CMS Draft Analysis Note

2015/01/16
Head Id: 273929
Archive Id: 273844:273929
Archive Date: 2015/01/15
Archive Tag: trunk

Search for Associated Production of a Single Top Quark and a Higgs Boson in Leptonic Channels

B. Stieger[1], C. Jorda Lope[2], F. Margaroli[3], and C. Rovelli[2]

[1] CERN
[2] INFN Roma
[3] Universitá di Roma La Sapienza and INFN Roma

Abstract

We perform a search for the production of a Higgs boson in association with a single top quark, using the $H \to W^+W^-$ and $H \to \tau^+\tau^-$ decay modes. The analysis uses the final states with two same-sign leptons, as well as three leptons, and uses the full 2012 CMS dataset at 8 TeV. Multivariate techniques are used to discriminate the signal from the dominant backgrounds. The analysis has an expected 95% confidence level upper limit on the single top quark plus Higgs boson production cross section of 5.0 times the expectation with C_t = -1.

Contents

1 Introduction . . . 2
2 Datasets, Triggers and Monte-Carlo Simulation . . . 3
2.1 MC pile-up reweighting . . . 4
3 Event Selection and Object Definitions . . . 5
3.1 Two same-sign lepton selection . . . 5
3.2 Three leptons selection . . . 9
4 Expected Background Composition . . . 15
5 Background involving fake leptons: $t\bar{t}$, Drell-Yan . . . 17
5.1 Fake Rate Method for Three Leptons Final State . . . 18
5.2 Fake Rate Method for Same-Sign Two Leptons Final State . . . 25
5.3 Background from Charge Mis-Assignment . . . 29
6 Shape Analysis . . . 30
6.1 Three Leptons Channel . . . 31
6.2 Same-Sign Dilepton Channel . . . 37
6.3 Other possible Background contributions . . . 45
7 Systematic uncertainties . . . 45
8 Results and Interpretation . . . 50
8.1 Checks on the effect of single systematic sources . . . 55
8.2 Checks on the post-fit nuisance parameters . . . 55
9 Conclusions . . . 64
A Additional Three-Lepton Plots . . . 69
B Additional Three-Lepton Tables . . . 72
C Additional Same-Sign Dilepton Plots . . . 72
C.1 Background Control Region Plots . . . 72
C.2 Cross-Check with Official POG ID . . . 75
D Systematic Shape Plots . . . 76
E Likelihood in Background Dominated Control Selection . . . 80
F Vetoing Hadronic Taus . . . 81
F.1 Tau Selection and Efficiencies . . . 81
F.2 Results with Tau Veto Applied . . . 82
F.3 Results without Tau Veto Applied . . . 82
G Overlap With $t\bar{t}H$ Multilepton Search . . . 83

1 Introduction

The discovery of a Higgs boson by the CMS and ATLAS experiments in 2012 [1, 2] opened a new field for explorations in the realm of particle physics. It is now critical to explore the coupling of this new particle with the other elementary particles to test whether it is the Higgs boson predicted by the standard model (SM). In particular, the study of the Yukawa structure of the coupling of the Higgs boson to fermions is at its beginning: as of today, the evidence for the Higgs boson coupling to fermions comes mostly from bottom quarks and tau leptons final states [3–5]. A fermionic coupling of special interest is the one of the new boson to the top quark. In fact it is widely believed that the top quark plays a special role in the electroweak symmetry breaking mechanism due to its very large mass [6, 7]. The Higgs boson has been discovered mainly through its direct coupling with the other known heavy bosons (W/Z) and photons where a SM interaction with top quark is assumed. New physics could alter the interaction between the top quark and the Higgs boson without exceeding current constraints.

The most straightforward way to study the coupling of top quarks to Higgs bosons is through the investigation of top-antitop-Higgs production. The search for $t\bar{t}H$ production has been performed in the Higgs boson decay to hadrons, photons, and leptons final states by the CMS collaboration. The combination of the above channels offered the first hint of direct coupling of the Higgs boson to the top quark [8]. The ATLAS collaboration has searched for $t\bar{t}H$ production using the Higgs boson to bottom/antibottom quarks and photons decays [9].

The coupling of the Higgs boson to a top quark (C_t) can be probed in a novel way by studying the associated production of a single top quark and a Higgs boson. Single top quark production has been observed at the LHC in the t-channel [10] and the s-channel production observation has been achieved at the Tevatron [11]. CMS recently reported the first observation of the associated production of a single top quark with a W boson [12]. Single top quark plus Higgs boson production proceeds mainly through t-channel diagrams, with the Higgs being emitted either from a top quark leg or from a W boson propagator (see Figure 1). In the SM, as the couplings of the Higgs to the W and the top quark have opposite sign, these two diagrams suffer from destructive interference, so that they almost cancel out: the process cross section, calculated at Next-to-Leading Order (NLO) precision, is approximately 18 fb [13–16].

Global fits of LHC and Tevatron Higgs data still allow for the possibility of a Higgs boson having negative couplings to fermions [17]. Because the interference between the two main production diagrams would become constructive, a negative coupling of the Higgs boson to the top quark (C_t = -1) would translate into a $\times 15$ increase in the single top quark plus Higgs boson cross section. Recent work in the literature thus suggested the investigation of tHq production in the Higgs boson to photons [18, 19], Higgs boson to bottom quarks [13] or Higgs boson to lepton final states [19].

The single top quark plus Higgs boson production cross section could be enhanced by other modifications of the SM. For example, new physics appearing at high mass scales could appear as flavor-changing-neutral-currents (FCNC) involving the top quark and the Higgs boson such as tHu, tHc, giving rise to final states overlapping with the final state under study here. While there is a large body of work in the literature on the phenomenology of such FCNC at colliders [20], the ATLAS and CMS collaborations produced only some preliminary investigations [21, 22]. Composite Higgs scenarios would predict the existence of a heavy top quark partner with large branching fraction into a top quark and a Higgs boson. The single top partner production cross section exceeds the pair production cross section under a large variety of assumptions of electroweak top quark partners couplings, for top quark partner masses above the the 700-800 GeV limit [23], that is to say above the current experimental constraints [24, 25].

This exotic production mode would give rise to final states that would overlap with the ones needed for studying single top plus Higgs boson production in the standard model.

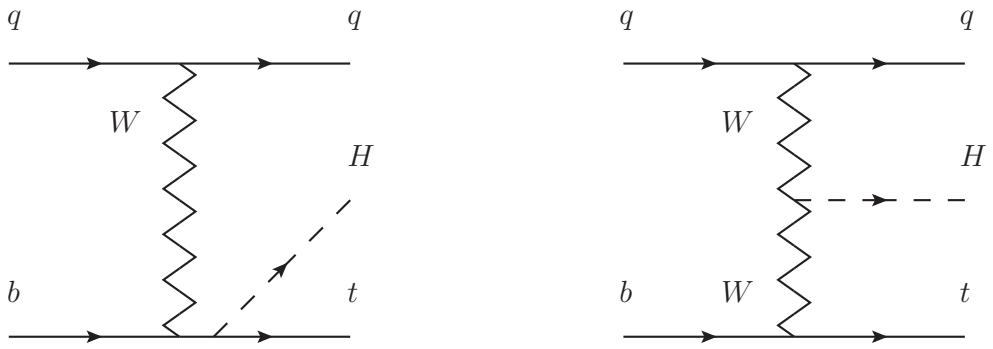

Figure 1: qtH production

A search for a negative value of top quark and Higgs boson Yukawa coupling, C_t, is performed studying the top quark plus Higgs boson production, qtH. For this search, the Higgs boson decay to a pair of W bosons and semileptonic decay of the top quark is considered, for two different final states: with two same-sign leptons or three leptons. For this analysis, the full pp collisions dataset collected by the CMS detector at a center of mass energy of 8 TeV is used.

2 Datasets, Triggers and Monte-Carlo Simulation

In this section the data and Monte Carlo samples used in the analysis are described.

The analysis is performed with data from pp collisions at 8 TeV. Since the selected events have two or three leptons in the final state the data collected with double lepton triggers are used, falling into the `DoubleMu`, `MuEG` and `DoubleElectron` Primary Datasets, see Table 1. The list of HLT trigger paths used to collect the events is shown in Table 2. The p_T thresholds for the leptons to fire these paths are 8 GeV and 17 GeV for both muons and electrons. For this reason, to be on the plateau of the trigger efficiency, two leptons with p_T larger than 10 GeV and 20 GeV respectively are required offline. The latest data reconstruction with `CMSSW` 53X (`22Jan2013 ReReco` [26]) is used, which includes the best calibration and alignment conditions available for the detector.

