT GOING TO BE CSVL OR CSVM? NEED TO DISCUSS THIS.

The M_{T} tail for single-lepton top events ($\mathrm{t\bar{t}} \rightarrow \ell + \mathrm{jets}$ and single top) is dominated by jet resolution effects. The W cannot be far off-shell because $M_{\mathrm{W}} < M_{\mathrm{top}}$. The modeling of the M_{T} tail from jet resolution effects is studied using Z+jets data and MC samples. Z events are selection by requiring 2 good leptons (satisfying ID and isolation requirements) and requiring the $M_{\ell\ell}$ to be in the range $81-101$ GeV. The negative lepton is treated as a neutrino and so is added to the MET: $E_{\mathrm{T}}^{\mathrm{miss}} \rightarrow p_{\mathrm{T}}(\ell^{-}) + E_{\mathrm{T}}^{\mathrm{miss}}$, and the M_{T} is recalculated with the positive lepton $M_{\mathrm{T}}(\ell^{+}, E_{\mathrm{T}}^{\mathrm{miss}})$. The resulting "pseudo-$M_{\mathrm{T}}$" is dominated by jet resolution effects, since no off-shell Z production enters the sample due to the $M_{\ell\ell}$ requirement. This section describes how well the MC predicts the tail of "pseudo-M_{T}".

The underlying distributions are shown in Fig. 4 and 5. The comparison of data and MC event counts is shown in Table 11. From this table we extract the data to MC scale factors SFR_{top}^{e} and SFR_{top}^{μ}.

Sample	CR2PRESEL0	CR2PRESEL1	CR2A	CR2B	CR2C	CR2D
DY MC	35 ± 2	30 ± 2	18 ± 2	32 ± 3	12 ± 2	5 ± 1
Data - non-DY MC	65 ± 9	50 ± 8	36 ± 6	49 ± 7	25 ± 5	14 ± 4
Data/MC SF	1.88 ± 0.29	1.68 ± 0.30	1.94 ± 0.40	1.54 ± 0.29	2.12 ± 0.58	2.96 ± 1.22

Table 11: Yields in M_{T} tail comparing the MC prediction (after applying SFs) to data. CR2PRESEL refers to a sample with $E_{\mathrm{T}}^{\mathrm{miss}} > 50$ GeV and $M_{\mathrm{T}} > 150$ GeV. The uncertainties are statistical only. NEED TO ADD THE SYMBOLS DEFINED IN THE TEXT FOR THESE SCALE FACTORS. IS THIS GOING TO BE DONE SEPARATELY FOR MUONS AND ELECTRONS???

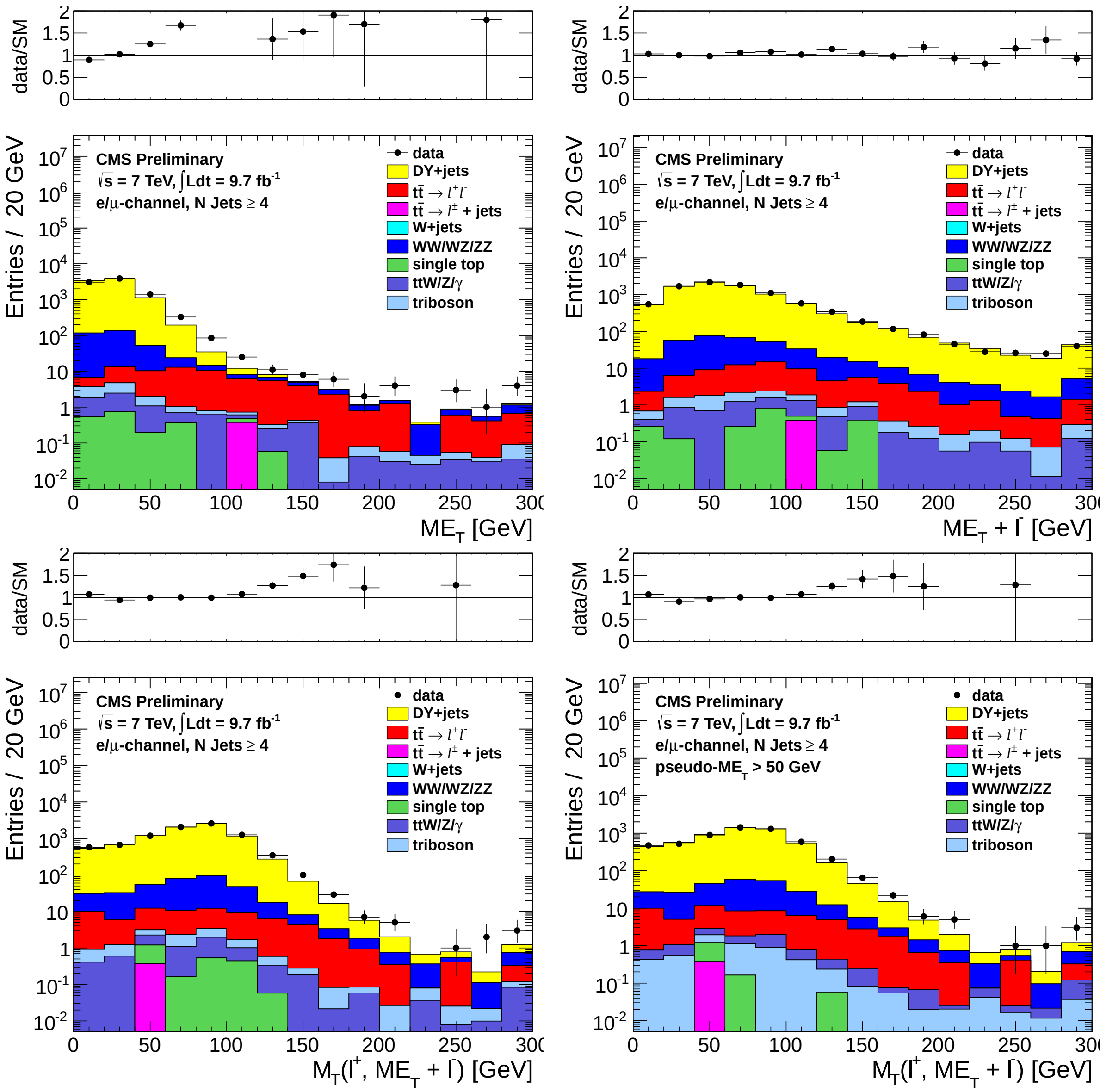

Figure 4: Comparison of the $E_{\mathrm{T}}^{\mathrm{miss}}$ (top, left), pseudo-$E_{\mathrm{T}}^{\mathrm{miss}}$ (top, right) and pseudo-M_{T} (bottom) distributions in data vs. MC for events satisfying the requirements of CR2, combining both the muon and electron channels. The pseudo-M_{T} distributions are shown before any additional requirements (bottom, left) and after requiring pseudo-$E_{\mathrm{T}}^{\mathrm{miss}}$¿50 GeV (bottom, right).

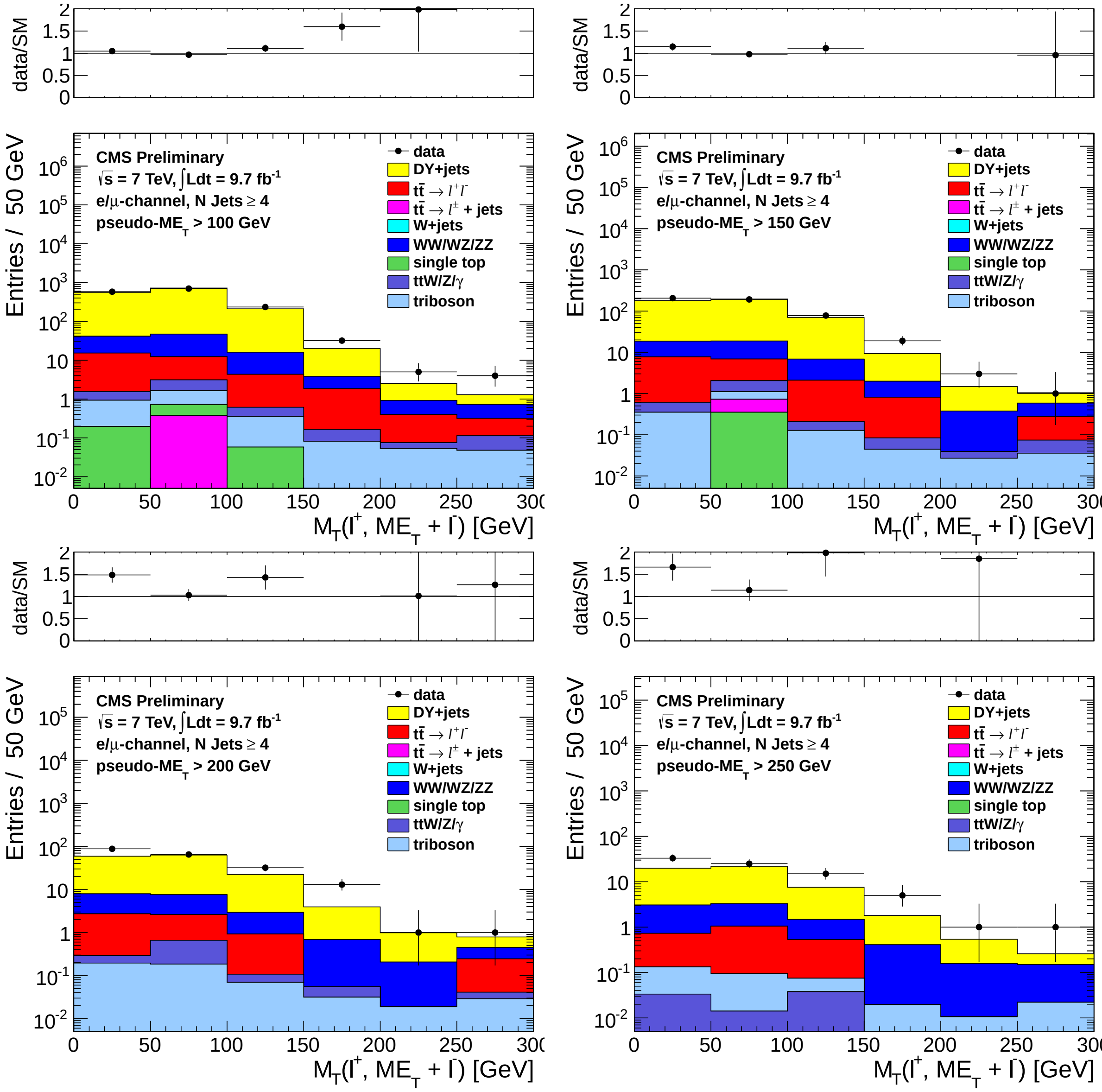

Figure 5: Comparison of the M_{T} distribution in data vs. MC for events satisfying the requirements of CR2, combining both the muon and electron channels. The pseudo-$E_{\mathrm{T}}^{\mathrm{miss}}$ requirements used are 100 GeV (top, left), 150 GeV (top, right), 200 GeV (bottom, left) and 250 GeV (bottom, right).

5.3 Dilepton studies in CR4

[DO WE NEED TO BETTER SPECIFY THE SELECTION FOR THIS REGION???]

5.3.1 Modeling of Additional Hard Jets in Top Dilepton Events

[THIS SUBSUBSECTION IS DONE...MODULO THE LATEST PLOTS AND THE LATEST NUMBERS IN THE TABLE]

Dilepton $t\bar{t}$ events have 2 jets from the top decays, so additional jets from radiation or higher order contributions are required to enter the signal sample. The modeling of addtional jets in $t\bar{t}$ events is checked in a $t\bar{t} \to \ell\ell$ control sample, selected by requiring

- exactly 2 selected electrons or muons with p_{T}> 20 GeV
- $E_{\mathrm{T}}^{\mathrm{miss}} > 100$ GeV
- ≥ 1 b-tagged jet
- Z-veto

Figure 6 shows a comparison of the jet multiplicity distribution in data and MC for this two-lepton control sample. After requiring at least 1 b-tagged jet, most of the events have 2 jets, as expected from the dominant process $t\bar{t} \to \ell\ell$. There is also a significant fraction of events with additional jets. The 3-jet sample is mainly comprised of $t\bar{t}$ events with 1 additional emission and similarly the $\geq$ 4-jet sample contains primarily $t\bar{t}+ \geq 2$ jet events.

It should be noted that in the case of $t\bar{t} \to \ell\ell$ events with a single reconstructed lepton, the other lepton may be mis-reconstructed as a jet. For example, a hadronic tau may be mis-identified as a jet (since no τ identification is used). In this case only 1 additional jet from radiation may suffice for a $t\bar{t} \to \ell\ell$ event to enter the signal sample. As a result, both the samples with $t\bar{t} + 1$ jet and $t\bar{t}+ \geq 2$ jets are relevant for estimating the top dilepton bkg in the signal region.

Table 12 shows scale factors (K_3 and K_4) used to correct the fraction of events with additional jets in MC to the observed fraction in data. These scale factors are calculated from Fig. 6 as follows:

- N_2 = data yield minus non-dilepton $t\bar{t}$ MC yield for $N_{\mathrm{jets}} \leq 2$
- N_3 = data yield minus non-dilepton $t\bar{t}$ MC yield for $N_{\mathrm{jets}} = 3$
- N_4 = data yield minus non-dilepton $t\bar{t}$ MC yield for $N_{\mathrm{jets}} \geq 4$
- M_2 = dilepton $t\bar{t}$ MC yield for $N_{\mathrm{jets}} \leq 2$
- M_3 = dilepton $t\bar{t}$ MC yield for $N_{\mathrm{jets}} = 3$
- M_4 = dilepton $t\bar{t}$ MC yield for $N_{\mathrm{jets}} \geq 4$

then

- $SF_2 = N_2/M_2$
- $SF_3 = N_3/M_3$
- $SF_4 = N_4/M_4$
- $K_3 = SF_3/SF_2$
- $K_4 = SF_4/SF_2$

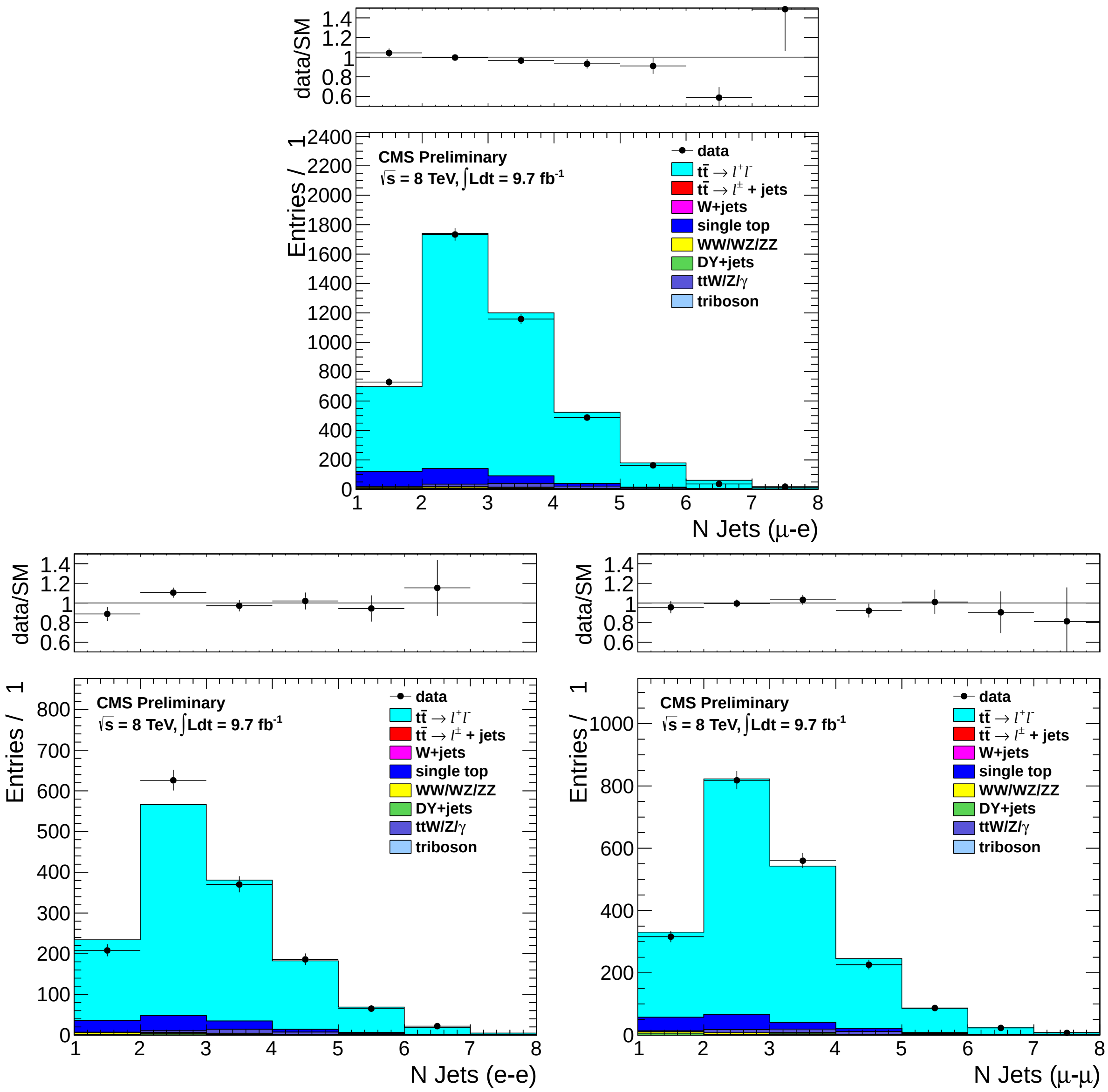

Figure 6: Comparison of the jet multiplicity distribution in data and MC for dilepton events in the e-μ (top), e-e (bottom left) and μ-μ (bottom right) channels.

This insures that $K_3 M_3/(M_2 + K_3 M_3 + K_4 M_4) = N_3/(N_2 + N_3 + N_4)$ and similarly for the ≥ 4 jet bin. The factors K_3 and K_4 are applied to the $t\bar{t} \to \ell\ell$ MC throughout the entire analysis, i.e. whenever $t\bar{t} \to \ell\ell$ MC is used to estimate or subtract a yield or distribution. In order to do so, it is first necessary to count the number of additional jets from radiation and exclude leptons mis-identified as jets. A jet is considered a mis-identified lepton if it is matched to a generator-level second lepton with sufficient energy to satisfy the jet p_T requirement ($p_T > 30$ GeV). Then $t\bar{t} \to \ell\ell$ events that need two radiation jets to enter our selection are scaled by K_4, while those that only need one radiation jet are scaled by K_3.

Jet Multiplicity Sample	Data/MC Scale Factor
N jets = 3 (sensitive to $t\bar{t} + 1$ extra jet from radiation)	$K_3 = 0.97 \pm 0.03$
N jets ≥ 4 (sensitive to $t\bar{t}+ \geq 2$ extra jets from radiation)	$K_4 = 0.91 \pm 0.04$

Table 12: Data/MC scale factors used to account for differences in the fraction of events with additional hard jets from radiation in $t\bar{t} \to \ell\ell$ events.

5.3.2 Validation of the "Physics" Modelling of the $t\bar{t} \to \ell\ell$ MC in CR4

5.3.3 sec:CR4-valid

[THE TEXT IN THIS SUBSECTION IS ESSENTIALLY COMPLETE]

As mentioned above, $t\bar{t} \to$ dileptons where one of the leptons is somehow lost constitutes the main background. The object of this test is to validate the M_T distribution of this background by looking at the M_T distribution of well identified dilepton events. We construct a transverse mass variable from the leading lepton and the $E_{\mathrm{T}}^{\mathrm{miss}}\dot{\mathrm{W}}$e distinguish between events with leading electrons and leading muons.