The integrated luminosity of the dataset used is 19.7 fb^{-1}, as estimated with `pixelLumiCalc.py` tool [27]. The JSON file used to select the data events, with all the subdetectors fully operational, is:
`Cert_190456-208686_8TeV_22Jan2013ReReco_Collisions12_JSON.txt`

For the background prediction official Monte Carlo samples are used, corresponding to the `Summer12_DR53X-PU` production, with pileup profile `START53_V7*`, see Table 3. The majority of the samples are simulated using the `MadGraph` [28] event generator, except for the QCD samples and the ttH production sample which are produced with `PYTHIA` [29]. In Table 3 also the estimated cross-section for each process is given. The NLO estimation is used, when available.

The signal is produced privately, as in the qtH (H $\rightarrow \gamma\gamma$) analysis [30]. Signal events are produced with `Madgraph5` [28], for two values of C_t: +1 and -1, and then passed to Pythia for underlying event/parton shower/hadronization. The masses for top and Higgs are m_{top} = 173 GeV

and m_H = 125 GeV, respectively. CTEQ6L1 Parton Distribution Functions (PDF) are used. The sample has been produced using the five flavor scheme (5FS), i.e. b quarks are present in the colliding partons. This choice does not model well the kinematics of the additional bjets; however, in this analysis no specific use is made of that information.

The signal samples contain nearly 1 million events, with three different decay modes for the Higgs boson: $\gamma\gamma$, WW and $\tau\tau$. The proportion of these decays is, respectively, 0.1, 0.696, and 0.204. As the analysis is designed to always accept the lepton coming from the top quark leptonic decay, and the second or third lepton from the leptonic decay of the W from the Higgs boson, the samples are produced with the top quark always decaying leptonically, in order to gain statistics. In Table 3 the signal samples are summarized.

Period	Primary Dataset
Run2012A	/DoubleMu/Run2012A-22Jan2013-v1/AOD
	/MuEG/Run2012A-22Jan2013-v1/AOD
	/DoubleElectron/Run2012A-22Jan2013-v1/AOD
Run2012B	/DoubleMuParked/Run2012B-22Jan2013-v1/AOD
	/MuEG/Run2012B-22Jan2013-v1/AOD
	/DoubleElectron/Run2012B-22Jan2013-v1/AOD
Run2012C	/DoubleMuParked/Run2012C-22Jan2013-v1/AOD
	/MuEG/Run2012C-22Jan2013-v1/AOD
	/DoubleElectron/Run2012C-22Jan2013-v1/AOD
Run2012D	/DoubleMuParked/Run2012D-22Jan2013-v1/AOD
	/MuEG/Run2012D-22Jan2013-v1/AOD
	/DoubleElectron/Run2012D-22Jan2013-v1/AOD

Table 1: Primary datasets used (2012 data).

Dataset	Trigger Path
/DoubleMu/Run2012	HLT_Mu17_Mu8_v*
	HLT_Mu17_TkMu8_v*
/MuEG/Run2012	HLT_Mu17_Ele8_CaloIdT_CaloIsoVL_TrkIdVL_TrkIsoVL_v*
	HLT_Mu8_Ele17_CaloIdT_CaloIsoVL_TrkIdVL_TrkIsoVL_v*
/DoubleElectron/Run2012	HLT_Ele17_CaloIdT_CaloIsoVL_TrkIdVL_TrkIsoVL_Ele8(...)
	(...)_CaloIdT_CaloIsoVL_TrkIdVL_TrkIsoVL_v*

Table 2: Trigger paths used to select data events.

2.1 MC pile-up reweighting

The Monte Carlo samples are produced with a different pile-up profile than the one in data, hence, simulated events need to be corrected for the different data conditions. For this, a per-event weight is estimated. The pile-up from simulation is taken directly from the MC truth information. The target pile-up distribution for data is derived using the per bunch-crossing and per luminosity section instantaneous luminosity, together with the total pp inelastic cross section, to generate an expected pile-up distribution, correctly weighted by the per-bunch-crossing per-lumi section integrated luminosity over the entire data-taking period.

The total pp inelastic cross section considered is 69.4 mb^{-1} [31]. In Figure 2 the distribution of the reconstructed primary vertices in a Monte Carlo sample after the pile-up reweighting correction is compared with the distribution in data, showing a good agreement.

Signal samples	σ[pb]
/CRAB_MC_LHE_1M-125_1/phys_higgs-qtH-blv_1M-mH125Ct1[*]-(...)/USER	0.0179
/CRAB_MC_LHE_1M-125_-1/phys_higgs-qtH-blv_1M-mH125Ct-1[*]-(...)/USER	0.2311
Background samples	$\sigma[\ textpb] \cdot$ BR
/WZJetsTo3LNu_TuneZ2_8TeV-madgraph-tauola/[1]	1.06
/ZZJetsTo4L_TuneZ2star_8TeV-madgraph-tauola/[1]	0.09
/WWJetsTo2L2Nu_TuneZ2star_8TeV-madgraph-tauola/[1]	5.81
/ZGToLLG_8TeV-madgraph/[1]	132.6
/DYJetsToLL_M-50_TuneZ2Star_8TeV-madgraph-tarball/[1]	3532.8
/DYJetsToLL_M-10To50filter_8TeV-madgraph/[1]	860.5
/TTJets_MassiveBinDECAY_TuneZ2star_8TeV-madgraph-tauola/[2]	234.0
/TTJets_FullLeptMGDecays_8TeV-madgraph-tauola/[3]	24.6
/TTWJets_8TeV-madgraph/[1]	0.232
/TTZJets_8TeV-madgraph/[1]	0.174
/WWWJets_8TeV-madgraph/[1]	0.0822
/WWZNoGstarJets_8TeV-madgraph/[1]	0.0633
/WZZNoGstarJets_8TeV-madgraph/[1]	0.0192
/TTH_Inclusive_M-125_8TeV_pythia6/[1]	0.1302
/QCD_Pt_20_30_BCtoE_TuneZ2star_8TeV_pythia6/[1]	167388
/QCD_Pt_30_80_BCtoE_TuneZ2star_8TeV_pythia6/[1]	167040
/QCD_Pt_80_170_BCtoE_TuneZ2star_8TeV_pythia6/[1]	12982
/QCD_Pt_20_30_EMEnriched_TuneZ2star_8TeV_pythia6/[1]	2914860
/QCD_Pt_30_80_EMEnriched_TuneZ2star_8TeV_pythia6/[1]	4615893
/QCD_Pt_80_170_EMEnriched_TuneZ2star_8TeV_pythia6/[1]	1832945
/QCD_Pt_20_MuEnrichedPt_15_TuneZ2star_8TeV_pythia6[1]	134680
[*] == Summer12_START50_V13_PU_2012_Startup_50ns_PoissonOOTPU	
[1] == Summer12_DR53X-PU_S10_START53_V7A-v1/AODSIM	
[2] == Summer12_DR53X-PU_S10_START53_V7C-v1/AODSIM	
[3] == Summer12_DR53X-PU_S10_START53_V7C-v2/AODSIM	

Table 3: Simulated datasets used for the analysis. The production cross sections times branching ratio are also listed.

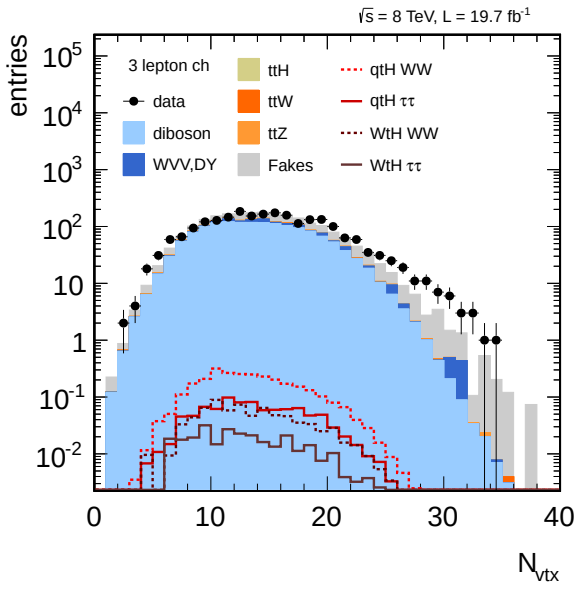

Figure 2: Number of reconstructed primary vertices for 2012 data compared with MC, for a loose preselection in the three lepton final state.

3 Event Selection and Object Definitions

A description of the physics objects and event selection is done, for two same-sign leptons and three leptons in the final state separately. This is due to the fact that the two analyses use different frameworks, as they inherit from previous analyses at CMS.

3.1 Two same-sign lepton selection

The objects used for the two same-sign lepton final state follow those used by the leptonic ttH analysis [32], which in turn are based on the $H \to ZZ^* \to 4l$ analysis, see references [33, 34]. Objects used for this analysis comprise muons, electrons, and jets, which can be b-tagged. Their exact selection procedure is described in [32] and is briefly summarized below.