The $t\bar{t}$ MC is corrected using the K_3 and K_4 factors from Section 5.3.1. It is also normalized to the total data yield separately for the $E_{\mathrm{T}}^{\mathrm{miss}}$ requirements of signal regions A, B, C, and D. These normalization factors are listed in Table 13 and are close to unity.

The underlying $E_{\mathrm{T}}^{\mathrm{miss}}$ and M_T distributions are shown in Figures 7 and **??**. The data-MC agreement is quite good. Quantitatively, this is also shown in Table 14.

Sample	CR4A	CR4B	CR4C	CR4D
Muon Data/MC-SF	0.91 ± 0.04	0.94 ± 0.07	1.06 ± 0.13	1.03 ± 0.22
Electron Data/MC-SF	0.95 ± 0.04	1.00 ± 0.08	0.85 ± 0.12	0.83 ± 0.19

Table 13: Data/MC scale factors for total yields, applied to compare the shapes of the distributions. The uncertainties are statistical only.

Sample	CR4A	CR4B	CR4C	CR4D
Muon MC	199 ± 7	102 ± 6	29 ± 3	8 ± 1
Muon Data	187	108	34	9
Muon Data/MC SF	0.94 ± 0.08	1.06 ± 0.12	1.17 ± 0.23	1.09 ± 0.40
Electron MC	203 ± 8	97 ± 5	26 ± 2	8 ± 1
Electron Data	201	102	25	5
Electron Data/MC SF	0.99 ± 0.08	1.06 ± 0.12	0.97 ± 0.21	0.60 ± 0.29

Table 14: Yields in M_{T} tail comparing the MC prediction (after applying SFs) to data. The uncertainties are statistical only.

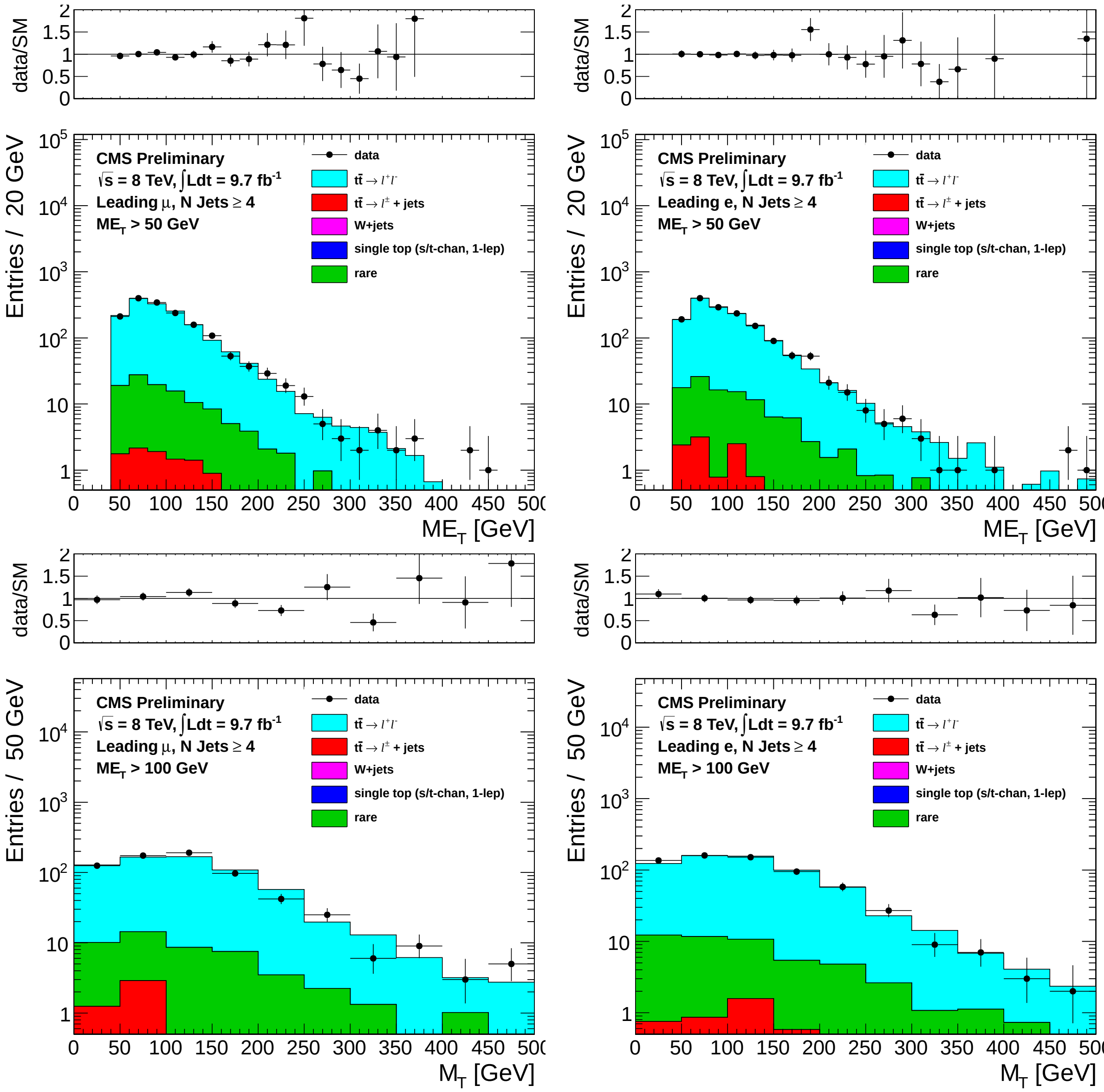

Figure 7: Comparison of the $E_{\mathrm{T}}^{\mathrm{miss}}$ (top) and M_{T} for $E_{\mathrm{T}}^{\mathrm{miss}} > 100$ (bottom) distributions in data vs. MC for events with a leading muon (left) and leading electron (right) satisfying the requirements of CR4.

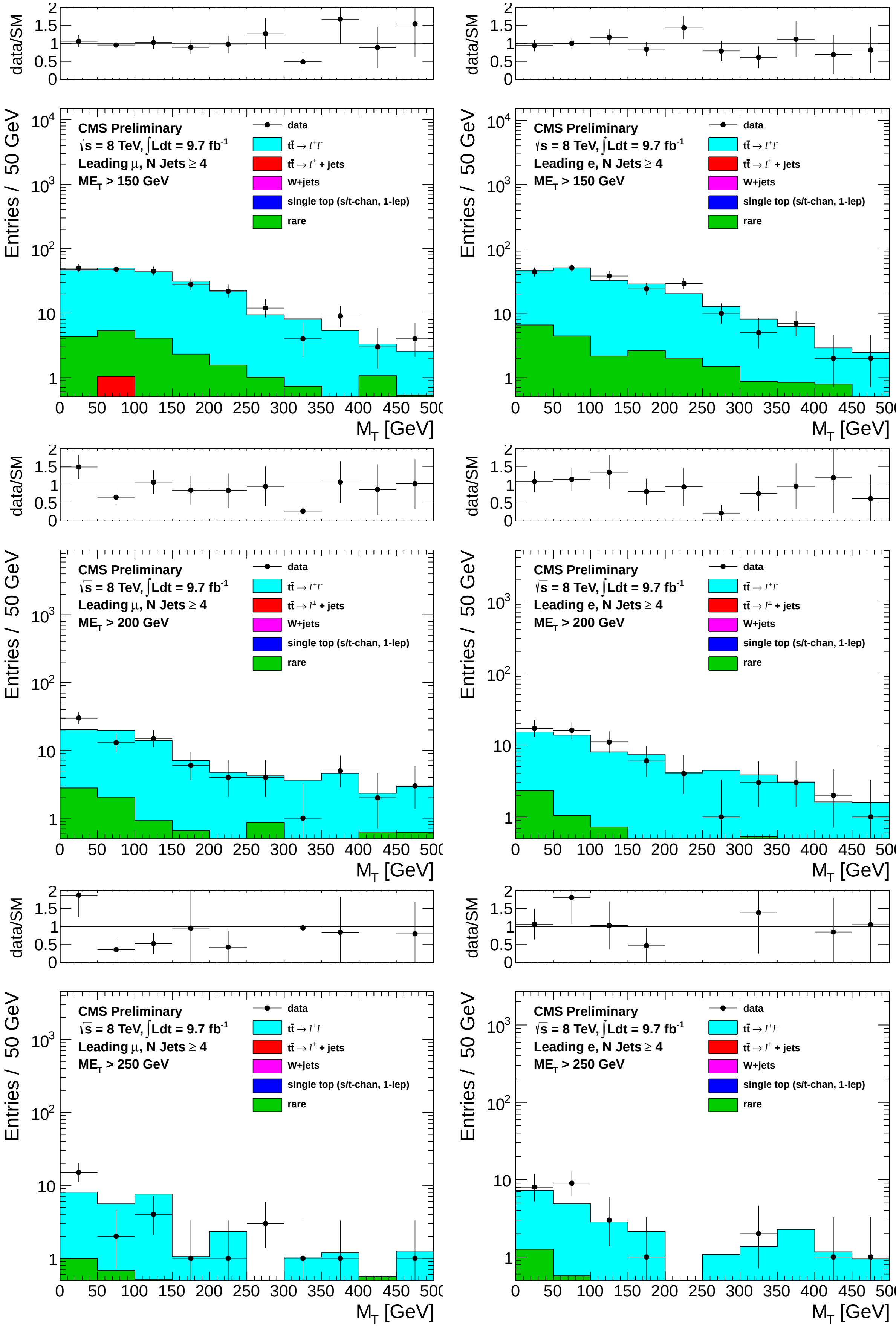

Figure 8: Comparison of the M_{T} distribution in data vs. MC for events with a leading muon (left) and leading electron (right) satisfying the requirements of CR4. The $E_{\mathrm{T}}^{\mathrm{miss}}$ requirements used are 150 GeV (top), 200 GeV (middle) and 250 GeV (bottom).