Both electrons and muons are loosely pre-selected with high efficiency for signal leptons and moderate rejection of leptons from b-hadrons and misidentification of light jets. In a second selection step, an optimized lepton MVA, developed for the leptonic ttH analysis is used to select the final objects.

Efficiencies to reconstruct and identify leptons have been computed in data and simulation using the tag-and-probe technique for both the pre-selection and lepton MVA steps (see below). From the measured values data/simulation scale factors are derived and applied in the analysis to correct the expected MC yields.

3.1.1 Muon selection

The muon objects considered in this analysis are reconstructed within the geometrical acceptance $|\eta^{\mu}| < 2.4$ and with transverse momentum $p_{\mathrm{T}}^{\mu} > 5\,\mathrm{GeV}$.

For the lepton preselection we choose the *Loose Muon selection*, which accepts all the particle-flow muon candidates (PF Muons) without additional quality criteria. The PF Muons are selected among the reconstructed muon track candidates by applying minimal requirements on the track components in the muon system and taking into account a matching with small energy deposits in the calorimeters. More details of the particle-flow muon selection are described in Ref. [35].
In addition we require a very loose selection on the distance of the muon track from the primary vertex in order to reject badly reconstructed tracks or tracks from pileup vertices. In particular we request the significance of the impact parameter to the event vertex, $\mathrm{SIP}_{3\mathrm{D}} = |\frac{\mathrm{IP}}{\sigma_{\mathrm{IP}}}|$, to be less than 10, and the distances in the transverse plane, d_{xy}, and along the z axis, d_z, to be less than 0.5 cm and 1 cm respectively.

Momentum scale calibrations are applied to muons both in data and in simulations, derived from the $1/p_{\mathrm{T}}$ distributions of muons from Z decays [36]. The corrections are designed to calibrate the overall momentum scale and to remove any dependency of the scale on the p_{T}, η, ϕ, and charge of the muon, thereby improving also the momentum resolution. In addition, on simulation a smearing of the momentum measurement is applied to better match the momentum resolution in data.

3.1.2 Electron selection

The electron objects considered in this analysis are reconstructed within the geometrical acceptance $|\eta^{\mathrm{e}}| < 2.5$ and with transverse momentum $p_{\mathrm{T}}^{\mathrm{e}} > 7\,\mathrm{GeV}$. Instead of starting from the PF electron candidate collections, reconstructed candidates are first obtained in an inclusive way to gain efficiency.

The identification of electrons relies on a Boosted Decision Tree (BDT) multivariate technique which was trained using a Higgs boson Monte Carlo sample for the signal and a $W + 1$-fake electron data sample for background. The working point was optimized using a $Z + 1$-fake electron data sample. More details can be found in [33].

The cut values on the BDT output resulting from the optimization procedure are summarized below:

- $5 < p_T < 10$ GeV:
  - $|\eta| < 0.8 : BDT > 0.47$
  - $0.8 < |\eta| < 1.479 : BDT > 0.004$
  - $|\eta| > 1.479 : BDT > 0.295$
- $p_T > 10$ GeV
  - $|\eta| < 0.8 : BDT > 0.5$
  - $0.8 < |\eta| < 1.479 : BDT > 0.12$
  - $|\eta| > 1.479 : BDT > 0.6$

Similar to the muons, we require a very loose selection on the distance of the electron track from the primary vertex in order to reject badly reconstructed tracks or tracks from pile-up vertices. We request the significance of the impact parameter to the event vertex, $\mathrm{SIP}_{3D} = |\frac{\mathrm{IP}}{\sigma_{\mathrm{IP}}}|$, to be less than 10, and the distances in the tranverse plane, d_{xy}, and along the z axis, d_z, to be less than 0.5 cm and 1 cm respectively. Finally the number of missing hits in the innermost tracker layer must be less than two.

The electron energy scale in data is corrected based on $Z \to ee$ data and simulated events to deal with discrepancies in the low p_T part of the spectrum. Electrons in simulated events are then smeared to have the best match between data and simulation, as described in [33, 34].

3.1.3 Lepton Isolation

The isolation of individual e or μ leptons is measured relative to their transverse momentum p_{T}^{ℓ}, by summing over charged and neutral particles in a cone $\Delta R = \sqrt{(\eta^{\ell} - \eta^{i})^2 + (\phi^{\ell} - \phi^{i})^2} < 0.4$ around the lepton direction at the interaction vertex:

$$R_{\mathrm{Iso}}^{\ell} \equiv \left(\sum_{\mathrm{charged}} p_{\mathrm{T}} + \mathrm{MAX}\left[0, \sum_{\mathrm{neutral}} p_{\mathrm{T}} + \sum_{\gamma} p_{\mathrm{T}} - 0.5 \sum_{\mathrm{charged,PU}} p_{\mathrm{T}} \right] \right) / p_{\mathrm{T}}^{\ell}, \quad (1)$$

where $\sum_{\mathrm{charged}} p_{\mathrm{T}}$, $\sum_{\mathrm{neutral}} p_{\mathrm{T}}$, and $\sum_{\gamma} p_{\mathrm{T}}$ are respectively the scalar sums of the transverse momenta of charged particles from the primary vertex, neutral hadrons, and photons located in the lepton cone. The contribution of pileup photons and neutral hadrons is estimated from the scalar sum of the transverse momenta of charged hadrons from pileup vertices in the cone, $\sum_{\mathrm{charged,PU}} p_{\mathrm{T}}$. This quantity is multiplied by a factor of 0.5, which corresponds approximately to the ratio of neutral to charged hadron production in the hadronization process of pileup interactions, as estimated from simulation. Possible double counting in the isolation evaluation, caused by small differences between reconstructed electron candidates and those identified from the PF algorithm, is avoided by applying specific vetoes [34]. The electrons or muons are considered isolated if $R_{\mathrm{Iso}}^{\ell} < 0.4$.

3.1.4 Lepton MVA Discriminator

To further suppress leptons from mis-identified light jets, B-hadron decays, and photon conversions, the leptonic ttH analysis developed and implemented a lepton MVA discriminator that

optimally selects signal leptons while rejecting such fakes. The event selection of the dilepton channel of this analysis and that of the ttH analysis are diverging only in the selection of the number of b-tagged jets and that of a forward jet and hence the background composition is expected to be similarly dominated by fake leptons from $t\bar{t}$ events. We therefore attempt no further re-optimization of the MVA and recycle it for this analysis.

We briefly summarize the MVA, its input variables, and the validation studies performed and refer to the original work for further details. [32]

Three groups of input variables have been considered for the MVA: variables related to the impact parameter of the lepton computed with respect to the primary vertex, variables related to the isolation of the lepton considering separately the neutral and charged candidate deposits in the lepton isolation cone, and variables related to the jet reconstructed in the event closest to the lepton. The definition of variables finally used in the training is given below:

- For both muons and electrons:
  - significance of the impact parameter with respect to the primary vertex, $|\mathrm{SIP}_{3\mathrm{D}} = \frac{\mathrm{IP}}{\sigma_{\mathrm{IP}}}|$;
  - the contribution of the charged hadrons in the lepton isolation cone, $\sum_{\mathrm{charged}} p_{\mathrm{T}}$;
  - the contribution of the neutral hadrons and photons in the lepton isolation cone, $\sum_{\mathrm{neutral}} p_{\mathrm{T}} + \sum_{\gamma} p_{\mathrm{T}}$;
  - ΔR distance between the lepton and the closest jet;
  - the ratio between the p_{T} of the lepton and the p_{T} of the closest jet;
  - the CSV b-tagging discriminator value of the closest jet;
- For muons only:
  - distance in the transverse place, d_{xy}, with respect to the primary vertex;
  - distance on the z axis, d_z, with respect to the primary vertex;
- For electrons only:
  - the multivariate discriminator used for the electron preselection;
  - number of missing hits in the innermost tracker layer.

A boosted decision tree (BDT) discriminator is trained on simulated prompt leptons in events from the $t\bar{t}H$ sample and on fake leptons from the $t\bar{t}$+jets sample, separately for muons and electrons, and in several categories of p_{T} and η.

The improvement in signal efficiency vs. background rejection compared to some other tight lepton selections is shown in Figure 3, taken from [32]. The working point used in this analysis is the one at a cut of the discriminator output of greater than 0.7.

The description of input variable shapes by the simulated MC samples has been extensively validated in signal and background dominated control regions, and a good agreement has been observed in general.

In the analysis, MC yields are corrected for differences in the MVA cut efficiency between data and MC, using scale factors derived from tag-and-probe efficiency measurements.

3.1.5 Jet selection

Jets are reconstructed by clustering PF candidates using the the anti-k_{T} algorithm with distance parameter $\Delta\mathrm{R} = 0.5$, as implemented in the FASTJET package.

The charged hadrons not coming from the primary vertices are subtracted from the PF can-

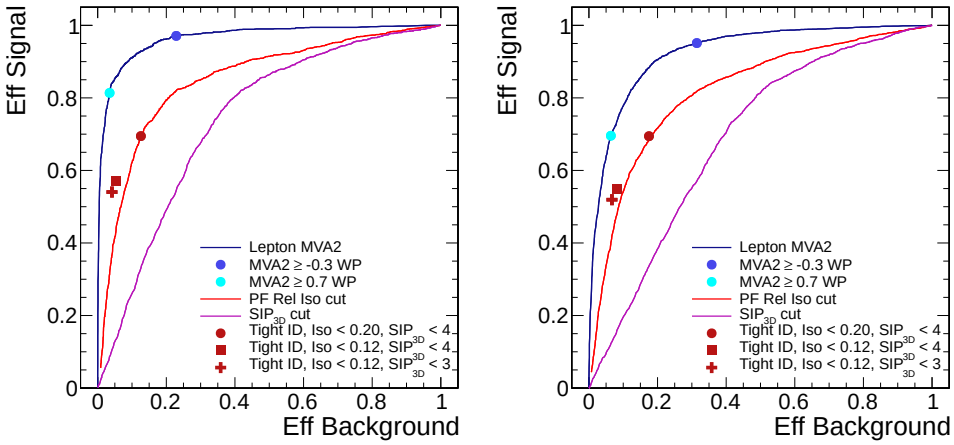

Figure 3: ROC curves to compare the performance of the lepton MVA discriminator with respect to the other usual tight lepton selections for the same-sign $\mu\mu$ channel (left) and same-sign $e\mu$ channel (right). [32]

didates considered in the clustering. Jet energy corrections are applied as a function of the jet E_T and η. In addition, a multivariate discriminator is applied to distinguish between jets coming from the primary vertex and jets coming from pile-up vertices. The discrimination is based on the differences in the jet shapes, in the relative multiplicity of charged and neutral components, and in the different fraction of transverse momentum which is carried by the hardest components. Within the tracker acceptance the jet tracks are also required to be compatible with the primary vertex. Jets are only considered if they have a transverse energy above 25 GeV. In addition, they have to be separated from any lepton candidates by requiring $\Delta R = \sqrt{(\eta^{\ell} - \eta^{jet})^2 + (\phi^{\ell} - \phi^{jet})^2} > 0.5$.

The CSV b-tagging algorithm is used to identify jets that are likely to originate from the hadronization of bottom quarks. The efficiency to tag b-jets and the rate of misidentification of non-b jets depend on the operating point chosen. Both the efficiency and the fake rate are parameterised as a function of the transverse momentum and pseudorapidity of the jets. These performance measurements are obtained directly from data in samples that can be enriched in b jets, such as $t\bar{t}$ and multijet events where a muon can be found inside the one of jets. Two working points for the CSV output discriminant are considered in the analysis. The *loose* one (CSV > 0.244) has approximately 85% efficiency to tag jets with b quarks and a 10% chance to tag jets with only light quarks or gluons. The *medium* working point (CSV > 0.679) has approximately 70% efficiency for tagging jets with b quarks and 1.5% efficiency to tag jets with only light quarks or gluons. Corrections that take into account the different performances of the CSV tagging in data and simulation are applied to the MC samples used in the analysis.

3.2 Three leptons selection

The object selection used for the three lepton final state is mainly based on that used for the fully leptonic HWW analysis at a center of mass energy of 8 TeV [37–40].

In this analysis several objects are used: muons, electrons, jets and missing transverse energy. A detailed description of the reconstruction of each object can be found elsewhere [41, 42]. In this section the final selection of each object is detailed.

Several validation plots with a loose three lepton pre-selection can be found in appendix A.

3.2.1 Muon selection

The muon selection used is fully described in References [37, 38]. It is based on the tight muon selection provided by the Muon POG, with some modifications. A loose and a tight selection are both used. The first one is used to compute the fake rate for the data driven background estimation, which is explained in Section 5. It is chosen to be loose enough but still tighter than the trigger requirements. The tight selection is used for the event selection.

1. **Loose muon selection**
   - $p_{\mathrm{T}} > 10$ GeV and $|\eta| < 2.4$
   - The muon is identified as a `PFMuon`. Also, it can be identified as Global or Tracker muon, with:
     - Global: `isGlobalMuon`, χ^2/ndof < 10, number of matches in the muons system higher than 1, and at least one muon chamber hit
     - Tracker: `isAllTrackerMuons` and `IsTMLastStationTight`
   - The expected relative uncertainty on p_{T} is lower than a 10%, $\delta p_{\mathrm{T}}/p_{\mathrm{T}} < 0.1$
   - The number of layers in the track is at least 5, and there is at least one pixel hit
   - The χ^2/ndof for the kink finder algorithm is lower than 20
   - The impact parameter cuts are: $|\mathrm{d}_z| < 0.1$ cm and $|\mathrm{d}_0| < 0.1$ cm
   - The output of the isolation MVA, described below, must be $\geq$ -0.6
2. **Tight muon selection**
   - The muon passes the Loose selection
   - The impact parameter cuts are: $|\mathrm{d}_0| < 0.01$ (0.02) cms for muons with $p_{\mathrm{T}} < 20$ GeV (> 20 GeV)
   - The output of the isolation MVA passes the selection below

The muon isolation is based on the Particle-Flow algorithm. A cone with ΔR =0.5 is considered and splitted into 5 isolation rings. Each of them, IsoRing, is corrected for the pile-up using the Fastjet technque as IsoRing = PFIso - $\rho \cdot \mathrm{A}_{eff}$. The effective area A_{eff} is optimized to maximize the selection of prompt muons, ρ is the median of the energy density distribution for the particles within the area of any jet in the event and PFIso is the energy deposit around the muon estimated with the PF algorithm. A MVA algorithm is used which combines the different IsoRings, exploiting the radial distributions of the Particle-Flow candidates inside the cone. For the tight selection, a muon is considered isolated when its MVA isolation value is greater than 0.86 (0.82) for muons with $p_{\mathrm{T}} < 20$ GeV in the barrel (endcap) and greater than 0.82 (0.86) for muons with $p_{\mathrm{T}} > 20$ GeV in the barrel (endcap).

The muon momentum scale is corrected using the Rochester Corrections [36, 43]. This method is based on the binned parametrization of data and MC momentum bias and resolution using $\mathrm{Z} \rightarrow \mu\mu$ invariant mass distributions.

3.2.2 Electron selection

The electron selection used is also based on that used in previous HWW fully leptonic analyses. As for the muons, a loose and a tight selection are used, where the loose one is tight enough with respect to the trigger:

1. **Loose electron selection**

- $p_T > 10$ GeV and $|\eta| < 2.5$
- Cuts on some identification variables as applied at HLT level. Numbers are given for $|\eta| < 1.479$ ($|\eta| > 1.479$) electrons:
  - $\sigma_{i\eta i\eta} < 0.01$ (0.03)
  - $|\Delta\eta_{in}| < 0.15$ (0.1)
  - $|\Delta\phi_{in}| < 0.007$ (0.009)
  - H/E < 0.12 (0.10)
- The electron passes the photon conversion veto and has no expected missing hits in the track
- max(Σ_{ECAL} $E_T - 1$) $/ p_T < 0.2$
- Σ_{HCAL} E_T $/ p_T < 0.2$
- Σ_{Tracker} E_T $/ p_T < 0.2$

2. **Tight electron selection**
   - The electron passes the Loose selection
   - The impact parameter cuts are: $|d_z| < 0.1$ cm and $|d_0| < 0.02$ cm
   - $[\Sigma_{\text{chargedH}} + \max(\Sigma_{\text{neutralH}} + \Sigma_{\text{photons}} - \rho \cdot A_{\text{eff}}, 0)] / p_T < 0.15$
   - The output of the triggering electrons identification MVA, provided by the EGamma POG, passes the cuts below, which depend on the p_T and η of the electron:
     - MVA > 0.0 (0.94) for $|\eta| < 0.8$ and $p_T < 20$ (> 20) GeV
     - MVA > 0.10 (0.85) for $0.8 < |\eta| < 1.479$ and $p_T < 20$ (> 20) GeV
     - MVA > 0.62 (0.92) for $|\eta| > 1.479$ and $p_T < 20$ (> 20) GeV

Energy scale corrections and resolution smearings are applied to improve the determination of the electron momentum on top of regression-based corrections to the ECAL energy and the E-p combination. This is done as in the H→ZZ→4l analysis and consistently with the POG recommendation [44].

3.2.3 Lepton efficiencies

No scale factor to account for the trigger efficiencies is applied in this analysis and a 100% efficiency is assumed. This is based on other analyses with three leptons in the final state, in which the HLT efficiency measurement gave a result close to 100% [45].

For the offline selection, for both muons and electrons data-to-MC scale factors are used to correct for possible differences in the lepton reconstruction, identification and isolation efficiency. These scale factors are estimated using the Tag & Probe method with Z→ $\ell\bar{\ell}$ events.