5.4 Test of control region with isolated track in CR5

[NEED TO VERIFY THAT THE DESCRIPTION OF SCALE FACTORS IS CORRECT AND ADD A LITTLE BIT OF DETAIL, AS NOTED IN THE TEXT]

This CR consists of events that pass all cuts but fail the isolated track veto cut. These events (especially in the tail of M_T) are predominantly $t\bar{t}$ dileptons. Thus the test in this control regions is similar to that performed in CR4 and described in Section **??**. There is some non-trivial complementarity because CR5 also includes events with taus and events with electrons or muons below the threshold of the CR4 selection. Also, this test is somewhat sensitive to the simulation of the track isolation requirement, since the number of dilepton events in CR5 depends on the (in)efficiency of that cut.

In CR5 there is also a significant component of $t\bar{t} \to \ell+$ jets, where one of the jets fluctuates to an isolated track. This component dominates at low M_T and is not necessarily well reproduced quantitatively by the simulation. This makes the normalization of the top MC a little bit tricky. We define a "pre-veto" sample as the sample of events that pass all cuts without any isolated track requirements. This sample is dominated by $t\bar{t} \to \ell+$ jets. We normalize the dilepton component of the top MC to that sample (NEED TO EXPLAIN EXACTLY HOW). Next we define a "post-veto" sample as the events that have an isolated track. The $t\bar{t} \to \ell+$ jets component is normalized in this sample (ALSO, NEED TO EXPLAIN HOW, EXACTLY). These normalization factors are summarized in Table 15.

The underlying $E_{\mathrm{T}}^{\mathrm{miss}}$ and M_T distributions are shown in Figures 9 and **??**. The data-MC agreement is quite good. Quantitatively, this is also shown in Table 16.

Sample	CR5A	CR5B	CR5C	CR5D
Muon pre-veto M_{T}-SF	0.98 ± 0.02	0.95 ± 0.04	0.99 ± 0.08	0.89 ± 0.15
Muon post-veto M_{T}-SF	1.28 ± 0.07	1.20 ± 0.13	1.22 ± 0.24	1.25 ± 0.43
Electron pre-veto M_{T}-SF	0.83 ± 0.02	0.75 ± 0.04	0.64 ± 0.07	0.63 ± 0.12
Electron post-veto M_{T}-SF	1.10 ± 0.08	1.02 ± 0.11	0.89 ± 0.19	1.27 ± 0.41

Table 15: M_{T} peak Data/MC scale factors. The pre-veto SFs are applied to the $\mathrm{t\bar{t}} \to \ell\ell$ sample, while the post-veto SFs are applied to the single lepton samples. The raw MC is used for backgrounds from rare processes. The uncertainties are statistical only.

Sample	CR5A	CR5B	CR5C	CR5D
Muon MC	293 ± 9	161 ± 7	51 ± 4	16 ± 2
Muon Data	315	165	62	13
Muon Data/MC SF	1.07 ± 0.07	1.03 ± 0.09	1.21 ± 0.18	0.82 ± 0.25
Electron MC	253 ± 8	126 ± 5	37 ± 3	12 ± 2
Electron Data	286	135	39	15
Electron Data/MC SF	1.13 ± 0.08	1.07 ± 0.10	1.07 ± 0.19	1.21 ± 0.35

Table 16: Yields in M_{T} tail comparing the MC prediction (after applying SFs) to data. The uncertainties are statistical only.

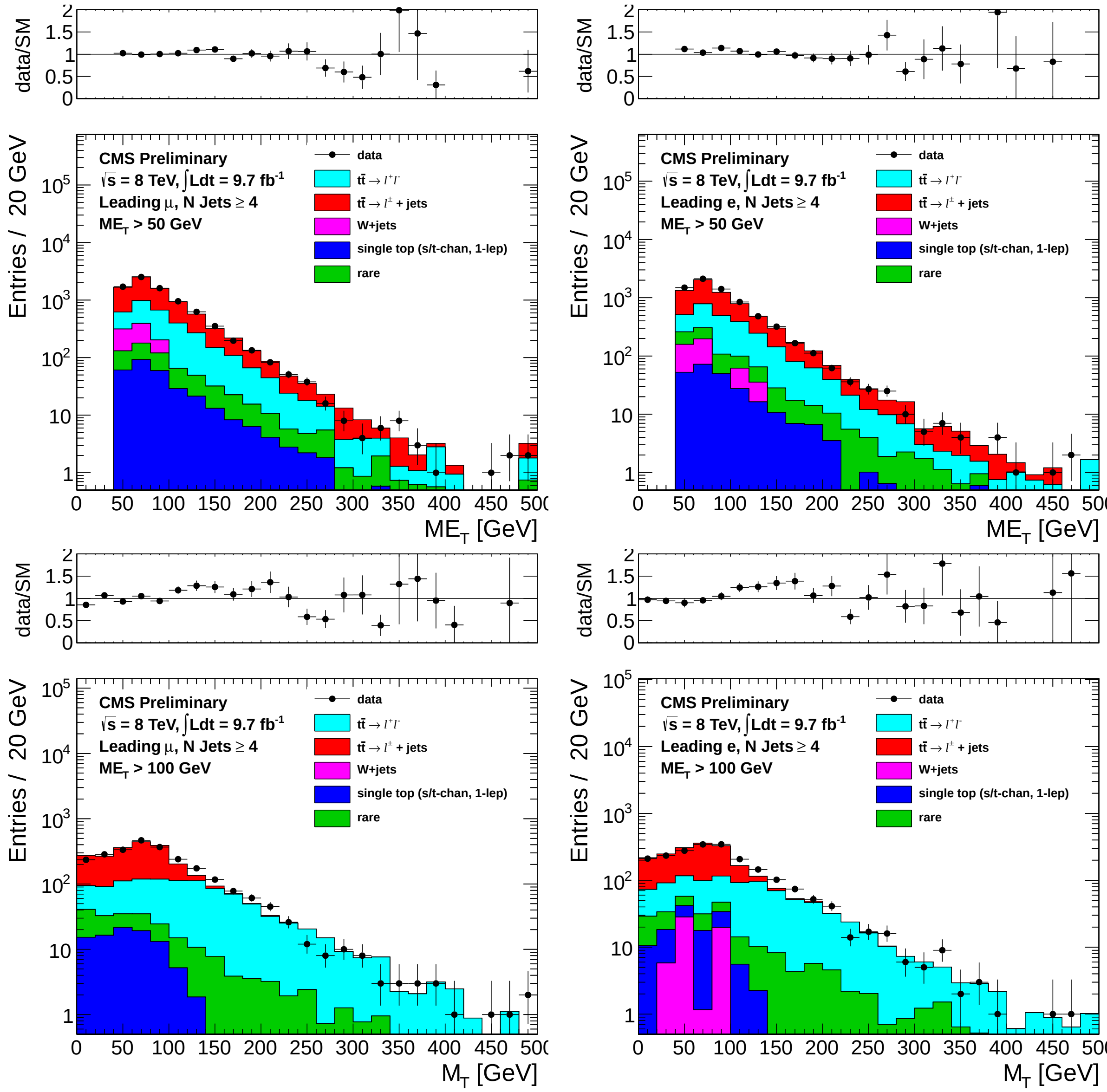

Figure 9: Comparison of the $E_{\mathrm{T}}^{\mathrm{miss}}$ (top) and M_{T} for $E_{\mathrm{T}}^{\mathrm{miss}} > 100$ (bottom) distributions in data vs. MC for events with a leading muon (left) and leading electron (right) satisfying the requirements of CR5.

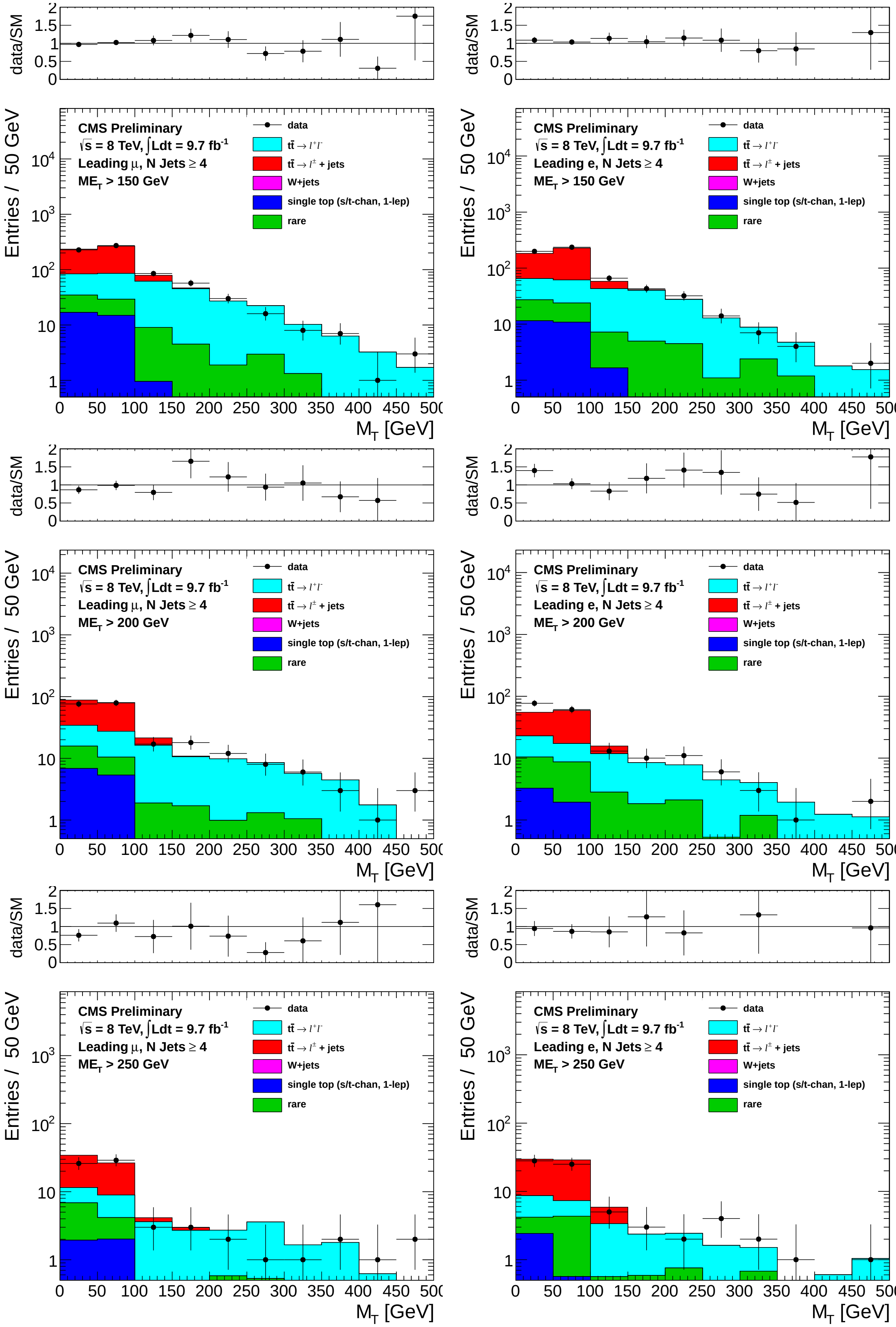

Figure 10: Comparison of the M_T distribution in data vs. MC for events with a leading muon (left) and leading electron (right) satisfying the requirements of CR5. The $E_T^{\rm miss}$ requirements used are 150 GeV (top), 200 GeV (middle) and 250 GeV (bottom).