For the electron selection used, official scale factors by the EGamma POG were already available for `22Jan2013ReReco` data [46] and are given in Table 4.

In the case of muons, the factors were not computed centrally, so needed to be reestimated. The datasets used are `SingleMu 22Jan2013ReReco` and MadGraph Drell-Yan, and the ntuples are provided centrally by the Muon POG [47]. The method used is the same as in the HWW fully leptonic analysis. Figure 4 shows the efficiency of the muon identification, impact parameter requirements and isolation, for both data and MC events, for three different regions in the detector: $0 < |\eta| \leq 0.8$, $0.8 < |\eta| \leq 1.2$ and $1.2 < |\eta| \leq 2.4$. The scale factors are summarized in Table 5.

p_T range (GeV)	$0 < \|\eta\| \leq 0.8$	$0.8 < \|\eta\| \leq 1.44$	$1.44 < \|\eta\| \leq 1.55$	$1.55 < \|\eta\| \leq 2.0$	$2.0 < \|\eta\| \leq 2.5$
$10 < p_T \leq 15$	0.919 ± 0.008	0.967 ± 0.011	1.022 ± 0.035	0.883 ± 0.033	0.625 ± 0.023
$15 < p_T \leq 20$	0.971 ± 0.004	0.971 ± 0.003	0.999 ± 0.017	0.930 ± 0.008	0.735 ± 0.012
$20 < p_T \leq 30$	0.985 ± 0.005	0.989 ± 0.001	1.001 ± 0.004	0.987 ± 0.002	0.934 ± 0.004
$30 < p_T \leq 40$	0.998 ± 0.000	0.998 ± 0.000	1.011 ± 0.002	0.987 ± 0.001	0.966 ± 0.001
$40 < p_T \leq 50$	1.000 ± 0.000	0.999 ± 0.000	1.006 ± 0.002	0.993 ± 0.001	0.980 ± 0.001
$50 < p_T$	1.000 ± 0.000	0.999 ± 0.000	1.005 ± 0.002	0.998 ± 0.001	0.987 ± 0.002

Table 4: Electron data to MC scale factors. Errors from the fit only.

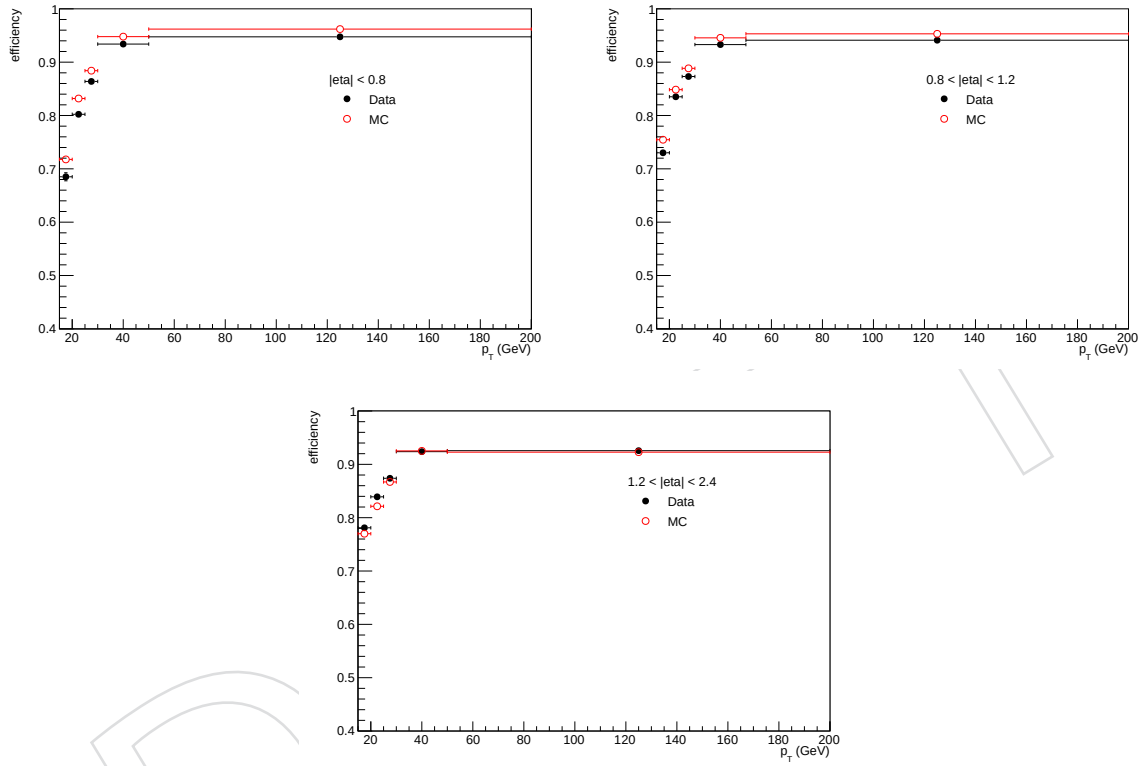

Figure 4: Muon identification, impact parameter and isolation efficiencies for data and MC, for $0 < |\eta| \leq 0.8$, $0.8 < |\eta| \leq 1.2$ and $1.2 < |\eta| \leq 2.4$

p_T range (GeV)	$0 < \|\eta\| \leq 0.8$	$0.8 < \|\eta\| \leq 1.2$	$1.2 < \|\eta\| \leq 2.4$
$10 < p_T \leq 15$	0.923 ± 0.010	0.962 ± 0.007	1.017 ± 0.004
$15 < p_T \leq 20$	0.955 ± 0.011	0.968 ± 0.005	1.014 ± 0.002
$20 < p_T \leq 25$	0.964 ± 0.001	0.984 ± 0.003	1.022 ± 0.001
$25 < p_T \leq 30$	0.977 ± 0.000	0.983 ± 0.000	1.008 ± 0.001
$30 < p_T \leq 50$	0.985 ± 0.000	0.987 ± 0.000	0.998 ± 0.000
$50 < p_T$	0.985 ± 0.000	0.987 ± 0.000	1.003 ± 0.001

Table 5: Muon data to MC scale factors. Statistical errors only.

3.2.4 Missing transverse energy selection

The missing transverse energy (MET) used in this analysis is the one computed using the PF algorithm, with Type-I corrections applied. Linear corrections are applied on both the x and y

components of the MET to reduce the observed ϕ modulation [48]:

$$\text{(MC) corr}_x = 1.62861\text{e-}01 - 2.38517\text{e-}02 \cdot N_{vtx}$$
$$\text{(MC) corr}_y = 3.60860\text{e-}01 - 1.30335\text{e-}01 \cdot N_{vtx}$$

$$\text{(Data) corr}_x = 4.83642\text{e-}02 + 2.48870\text{e-}01 \cdot N_{vtx}$$
$$\text{(Data) corr}_y = -1.50135\text{e-}01 - 8.27917\text{e-}02 \cdot N_{vtx}$$

With this correction factors, the x, y MET components (MET_x and MET_y) are corrected as:

$$MET_x = (MET_x - corr_x)$$
$$MET_y = (MET_y - corr_y)$$

3.2.5 Jet selection

The jets used in the analysis are reconstructed with the PF algorithm and the anti-k_T clustering algorithms with R = 0.5. All levels of energy corrections are applied: L1L2L3 and residual on data.

A loose jet identification is applied to clean the sample while keeping a sufficiently high efficiency on the signal. First the jets are required to pass the loose JetID selection recommended by the JetMet POG. Then the same cut based ID agains pile-up as in the qtH, $H \to \gamma\gamma$ analysis is used [49]. This selection relies on the usage of two among the most discriminating variables against pileup: β^* and $\langle \Delta R^2 \rangle$. The variable β^* is defined as the sum of the p_T of all the PF charged candidates associated to a primary vertex different from the chosen one, divided by the sum of the p_T of all charged candidates in the jet:

$$\beta^* = \frac{\Sigma_{\text{other PV}} p_{Ti}}{\Sigma p_{Ti}}$$

$\langle \Delta R^2 \rangle$ is defined as:

$$\langle \Delta R^2 \rangle = \frac{\Sigma \Delta R_i^2 p_{Ti}^2}{\Sigma p_{Ti}^2}$$

The value of β^* increases (thus, deteriorates vertex separation) when large pile-up is present. The cut based pileup ID requirements used in the analysis are:

- $\langle \Delta R^2 \rangle < 0.06$, if $|\eta| < 2.5$, < 0.05, if $|\eta| < 3.0$, or < 0.055, if $|\eta| > 3.0$
- $\beta^* < 0.2 \cdot \text{Log}(N_{vtx} - 0.64)$

Jets coming from b quark hadronization can be identified with the so called b-tagging technique, which exploits the distinct properties of the decay of b-hadrons, which are the product of b quark fragmentation. In this analysis, the Combined Secondary Vertex (CSV) algorithm is used to tag the jets and the medium working point is chosen.

Because of the signal kinematics, one jet from b and at least one forward jet are expected in the final state. Jets are selected with a $p_T > 25$ GeV and no limit on η. They are defined as b-tagged, forward or central with the requirements shown as Table 6:

As shown in Table 6, the forward jets are defined as those which have $|\eta| > 1.5$. In the part of this region where the tracker is present, $1.5 < |\eta| < 2.4$, jets are considered as forward if they are not b-tagged. The cut at $|\eta| = 1.5$ is not arbitrary, but is close to that in which the

Jet type	Conditions
b-tagged	$\|\eta\| < 2.4$ && CSV > 0.679
forward	($\|\eta\| > 1.5$ && CSV < 0.679) OR ($\|\eta\| > 2.4$)]
central	not b-tagged and no forward

Table 6: Jet selection

distributions for b quark and forward quark at generation level cross each other. This is shown in Figure 5 where the η and p_T distributions for qtH production for the light q quark and b quark at generation level are shown.

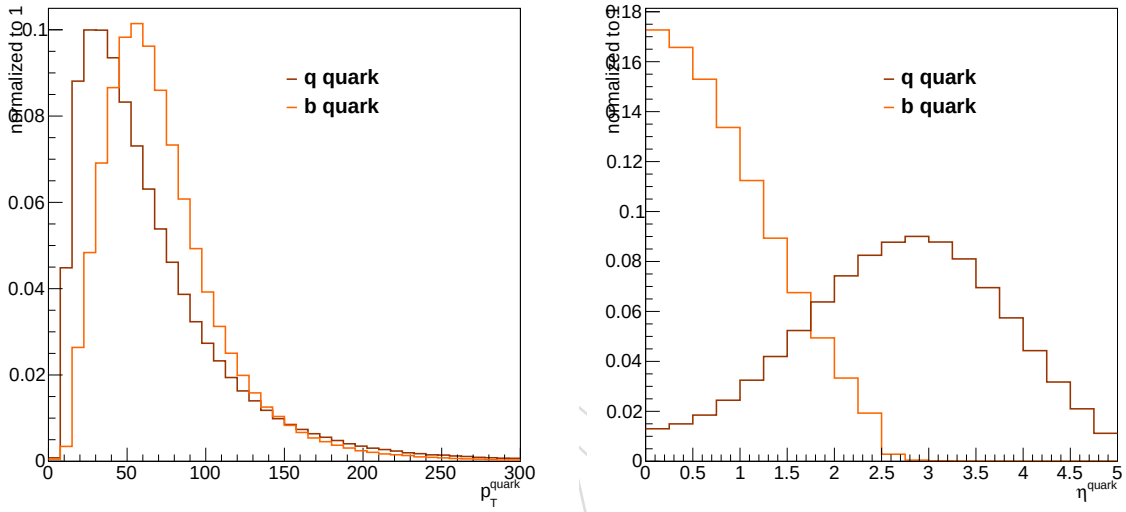

Figure 5: Distribution of p_T (left) and η (right) for the forward and b quarks on the qtH signal at generation level

3.2.6 b-Tagging corrections

As for the leptons, data to MC scale factors are also needed to correct the b-tagging efficiency measured on MC. The BTV POG provides centrally a set of b-tagging efficiency scale factors [50], SF, for b and light flavour jets, where:

$$SF = \epsilon_{Data} / \epsilon_{MC}$$

These SF depend on the jet p_T, η and flavour. Once these are provided, different methods can be chosen to correct the different efficiencies. In this analysis, the correct event yield in data for signal is predicted by only changing the weight of the selected MC events. In this method, the probabilities of a given configuration of jets in MC simulation and data are defined as:

$$P(MC) = \prod_{i\,=\,\text{tagged}} \varepsilon_i \prod_{j\,=\,\text{not tagged}} (1 - \varepsilon_j)$$

$$P(Data) = \prod_{i\,=\,\text{tagged}} SF_i \varepsilon_i \prod_{j\,=\,\text{not tagged}} (1 - SF_j \varepsilon_j)$$

where ε_i is the MC b-tagging efficiency for the selected CSV working point, and SF_i and ε_i are,

functions of the jet flavor, jet p_T, and jet η. The event weight is calculated as:

$$w = \frac{\mathrm{P(Data)}}{\mathrm{P(MC)}}$$

The SF values are provided centrally, and they are summarized in Table 7.

b and c flavours	
η range (GeV)	SF function
$0.0 < \eta < 2.4$	(0.938887 + (0.00017124·x) + (-2.76366x10^{-7} ·x^2)
light flavours	
η range (GeV)	SF function
$0.0 < \eta < 0.8$	(1.07541 + 0.00231827·x) + (-4.74249x10^{-6} ·x^2) + (2.70862x10^{-9} ·x^3)
$0.0 < \eta < 1.6$	(1.05613 + 0.00114031·x) + (-2.56066x10^{-6} ·x^2) + (1.67792x10^{-9} ·x^3)
$1.6 < \eta < 2.4$	(1.05625 + 0.000487231·x) + (-2.22792x10^{-6} ·x^2) + (1.70262x10^{-9} ·x^3)

Table 7: b-tagging scale factors

The tagging efficiencies ε_i are estimated using simulated signal events with a given loose preselection, to account for the signal kinematics. These efficiencies are calculated as a function of p_T and η of the jet for b and light flavours, and for the full p_T, η range for c flavour, because of the lack of statistics. In this last case, the estitimated efficiency is ε = (0.14 $\pm$ 0.03). For b and light flavoured jets, the efficiency maps are shown in Tables 8 and 9.

b flavour jets efficiency			
p_T range (GeV)	$0 < \lvert\eta\rvert \leq 0.8$	$0.8 < \lvert\eta\rvert \leq 1.6$	$1.6 < \lvert\eta\rvert \leq 2.4$
$25 < p_T \leq 50$	0.632 ± 0.040	0.578 ± 0.043	0.600 ± 0.062
$50 < p_T \leq 80$	0.720 ± 0.039	0.689 ± 0.048	0.660 ± 0.075
$80 < p_T \leq 120$	0.780 ± 0.051	0.738 ± 0.065	0.630 ± 0.114
$120 < p_T$	0.756 ± 0.064	0.675 ± 0.082	0.669 ± 0.177

Table 8: CSV Medium working point efficiency for qtH signal C_t = -1 events for b flavour jets. Statistical errors only.

light flavour jets efficiency		
p_T range (GeV)	$0 < \lvert\eta\rvert \leq 1.2$	$1.2 < \lvert\eta\rvert \leq 2.4$
$25 < p_T \leq 50$	0.042 ± 0.010	0.011 ± 0.003
$50 < p_T$	0.010 ± 0.004	0.025 ± 0.006

Table 9: CSV Medium working point efficiency for qtH signal C_t = -1 events for light flavour jets. Statistical errors only.

4 Expected Background Composition

In this section we describe the composition of the background processes for both channels. Figure 6 shows sketches of the signature of our signal process in the two considered channels.

While diboson backgrounds have a larger cross section, they are highly reduced imposing a Z-veto (WZ,ZZ) and requiring no extra lepton in the event (WZ,ZZ) or a b-tagged or forward

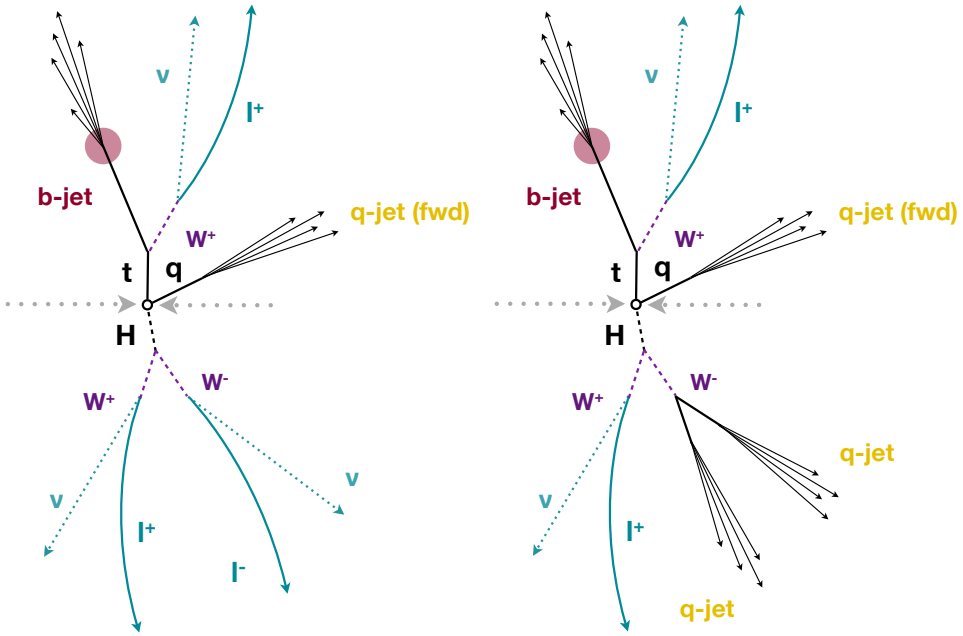

Figure 6: Sketch of tHq final state for the three lepton (left) and same-sign dilepton (right) channels.

jet (WW,WZ,ZZ). The Drell-Yan contribution is highly reduced in the case of three lepton final state because of the third lepton requirement, together with the jet counting, and the Z-veto.