6 Other Backgrounds

Additional background contributions from rare processes include

- $\mathrm{t\bar{t}}$ in association with a boson $\mathrm{t\bar{t}}$ + WZ/γ^*
- Z/γ^*+Jets
- diboson WW/WZ/ZZ
- triboson WWW/WWZ/WZZ/ZZZ
- dilepton single top tW.

These backgrounds are small, contributing at the $\sim 5\%$ level and their predictions are taken from MC, normalized to the corresponding cross sections. A 50% systematic uncertainty is assigned for all these backgrounds.

Backgrounds from QCD are expected to be small in the signal regions with large M_T and $E_\mathrm{T}^\mathrm{miss}$

7 Tail-to-Peak ratio for lepton + jets top and W events

[FILL IN SOME XX]

An important component of the background calculation is the ratio of the number of events with M_T in the signal region to the number of events with XX $< M_T <$ XX GeV. As discussed in Section **??**, these ratios are different for $W+$ jets and top events.

Sample	SRA	SRB	SRC	SRD	SRE
Muon TTP SL Top (R^{μ}_{top})	0.012	0.030	0.019	0.021	0.024
Muon TTP W+Jets (R^{μ}_{wjet})	0.032	0.058	0.054	0.060	0.058
Electron TTP SL Top (R^{e}_{top})	0.012	0.027	0.021	0.023	0.010
Electron TTP W+Jets (R^{e}_{wjet})	0.032	0.070	0.071	0.073	0.070

Table 17: Ratio of MC events in the M_T-tail over events in the M_T-peak for $\mathrm{t\bar{t}} \to \ell +$ jets (also used for 1-lepton single top) and W+jets. These are derived before applying the b-tagging requirement. ADD STAT. UNCERTAINTIES. VERENA....MAKE SURE TO ADD THE SYMBOLS

The MC value of these ratios are shown in Table 22. In addition the studies of CR1 and CR2 (Sections 5.1 and 5.2) lead to data/MC scale factors SFR^e_{wjets} and SFR^μ_{wjets} (Table 10) and SFR^e_{top} and SFR^μ_{top} (Table 11)

8 Background Prediction

[THIS IS WHERE THE NUMBERS TO DERIVE THE BACKGROUND ESTIMATES ARE DUMPED AND THE CALCULATION EXPLAINED]

Here we give the details of how we arrive at the background prediction in a given signal region. Here we concentrate on the method used to arrive at the central value of the background prediction. The systematic uncertainties will be discussed in Section XX.

As mentioned in Section 2, we normalize the main $t\bar{t}$ background to the M_T peak. This is actually a bit tricky because we want to minimize the effect of the isolated track veto on lepton + jets events, which may not be terribly well reproduced. Thus, we define two normaliztion region in the M_T peak (XX $< M_T <$ XX), one before and one after the application of the isolated track veto.

The dominant dilepton background is normalized to the pre-veto normalization region. A pre-veto scale factor (SF_{pre}) is defined as a common scale factors that needs to be applied to the $t\bar{t}$, single-top, and $W+$ jets MC to make the data yield in the pre-veto normalization agree with the MC prediction (the small rare MC component is held fixed). Then, the dilepton background prediction is obtained by multiplying the dilepton BG Monte Carlo by SF_{pre}.

The $t\bar{t}$ lepton + jet BG is normalized to post-veto normalization region. A post-veto scale factor (SF_{post}) is defined in (almost) the same way as the pre-veto scale factor. The difference here is that this scale factor applies only to the lepton + jets components and not the dilepton component, since that component is already rescaled by SF_{pre}. This procedure minimizes the reliance on the understanding of the isolated track veto. Then the $t\bar{t}$ lepton + jet BG is obtained by taking the number of MC-predicted $t\bar{t}$ lepton + jets in the post-veto normalization region, scaling it by SF_{post}, multiplying it by the tail-to-peak ratio R_{top} of Table 22, and finally the data-MC scale factor SFR_{top} from Table 11.

The single top background is obtained in exactly the same way as the $t\bar{t}$ lepton + jet BG, using the same tail-to-peak ratio and the same data-MC scale-factor. The W background is done in a similar way, but using a different tail-to-peak ratio (R_{wjets} of Table 22), and a different data-MC scale factor (SFR_{wjet} from Table 10).

Other (small) backgrounds are taken straight from Monte Carlo, as described in Section 6.

Sample	SRA	SRB	SRC	SRD
Muon				
$t\bar{t} \to \ell\ell$	536 ± 13	145 ± 7	42 ± 4	16 ± 2
$t\bar{t} \to \ell +$ jets & single top (1ℓ)	4953 ± 41	1408 ± 22	429 ± 12	149 ± 7
W+jets	332 ± 88	74 ± 38	42 ± 30	42 ± 30
Rare	152 ± 6	55 ± 4	19 ± 2	8 ± 1
Total	5973 ± 99	1682 ± 45	532 ± 32	214 ± 31
Data	5861	1608	526	192
Electron				
$t\bar{t} \to \ell\ell$	498 ± 12	146 ± 7	49 ± 4	17 ± 2
$t\bar{t} \to \ell +$ jets & single top (1ℓ)	4581 ± 39	1307 ± 21	407 ± 12	134 ± 7
W+jets	560 ± 111	224 ± 72	140 ± 58	46 ± 33
Rare	135 ± 6	51 ± 4	21 ± 2	9 ± 1
Total	5774 ± 119	1728 ± 76	617 ± 59	207 ± 34
Data	4822	1314	405	134
Muon+Electron Combined				
$t\bar{t} \to \ell\ell$	1034 ± 18	291 ± 9	92 ± 5	33 ± 3
$t\bar{t} \to \ell +$ jets & single top (1ℓ)	9534 ± 57	2715 ± 30	836 ± 17	283 ± 10
W+jets	891 ± 142	298 ± 82	182 ± 65	88 ± 44
Rare	287 ± 9	106 ± 5	41 ± 3	17 ± 2
Total	11747 ± 154	3410 ± 88	1149 ± 67	421 ± 45
Data	10683	2922	931	326

Table 18: Preveto MC and data yields in M_{T} peak region. The n-jets k-factors have been applied to the $t\bar{t} \to \ell\ell$. The uncertainties are statistical only.

Sample	SRA	SRB	SRC	SRD
Muon				
$t\bar{t} \to \ell\ell$	196 ± 8	53 ± 4	16 ± 2	7 ± 1
$t\bar{t} \to \ell +$ jets & single top (1ℓ)	4459 ± 39	1275 ± 21	389 ± 11	136 ± 7
W+jets	332 ± 88	74 ± 38	42 ± 30	42 ± 30
Rare	109 ± 5	38 ± 3	13 ± 2	5 ± 1
Total	5096 ± 97	1440 ± 44	460 ± 32	189 ± 31
Data	4842	1343	447	164
Electron				
$t\bar{t} \to \ell\ell$	200 ± 8	62 ± 4	24 ± 3	8 ± 2
$t\bar{t} \to \ell +$ jets & single top (1ℓ)	4117 ± 37	1163 ± 20	367 ± 11	121 ± 6
W+jets	541 ± 110	224 ± 72	140 ± 58	46 ± 33
Rare	96 ± 5	37 ± 3	15 ± 2	7 ± 1
Total	4954 ± 117	1486 ± 75	546 ± 59	182 ± 33
Data	3980	1086	345	107
Muon+Electron Combined				
$t\bar{t} \to \ell\ell$	396 ± 11	115 ± 6	40 ± 4	14 ± 2
$t\bar{t} \to \ell +$ jets & single top (1ℓ)	8576 ± 54	2438 ± 29	756 ± 16	256 ± 9
W+jets	872 ± 141	298 ± 82	182 ± 65	88 ± 44
Rare	205 ± 7	75 ± 4	28 ± 3	12 ± 2
Total	10050 ± 152	2926 ± 87	1006 ± 67	371 ± 45
Data	8822	2429	792	271

Table 19: MC and data yields in M_{T} peak region after full selection. The n-jets k-factors have been applied to the $t\bar{t} \to \ell\ell$. The uncertainties are statistical only.

Sample	SRA	SRB	SRC	SRD
Muon				
$t\bar{t} \to \ell\ell$	337 ± 10	198 ± 8	71 ± 5	23 ± 3
$t\bar{t} \to \ell +$ jets & single top (1ℓ)	52 ± 4	39 ± 3	9 ± 2	3 ± 1
W+jets	29 ± 26	29 ± 26	29 ± 26	26 ± 26
Rare	33 ± 3	22 ± 2	9 ± 1	4 ± 1
Total	452 ± 28	288 ± 27	118 ± 26	56 ± 26
Electron				
$t\bar{t} \to \ell\ell$	312 ± 10	180 ± 7	53 ± 4	17 ± 2
$t\bar{t} \to \ell +$ jets & single top (1ℓ)	59 ± 4	32 ± 3	6 ± 1	1 ± 1
W+jets	29 ± 24	28 ± 24	0 ± 0	0 ± 0
Rare	30 ± 2	18 ± 2	8 ± 1	3 ± 1
Total	429 ± 27	258 ± 25	67 ± 4	21 ± 2
Muon+Electron Combined				
$t\bar{t} \to \ell\ell$	649 ± 14	378 ± 11	125 ± 6	40 ± 4
$t\bar{t} \to \ell +$ jets & single top (1ℓ)	111 ± 6	71 ± 5	15 ± 2	4 ± 1
W+jets	58 ± 35	57 ± 35	29 ± 26	26 ± 26
Rare	63 ± 3	40 ± 3	17 ± 2	7 ± 1
Total	881 ± 39	545 ± 37	185 ± 27	77 ± 26

Table 20: MC yields in M_T tail region after full selection. The n-jets k-factors have been applied to the $t\bar{t} \to \ell\ell$. The uncertainties are statistical only. Note these values are only used for the rare backgrounds prediction.