Production of three electroweak bosons has a very small cross section and is further reduced by rejecting events with extra leptons. In addition both diboson and triboson production does in general not include forward jets or jets coming from b-quark decays.

Processes like $t\bar{t}W^{\pm}$, $t\bar{t}Z$ and $t\bar{t}H$, although with a small cross section, also have a high lepton and jets multiplicity (depending on the boson and top decays) and can be therefore important backgrounds.

Due to the relatively large cross section, the $t\bar{t}$ production is the major source of background for both channels, when additional leptons get produced in the decay of a B-hadron or a light jet gets mis-identified as a lepton, as explained in the following sections.

An additional background in the case of same-sign dileptons arises when the charge of a lepton in events with an originally opposite-sign pair is mis-identified. This happens not so much due to a track mis-reconstruction, but rather due to strongly asymmetric conversions of hard bremsstrahlung photons emitted from the initial lepton, and is therefore much more likely to happen for electrons than for muons. In the case where the original electron loses most of its energy to the radiated photon and out of the converted $e^{+}e^{-}$ pair the one with opposite charge carries most of the momentum, the resulting track can have opposite curvature to the original lepton.

Furthermore, the same-sign channel gets some contribution from the associated production of two W bosons with same charge, and two light jets.

The main sources of background, $t\bar{t}$ and events with charge mis-identification, are estimated using data driven methods, see Section 5. In the case of the other processes, they are estimated using MC prediction, with datasets summarized in Table 3. The expected number of events for each background is estimated taking into account the integrated luminosity, and corrected

by the proper scale factors, when necessary. Table 10 summarizes the background contribution considered per each channel. In the following, for the three lepton final state, we refer as WVV+DY as the sum of WWW, WWZ, WZZ, and Drell-Yan backgrounds, as the total contribution of these to the final likelihood selection is small. Diboson label refers to the sum of WW, WW, and ZZ.

Channel	Background from data driven methods	Background from MC prediction
$e\mu$ ee $\mu\mu$	$t\bar{t}$, DY(charge-misid.)	$t\bar{t}W^{\pm}$, $t\bar{t}Z$, $t\bar{t}H$, $t\bar{t}\gamma$, $t\bar{t}\gamma^{*}$ $W^{+}W^{+}qq$, WZ
$\ell\ell\ell$	$t\bar{t}$	WW, WZ, ZZ, WWW, WWZ, WZZ, DY, $t\bar{t}W^{\pm}$, $t\bar{t}Z$, $t\bar{t}H$

Table 10: Background composition and method used for its estimation, per channel

5 Background involving fake leptons: $t\bar{t}$, Drell-Yan

The main source of background for both the two same-sign leptons and three leptons final states is the combination of prompt leptons from W and Z boson decays with leptons produced in the hadronization of heavy flavor quarks or from mis-reconstructed light quarks, mainly from $t\bar{t}$ production. The methods used for the estimation of this background for the two channels are very close, but since the objects and the measured properties are different (as different lepton definitions are used) they are explained separately, to avoid confusion. Sec. 5.1 describes the method itself and its exact implementation in the context of the three lepton selection. Sec. 5.2 gives the parameters of the method for the two leptons final state, which are inherited from the leptonic ttH analysis.

Another type of background has to be taken into consideration for the same-sign dilepton channel, where the charge of a lepton is mis-identified and an event with a generated opposite-sign pair is reconstructed as having equal charges. Depending on the event selection, the source of this background is mainly $t\bar{t}$, but can be dominated by Drell-Yan for looser selection and for the *ee* channel. Sec. 5.3 describes how this contribution is estimated using the charge misidentification probability and a control selection with opposite-sign dileptons.

5.1 Fake Rate Method for Three Leptons Final State

In the case of the three lepton final state, $t\bar{t}$ events should be in principle highly rejected when requiring three prompt leptons, but still some contribution remain. This is due to the fact that two prompt leptons are expected from the leptonic decay of the W bosons, and a third lepton can be originated from a b-quark decay and identified as a prompt lepton. These leptons are called *fake leptons* in the contest of this analysis, since they are real but not prompt. An example of this situation is shown in Figure 7.

In the following, a detailed description of the data driven method used for the estimation of $t\bar{t}$ background is done. It is called Fake Rate method, as it based on the estimation of the

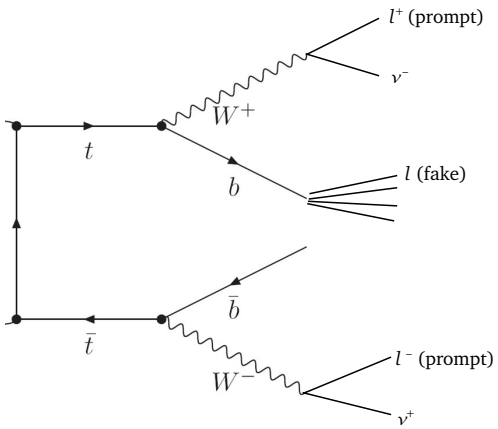

Figure 7: Diagram of a $t\bar{t}$ event in which the third lepton is coming from a b-quark decay.

probability for a lepton coming from a b-quark decay to be mis-identified as a prompt lepton. A detailed description can be found at Ref. [51]. According to this method, leptons in the event can be categorized into two types: fake and prompt. For this analysis, prompt leptons are real and isolated ones coming from the decay of the W boson. Fakes are real leptons in jets from heavy quarks which are not isolated.

Once the fake rate is measured, for both muons and electrons, it is used as a weight to extrapolate from a *loose+loose+ loose* (leptons) control region the estimated number of events from mis-identified leptons in the signal region. In our case the main contribution is coming from the $t\bar{t}$ process. The method is then separated into two main steps: (1) measurement of the fake rate for leptons and (2) extrapolation with this FR to the signal region.

5.1.1 Measurement of the Fake Rate for leptons

The Fake Rate (FR) is the probability for a loose lepton (fakeable object) to be identified as a prompt lepton (tight lepton). The FR is expressed as:

$$\mathrm{FR} = \frac{\mathrm{N_{tight}}}{\mathrm{N_{loose}}} \tag{2}$$

where $\mathrm{N_{loose}}$ is the number of candidates that pass the loose selection, and $\mathrm{N_{tight}}$ the number of loose leptons that also fulfill the tight requirements used in the analysis. The selection for loose and tight leptons was presented in Sections 3.2.1 and 3.2.2.

The FR is measured in a QCD enriched data sample. In contrast to what was done for the HWW fully leptonic analysis, in which the main contribution to fakes was coming from mis-identified light-jets, in this analysis the jets are produced from heavy-quark decays. For this reason, the QCD control region used to measure the FR must be enriched in b-quark decays. To define this region, a similar procedure as in the ttH fully leptonic analysis is followed [32].

To select a $b\bar{b}$ QCD control region as pure as possible, a back-to-back pair of leptons is selected. In general, this kind of events are reconstructed as a pair of jets with very high $\Delta\phi$ separation. For this measurement, only the cases in which the b-quarks decay to a lepton are considered. To select QCD events and measure the FR, a tag-and-probe method is defined, where a *tag* lepton helps to select the event, and a *probe* lepton serves to measure the rate. The *tag* lepton that identifies the event is a highly anti-isolated ($\mathrm{mva_{iso}} < -0.6$) muon with a high displacement from the primary vertex (SIP3D > 7) of the event. The *probe* is selected as a loose lepton, either

a muon or an electron as defined in Section 3.2, so that $\Delta\phi$(tag,probe) is higher than 2.5 rad. To further improve the purity of the sample for QCD events, a p_T balance between leptons is required, by asking the pair to pass the following requirement: $p_{T\ \text{probe}}/[p_{T\ \text{tag}}*(1+\text{relIso}_{\text{tag}})]$. In this condition, the tag momentum is corrected by the relative isolation since the tag lepton is by construction less isolated than the probe, and therefore otherwise carrying a smaller fraction of the b-jet momentum.

QCD enriched data events were selecting using the `MuEG` and `DoubleMu` Primary datasets (PD). These PDs do not bias the measurement, as the HLT paths used (the same as for the standard analysis, see Table 2) have loose or none isolation requirements, less tight than the ones used for the loose lepton definition (in fact, these loose definitions were originally defined to take this into account).