Sample	SRA	SRB	SRC	SRD
Muon pre-veto M_T-SF	0.98 ± 0.02	0.96 ± 0.04	0.99 ± 0.08	0.90 ± 0.15
Muon post-veto M_T-SF	0.95 ± 0.01	0.93 ± 0.03	0.97 ± 0.05	0.86 ± 0.07
Muon veto M_T-SF	0.97 ± 0.01	0.97 ± 0.01	0.98 ± 0.02	0.96 ± 0.04
Electron pre-veto M_T-SF	0.83 ± 0.02	0.75 ± 0.04	0.64 ± 0.07	0.63 ± 0.12
Electron post-veto M_T-SF	0.80 ± 0.01	0.72 ± 0.02	0.62 ± 0.04	0.57 ± 0.06
Electron veto M_T-SF	0.96 ± 0.01	0.96 ± 0.02	0.96 ± 0.03	0.90 ± 0.05

Table 21: M_T peak Data/MC scale factors. The pre-veto SFs are applied to the $t\bar{t} \to \ell\ell$ sample, while the post-veto SFs are applied to the single lepton samples. The veto SF is shown for comparison across channels. The raw MC is used for backgrounds from rare processes. The uncertainties are statistical only.

Sample	SRA	SRB	SRC	SRD
Muon TTP SL Top	0.011	0.025	0.020	0.015
Muon TTP W+Jets	0.039	0.021	0.038	0.085
Electron TTP SL Top	0.012	0.025	0.018	0.011
Electron TTP W+Jets	0.035	0.020	0.000	0.000

Table 22: Ratio of MC events in the M_T-tail over events in the M_T-peak for $t\bar{t} \to \ell +$ jets (also used for 1-lepton single top) and W+jets. These are derived before applying the b-tagging requirement. ADD STAT. UNCERTAINTIES

9 Systematic Uncertainties

[ADD INTRODUCTORY BLURB ON UNCERTAINTIES
ADD COMPARISONS OF ALL THE ALTERNATIVE SAMPLES FOR ALL THE SIGNAL REGIONS
LIST ALL THE UNCERTAINTIES INCLUDED AND THEIR VALUES]

9.1 Uncertainty on the $t\bar{t} \to \ell\ell$ Acceptance

The $t\bar{t}$ background prediction is obtained from MC, with corrections derived from control samples in data. The uncertainty associated with the theoretical modeling of the $t\bar{t}$ production and decay is estimated by comparing the background predictions obtained using alternative MC samples. It should be noted that the full analysis is performed with the alternative samples under consideration, including the derivation of the various data-to-MC scale factors. The variations considered are

- Top mass: The alternative values for the top mass differ from the central value by 5 GeV: $m_{\mathrm{top}} = 178.5$ GeV and $m_{\mathrm{top}} = 166.5$ GeV.
- Jet-parton matching scale: This corresponds to variations in the scale at which the Matrix Element partons from Madgraph are matched to Parton Shower partons from Pythia. The nominal value is $x_q > 20$ GeV. The alternative values used are $x_q > 10$ GeV and $x_q > 40$ GeV.
- Renormalization and factorization scale: The alternative samples correspond to variations in the scale $\times 2$ and $\times 0.5$. The nominal value for the scale used is $Q^2 = m_{\mathrm{top}}^2 + \sum_{\mathrm{jets}} p_{\mathrm{T}}^2$.
- Alternative generators: Samples produced with different generators include MC@NLO and Powheg (NLO generators) and Pythia (LO). It may also be noted that MC@NLO uses Herwig6 for the hadronisation, while POWHEG uses Pythia6.
- Modeling of taus: The alternative sample does not include Tauola and is otherwise identical to the Powheg sample. [DONE AT 7TEV AND FOUND TO BE NEGLIGIBLE]
- The PDF uncertainty is estimated following the PDF4LHC recommendations[CITE]. The events are reweighted using alternative PDF sets for CT10 and MSTW2008 and the uncertainties for each are derived using the alternative eigenvector variations and the "master equation". In addition, the NNPDF2.1 set with 100 replicas. The central value is determined from the mean and the uncertainty is derived from the 1σ range. The overall uncertainty is derived from the envelope of the alternative predictions and their uncertainties. [DONE AT 7 TEV AND FOUND TO BE NEGLIGIBLE]

9.2 Isolated Track Veto: Tag and Probe Studies

[EVERYTHING IS 7TEV HERE, UPDATE WITH NEW RESULTS
ADD TABLE WITH FRACTION OF EVENTS THAT HAVE A TRUE ISOLATED TRACK]

In this section we compare the performance of the isolated track veto in data and MC using tag-and-probe studies with samples of Z$\to$ee and Z$\to \mu\mu$. The purpose of these studies is to demonstrate that the efficiency to satisfy the isolated track veto requirements is well-reproduced in the MC, since if this were not the case we would need to apply a data-to-MC scale factor in order to correctly predict the $t\bar{t} \to \ell\ell$ background. This study addresses possible data vs. MC discrepancies for the **efficiency** to identify (and reject) events with a second **genuine** lepton (e, μ, or $\tau \to$1-prong). It does not address possible data vs. MC discrepancies in the fake rate for rejecting events without a second genuine lepton; this is handled separately in the top normalization procedure by scaling the $t\bar{t} \to \ell+$jets contribution to match the data in the M_{T} peak after applying the isolated track veto. Furthermore, we test the data and MC isolated track veto efficiencies for electrons and muons since we are using a Z tag-and-probe technique, but we do not directly test the performance for hadronic tracks from τ decays. The performance for hadronic τ decay products may differ from that of electrons and muons for two reasons. First, the τ may decay to a hadronic track plus one or two π^0's, which may decay to $\gamma\gamma$ followed by a photon conversion. As shown in Figure 14, the isolation distribution for charged tracks from τ decays that are not produced in association with π^0s are consistent with that from es and μs. Since events from single prong τ decays produced in

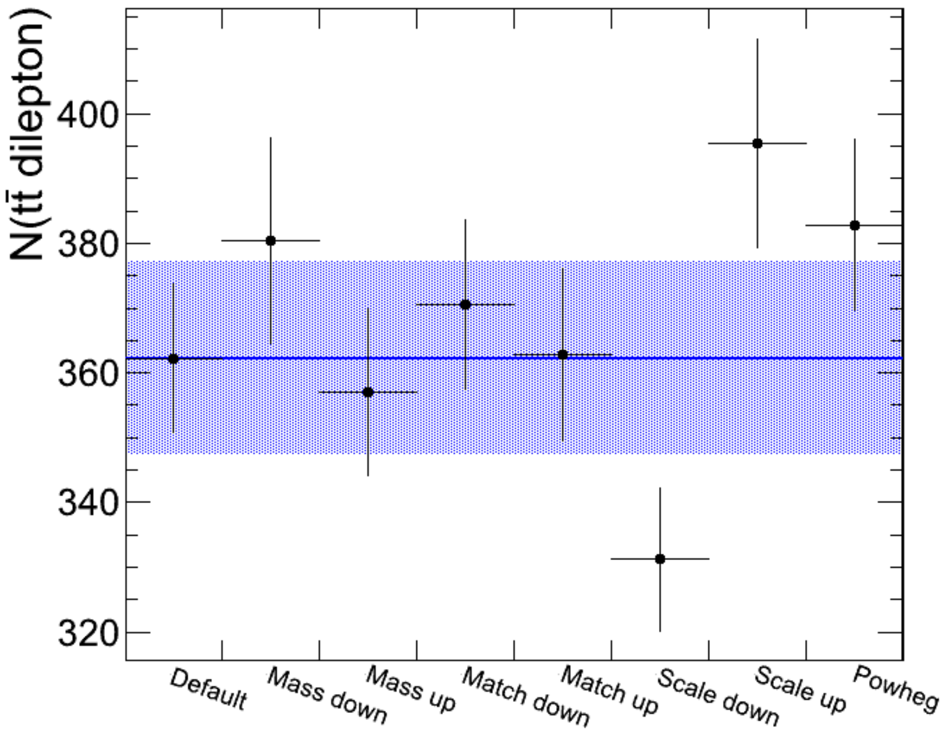

Figure 11: Comparison of the $t\bar{t} \to \ell\ell$ central prediction with those using alternative MC samples. The blue band corresponds to the total statistical error for all data and MC samples. The alternative sample predictions are indicated by the datapoints. The uncertainties on the alternative predictions correspond to the uncorrelated statistical uncertainty from the size of the alternative sample only.

association with π^0s comprise a small fraction of the total sample, and since the kinematics of τ, π^0 and $\gamma \to e^+e^-$ decays are well-understood, we currently demonstrate that the isolation is well-reproduced for electrons and muons only. Second, hadronic tracks may undergo nuclear interactions and hence their tracks may not be reconstructed. As discussed above, independent studies show that the MC reproduces the hadronic tracking efficiency within 4%, leading to a total background uncertainty of less than 0.5% (after taking into account the fraction of the total background due to hadronic τ decays with $p_T > 10$ GeV tracks), and we hence regard this effect as neglgigible.

The tag-and-probe studies are performed in the full 2011 data sample, and compared with the DYJets madgraph sample. All events must contain a tag-probe pair (details below) with opposite-sign and satisfying the Z mass requirement 76–106 GeV. We compare the distributions of absolute track isolation for probe electrons/muons in data vs. MC. The contributions to this isolation sum are from ambient energy in the event from underlying event, pile-up and jet activitiy, and hence do not depend on the p_T of the probe lepton. We therefore restrict the probe p_T to be > 30 GeV in order to suppress fake backgrounds with steeply-falling p_T spectra. To suppress non-Z backgrounds (in particular $t\bar{t}$) we require $E_T^{miss} < 30$ GeV and 0 b-tagged events. The specific criteria for tags and probes for electrons and muons are:

- Electrons
  - Tag criteria
    - Electron passes full analysis ID/iso selection
    - $p_T > 30$ GeV, $|\eta| < 2.5$
    - Matched to 1 of the 2 electron tag-and-probe triggers
      - `HLT_Ele17_CaloIdVT_CaloIsoVT_TrkIdT_TrkIsoVT_SC8_Mass30_v*`

        · HLT_Ele17_CaloIdVT_CaloIsoVT_TrkIdT_TrkIsoVT_Ele8_Mass30_v*

  - Probe criteria
    * Electron passes full analysis ID selection
    * $p_{\mathrm{T}} > 30$ GeV

- Muons
  - Tag criteria
    * Muon passes full analysis ID/iso selection
    * $p_{\mathrm{T}} > 30$ GeV, $|\eta| < 2.1$
    * Matched to 1 of the 2 electron tag-and-probe triggers
      · HLT_IsoMu30_v*
      · HLT_IsoMu30_eta2p1_v*
  - Probe criteria
    * Muon passes full analysis ID selection
    * $p_{\mathrm{T}} > 30$ GeV

The absolute track isolation distributions for passing probes are displayed in Fig. 12. In general we observe good agreement between data and MC. To be more quantitative, we compare the data vs. MC efficiencies to satisfy absolute track isolation requirements varying from > 1 GeV to > 5 GeV, as summarized in Table 23. In the $\geq$0 and $\geq$1 jet bins where the efficiencies can be tested with statistical precision, the data and MC efficiencies agree within 7%, and we apply this as a systematic uncertainty on the isolated track veto efficiency. For the higher jet multiplicity bins the statistical precision decreases, but we do not observe any evidence for a data vs. MC discrepancy in the isolated track veto efficiency.

Figure 12: Comparison of the absolute track isolation in data vs. MC for electrons (left) and muons (right) for events with the N_{jets} requirement varied from $N_{\mathrm{jets}} \geq 0$ to $N_{\mathrm{jets}} \geq 4$.

Table 23: Comparison of the data vs. MC efficiencies to satisfy the indicated requirements on the absolute track isolation, and the ratio of these two efficiencies. Results are indicated separately for electrons and muons and for various jet multiplicity requirements.

e + ≥0 jets	> 1 GeV	> 2 GeV	> 3 GeV	> 4 GeV	> 5 GeV
data	0.088 ± 0.0003	0.030 ± 0.0002	0.013 ± 0.0001	0.007 ± 0.0001	0.005 ± 0.0001
mc	0.087 ± 0.0001	0.030 ± 0.0001	0.014 ± 0.0001	0.008 ± 0.0000	0.005 ± 0.0000
data/mc	1.01 ± 0.00	0.99 ± 0.01	0.97 ± 0.01	0.95 ± 0.01	0.93 ± 0.01
μ + ≥0 jets	> 1 GeV	> 2 GeV	> 3 GeV	> 4 GeV	> 5 GeV
data	0.087 ± 0.0002	0.031 ± 0.0001	0.015 ± 0.0001	0.008 ± 0.0001	0.005 ± 0.0001
mc	0.085 ± 0.0001	0.030 ± 0.0001	0.014 ± 0.0000	0.008 ± 0.0000	0.005 ± 0.0000
data/mc	1.02 ± 0.00	1.06 ± 0.00	1.06 ± 0.01	1.03 ± 0.01	1.02 ± 0.01
e + ≥1 jets	> 1 GeV	> 2 GeV	> 3 GeV	> 4 GeV	> 5 GeV
data	0.099 ± 0.0008	0.038 ± 0.0005	0.019 ± 0.0004	0.011 ± 0.0003	0.008 ± 0.0002
mc	0.100 ± 0.0004	0.038 ± 0.0003	0.019 ± 0.0002	0.012 ± 0.0002	0.008 ± 0.0001
data/mc	0.99 ± 0.01	1.00 ± 0.02	0.99 ± 0.02	0.98 ± 0.03	0.97 ± 0.03
μ + ≥1 jets	> 1 GeV	> 2 GeV	> 3 GeV	> 4 GeV	> 5 GeV
data	0.100 ± 0.0006	0.041 ± 0.0004	0.022 ± 0.0003	0.014 ± 0.0002	0.010 ± 0.0002
mc	0.099 ± 0.0004	0.039 ± 0.0002	0.020 ± 0.0002	0.013 ± 0.0001	0.009 ± 0.0001
data/mc	1.01 ± 0.01	1.05 ± 0.01	1.05 ± 0.02	1.06 ± 0.02	1.06 ± 0.03
e + ≥2 jets	> 1 GeV	> 2 GeV	> 3 GeV	> 4 GeV	> 5 GeV
data	0.105 ± 0.0020	0.042 ± 0.0013	0.021 ± 0.0009	0.013 ± 0.0007	0.009 ± 0.0006
mc	0.109 ± 0.0011	0.043 ± 0.0007	0.021 ± 0.0005	0.013 ± 0.0004	0.009 ± 0.0003
data/mc	0.96 ± 0.02	0.97 ± 0.03	1.00 ± 0.05	1.01 ± 0.06	0.97 ± 0.08
μ + ≥2 jets	> 1 GeV	> 2 GeV	> 3 GeV	> 4 GeV	> 5 GeV
data	0.106 ± 0.0016	0.045 ± 0.0011	0.025 ± 0.0008	0.016 ± 0.0007	0.012 ± 0.0006
mc	0.108 ± 0.0009	0.044 ± 0.0006	0.024 ± 0.0004	0.016 ± 0.0004	0.011 ± 0.0003
data/mc	0.98 ± 0.02	1.04 ± 0.03	1.04 ± 0.04	1.04 ± 0.05	1.06 ± 0.06
e + ≥3 jets	> 1 GeV	> 2 GeV	> 3 GeV	> 4 GeV	> 5 GeV
data	0.117 ± 0.0055	0.051 ± 0.0038	0.029 ± 0.0029	0.018 ± 0.0023	0.012 ± 0.0019
mc	0.120 ± 0.0031	0.052 ± 0.0021	0.027 ± 0.0015	0.018 ± 0.0012	0.013 ± 0.0011
data/mc	0.97 ± 0.05	0.99 ± 0.08	1.10 ± 0.13	1.03 ± 0.15	0.91 ± 0.16
μ + ≥3 jets	> 1 GeV	> 2 GeV	> 3 GeV	> 4 GeV	> 5 GeV
data	0.111 ± 0.0044	0.050 ± 0.0030	0.029 ± 0.0024	0.019 ± 0.0019	0.014 ± 0.0017
mc	0.115 ± 0.0025	0.051 ± 0.0017	0.030 ± 0.0013	0.020 ± 0.0011	0.015 ± 0.0009
data/mc	0.97 ± 0.04	0.97 ± 0.07	0.95 ± 0.09	0.97 ± 0.11	0.99 ± 0.13
e + ≥4 jets	> 1 GeV	> 2 GeV	> 3 GeV	> 4 GeV	> 5 GeV
data	0.113 ± 0.0148	0.048 ± 0.0100	0.033 ± 0.0083	0.020 ± 0.0065	0.017 ± 0.0062
mc	0.146 ± 0.0092	0.064 ± 0.0064	0.034 ± 0.0048	0.024 ± 0.0040	0.021 ± 0.0037
data/mc	0.78 ± 0.11	0.74 ± 0.17	0.96 ± 0.28	0.82 ± 0.30	0.85 ± 0.34
μ + ≥4 jets	> 1 GeV	> 2 GeV	> 3 GeV	> 4 GeV	> 5 GeV
data	0.130 ± 0.0128	0.052 ± 0.0085	0.028 ± 0.0063	0.019 ± 0.0052	0.019 ± 0.0052
mc	0.105 ± 0.0064	0.045 ± 0.0043	0.027 ± 0.0034	0.019 ± 0.0028	0.014 ± 0.0024
data/mc	1.23 ± 0.14	1.18 ± 0.22	1.03 ± 0.27	1.01 ± 0.32	1.37 ± 0.45

fix me: What you have written in the next paragraph does not explain how ϵ_{fake} is measured. Why not measure ϵ_{fake} in the b-veto region?

10 Results

11 Conclusion

References

[1] https://hypernews.cern.ch/HyperNews/CMS/get/SUS-12-007/32.html

[2] arXiv:1204.3774v1 [hep-ex]

[3] D. Barge, CMS AN-2011/464

[4] CMS Collaboration, *Search for supersymmetry in events with a Z boson, jets and momentum imbalance* PAS SUS-11-021

[5] CMS Collaboration, *"Measurement of the b-tagging efficiency using* $t\bar{t}$ *events"*, PAS BTV-11-003, in preparation.

[6] CMS Collaboration, *Measurement of Tracking Efciency*, PAS TRK-10-00

[7] CMS Collaboration, *Performance of b-jet identification in CMS*, PAS BTV-11-001

[8] M. Narain for BTV POG, https://indico.cern.ch/getFile.py/access?contribId=0&resId=1&materialId=slides&confId=163892

[9] arXiv:1103.1348v1, D. Barge *at al.*, CMS AN-CMS2011/269.

[10] V. Pavlunin, Phys. Rev. **D81**, 035005 (2010).

[11] V. Pavlunin, CMS AN-2009/125

[12] A reference to the top paper, once it is submitted. Also D. Barge *et al.*, AN-CMS2010/258.

[13] Changes to the selection for the 38x CMSSW release are given in https://twiki.cern.ch/twiki/bin/viewauth/CMS/TopDileptonRefAnalysis2010Pass5.

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

[15] https://twiki.cern.ch/twiki/bin/viewauth/CMS/EgammaWorkingPointsv3

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

[17] B. Mangano *et al.*, AN-CMS2010/283.

[18] https://twiki.cern.ch/twiki/bin/viewauth/CMS/CrossSections_3XSeries, https://twiki.cern.ch/twiki/bin/view/CMS/ProductionSpring2011

[19] CMS Collaboration, "Measurement of CMS luminosity", *CMS-PAS* **EWK-10-004** (2010).

[20] D. Barge *at al.*, AN-CMS2009/130.

[21] W. Andrews *et al.*, AN-CMS2009/023.

[22] D. Barge *at al.*, AN-CMS2010/257.

[23] L. Bauerdick *at al.*, AN-CMS2011/155.

[24] CMS-PAS-JME-10-010.

[25] arXiv:1103.6083v1, J. T. Ruderman, D. Shih

[26] H. Haber, G. Kane, Phys. Reports 117, Nos. 2-4 (1985) 75-263.

[27] http://cmssw.cvs.cern.ch/cgi-bin/cmssw.cgi/UserCode/SusyAnalysis/SLHAFILES/TChiwz/

[28] http://cmssw.cvs.cern.ch/cgi-bin/cmssw.cgi/UserCode/SusyAnalysis/SLHAFILES/TChizz/

A Performance of the Isolation Requirement

The last requirement used in the analysis is an isolated track veto. This selection criteria rejects events containing a track of $p_{\mathrm{T}} > 10$ GeV with relative track isolation $\sum p_{\mathrm{T}}/p_{\mathrm{T}}(trk)$ in a cone of size $R = 0.3 <$ 0.1. It may be noted that only tracks consistent with the vertex with highest $\sum p_{\mathrm{T}}^2$ are considered in order to reduce the impact of spurious tracks, for example from pileup interactions. This requirement has very good performance. Figure 13 shows the efficiency for rejecting dilepton events compared to the efficiency for selecting single lepton events for various cone sizes and cut values. The chosen working point provides a signal efficiency of $\epsilon(sig) = 92\%$ for a background rejection of $\epsilon(bkg) = 53\%$ in MC. With "signal" ("background") we are referring to $\mathrm{t\bar{t}} \to \ell + \mathrm{jets}$ ($\mathrm{t\bar{t}} \to \ell\ell$).

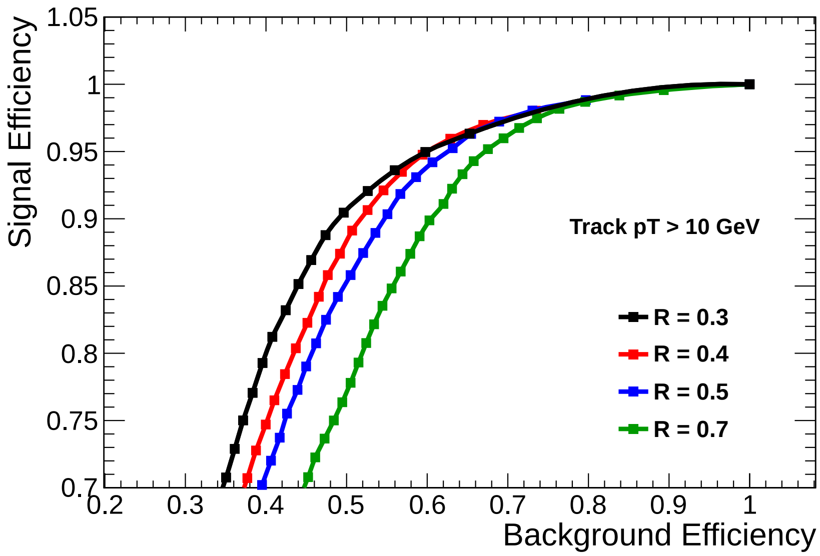

Figure 13: Comparison of the performance in terms of signal (single lepton events) efficiency and background (dilepton events) rejection for various cone sizes and cut values. The current isolation requirement uses a cone of size $\Delta R = 0.3$ and a cut value of 0.1, corresponding to $\epsilon(sig) = 92\%$ for $\epsilon(bkg) = 53\%$. ADD ARROW OR LINE TO INDICATE WORKING POINT.

It should be emphasized that the isolated track veto has a different impact on the samples with a single lepton (mainly $\mathrm{t\bar{t}} \to \ell + \mathrm{jets}$ and W+jets) and that with two leptons (mainly $\mathrm{t\bar{t}} \to \ell\ell$). For the dilepton background, the veto rejects events which have a genuine second lepton. Thus the performance may be understood as an efficiency $\epsilon_{iso\ trk}$ to identify the isolated track. In the case of the single lepton background, the veto rejects events which do not have a genuine second lepton, but rather which contain a "fake" isolated track. The isolated track veto thus effectively scales the single lepton sample by (1-ϵ_{fake}), where ϵ_{fake} is the probability to identify an isolated track with p_{T}> 10 GeVin events which contain no genuine second lepton. It is thus necessary to study the isolated track efficiency $\epsilon(trk)$ and ϵ_{fake} in order to fully characterize the veto performance.

The veto efficiency for dilepton events is calculated using the tag and probe method in Z events. A good lepton satisfying the full ID and isolation criteria and matched to a trigger object serves as the tag. The probe is defined as a track with $p_{\mathrm{T}} > 10$ GeV that has opposite charge to the tag and has an invariant mass with the probe consistent with the Z mass.

Fix me: fkw does not understand why you refer to p_{T}> 10 GeVhere, given that in the very next paragraph you state that this is measured via the absolute track isolation, implying, but not explicitly stating, that a much higher p_{T} threshold is used to get a clean Z signal. ???

The variable used to study the performance of the veto is the absolute track isolation, since it removes the dependence of the isolation variable on the p_{T} of the object under consideration. This is particularly useful because the underlying p_{T} distribution is different for second leptons in $\mathrm{t\bar{t}} \to \ell\ell$ events compared to Z events, particularly due to the presence of τs that have softer decay products. As shown in Figure 14, the absolute isolation is consistent between Z + 4 jet events and $t\bar{t} \to \ell\ell$ events, including leptons from W and τ decays. This supports the notion that the isolation, defined as the energy surrounding the object under consideration, depends only on the environment of the object and not on the object itself. The isolation is thus sensitive to the ambient pileup and jet activity in the event, which is uncorrelated with the lepton p_{T}. It is thus justified to use tag and probe in Z + 4 jet events, where the jet activity is similar to $t\bar{t} \to \ell\ell$ events in our $N_{\mathrm{jets}} > 4$ signal region, in order to estimate the performance of the isolation requirement for the various leptonic categories of $t\bar{t} \to \ell\ell$ events.

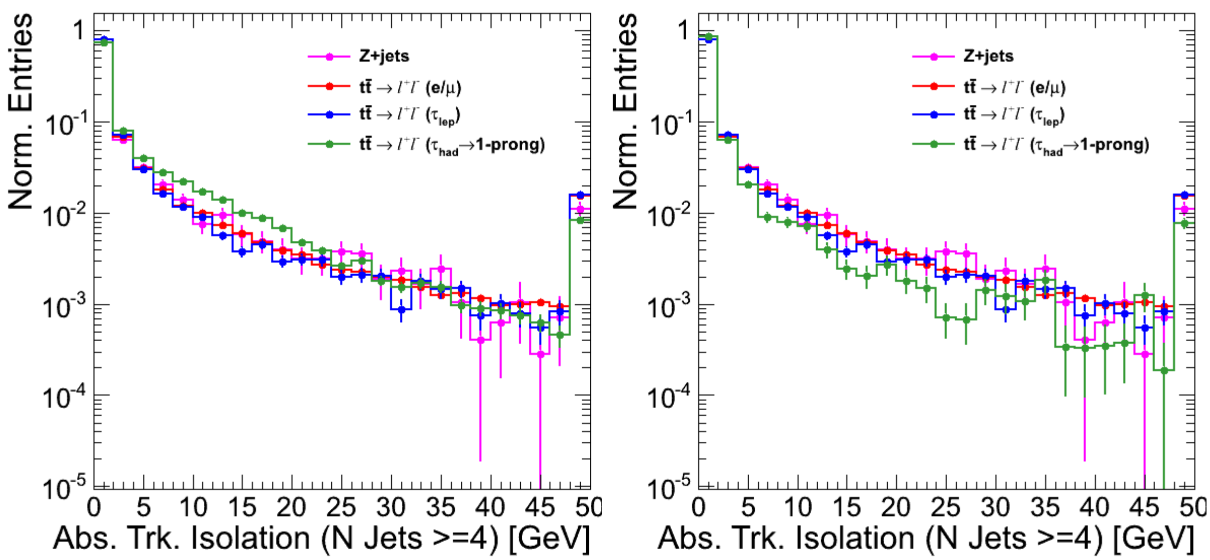

Figure 14: Comparison of absolute track isolation for track probes in Z + 4 jet and $t\bar{t} \to \ell\ell$ events for different lepton types. The isolation variables agree across samples, except for single prong τs, that tend to be slightly less isolated (left). The agreement across isolation distributions is recovered after removing single prong τ events produced in association with π^0s from the sample (right).

B Glossary of abbreviations

$R^e_{wjet}, R^\mu_{wjet}, R_{wjet}$
Monte Carlo ratio of $W+$ jet events in the M_T tail to the M_T peak. Seaparately for electrons and muons, or combined.

$R^e_{top}, R^\mu_{top}, R_{top}$
Monte Carlo ratio of $t\bar{t}$ or single-top $\ell+$ jets events in the M_T tail to the M_T peak. Separately for electrons and muons, or combined.

$SFR^e_{wjet}, SFR^\mu_{wjet}, SFR_{wjet}$
Data/MC scale factors for $R^e_{wjet}, R^\mu_{wjet}, R_{wjet}$

$SFR^e_{top}, SFR^\mu_{top}, SFR_{top}$
Data/MC scale factors for $R^e_{top}, R^\mu_{top}, R_{top}$

K_3 and K_4
Scale factors for events with one or two extra jets from radiation

SF_{pre} and SF_{post}
Scale factors to be applied to MC to normalize to the yields in the M_T control region before and after the application of the isolated track veto.