In the control region described above (for both numerator and denominator) there is contribution from other processes: W+jets, Z+jets and top events mainly. To reduce this contribution, no extra lepton in the event is admitted. For the final fake rate estimate a removal of the expected number of events (as estimated from simulation) for the numerator and denominator of Eq. 2 is done.

The results for the fake rate measurement, for both muons and electrons, once the EWK contamination is removed, are shown in Tables 11 and 12. The same information is given on Figures 8 and 9, as a two-dimension graphs.

Muon fake rate				
p_T range (GeV)	$0 < \lvert\eta\rvert \leq 0.500$	$0.500 < \lvert\eta\rvert \leq 1.479$	$1.479 < \lvert\eta\rvert \leq 2.000$	$2.000 < \lvert\eta\rvert \leq 2.500$
$10 < p_T \leq 15$	0.056 ± 0.001	0.065 ± 0.001	0.088 ± 0.002	0.122 ± 0.001
$15 < p_T \leq 20$	0.053 ± 0.002	0.064 ± 0.002	0.102 ± 0.002	0.126 ± 0.002
$20 < p_T \leq 25$	0.080 ± 0.005	0.111 ± 0.002	0.106 ± 0.002	0.127 ± 0.004
$25 < p_T \leq 30$	0.090 ± 0.009	0.120 ± 0.005	0.113 ± 0.007	0.133 ± 0.005
$30 < p_T \leq 35$	0.112 ± 0.015	0.122 ± 0.014	0.132 ± 0.010	0.120 ± 0.009
$35 < p_T \leq 45$	0.114 ± 0.029	0.168 ± 0.023	0.151 ± 0.025	0.148 ± 0.032
$45 < p_T \leq 60$	0.231 ± 0.060	0.199 ± 0.070	0.190 ± 0.045	0.157 ± 0.082

Table 11: Measured **fake rate** in bins of η and p_T of the **Muon**. Statistical errors only.

Electron fake rate				
p_T range (GeV)	$0 < \lvert\eta\rvert \leq 0.500$	$0.500 < \lvert\eta\rvert \leq 1.479$	$1.479 < \lvert\eta\rvert \leq 2.000$	$2.000 < \lvert\eta\rvert \leq 2.500$
$10 < p_T \leq 15$	0.064 ± 0.004	0.062 ± 0.003	0.022 ± 0.002	0.009 ± 0.002
$15 < p_T \leq 20$	0.069 ± 0.012	0.098 ± 0.009	0.029 ± 0.004	0.020 ± 0.003
$20 < p_T \leq 25$	0.070 ± 0.012	0.081 ± 0.004	0.034 ± 0.003	0.018 ± 0.004
$25 < p_T \leq 30$	0.084 ± 0.015	0.091 ± 0.008	0.037 ± 0.004	0.021 ± 0.002
$30 < p_T \leq 35$	0.095 ± 0.024	0.091 ± 0.012	0.046 ± 0.008	0.021 ± 0.002
$35 < p_T \leq 45$	0.104 ± 0.035	0.083 ± 0.019	0.047 ± 0.008	0.022 ± 0.003
$45 < p_T \leq 60$	0.134 ± 0.041	0.077 ± 0.030	0.038 ± 0.010	0.028 ± 0.005
$60 < p_T \leq 300$	0.130 ± 0.071	0.085 ± 0.027	0.027 ± 0.008	0.038 ± 0.010

Table 12: Measured **fake rate** in bins of η and p_T of the **Electron**. Statistical errors only.

In Figures 10 and 11, the FR for both muons and electrons is shown for four different detector regions: $0 < |\eta| \leq 0.5$, $0.5 < |\eta| \leq 1.479$, $1.479 < |\eta| \leq 2.0$ and $2.0 < |\eta| \leq 2.5$, before and after the EWK contribution has been taken into account. The EWK contamination clearly affects the

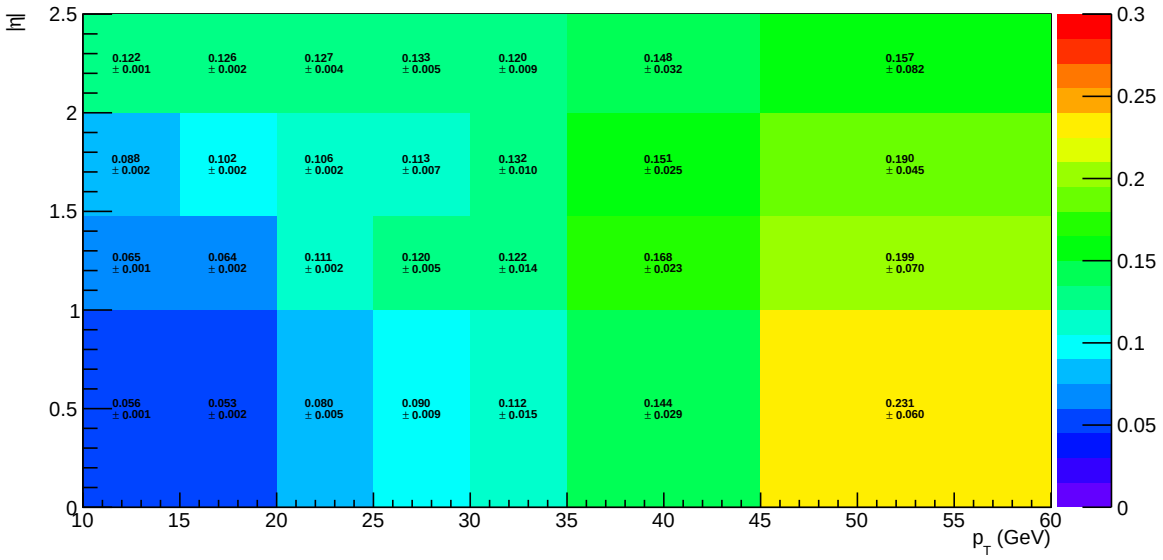

Figure 8: Measured **fake rate** p_{T}-$|\eta|$ map for **muons**

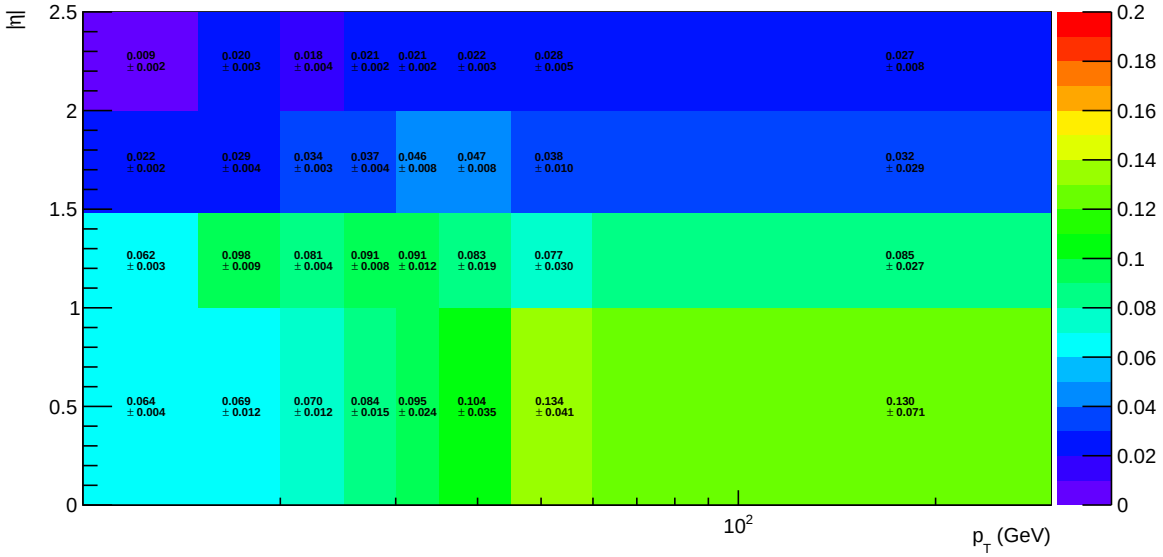

Figure 9: Measured **fake rate** p_{T}-$|\eta|$ map for **electrons**

higher p_{T} bins, thus the necessity of this correction.

5.1.2 Extrapolation from a loose+loose+loose region to the signal region

The FR estimated in the QCD enriched sample is used to weight four different data samples, defined by the number of leptons that pass the tight selection. These events are indicated as N_{t0}, N_{t1}, N_{t2} and N_{t3} when zero, one, two or three leptons fulfill the tight requirements respectively.

There could be some contribution to the loose component coming from prompt leptons eg from W or Z decays which do not pass the tight selection due to inefficiencies. To take into account this contamination, the prompt rate, PR, is measured on data. It is defined as the rate of real prompt leptons that fulfill the loose selection, but fail the tight requirements. The PR is

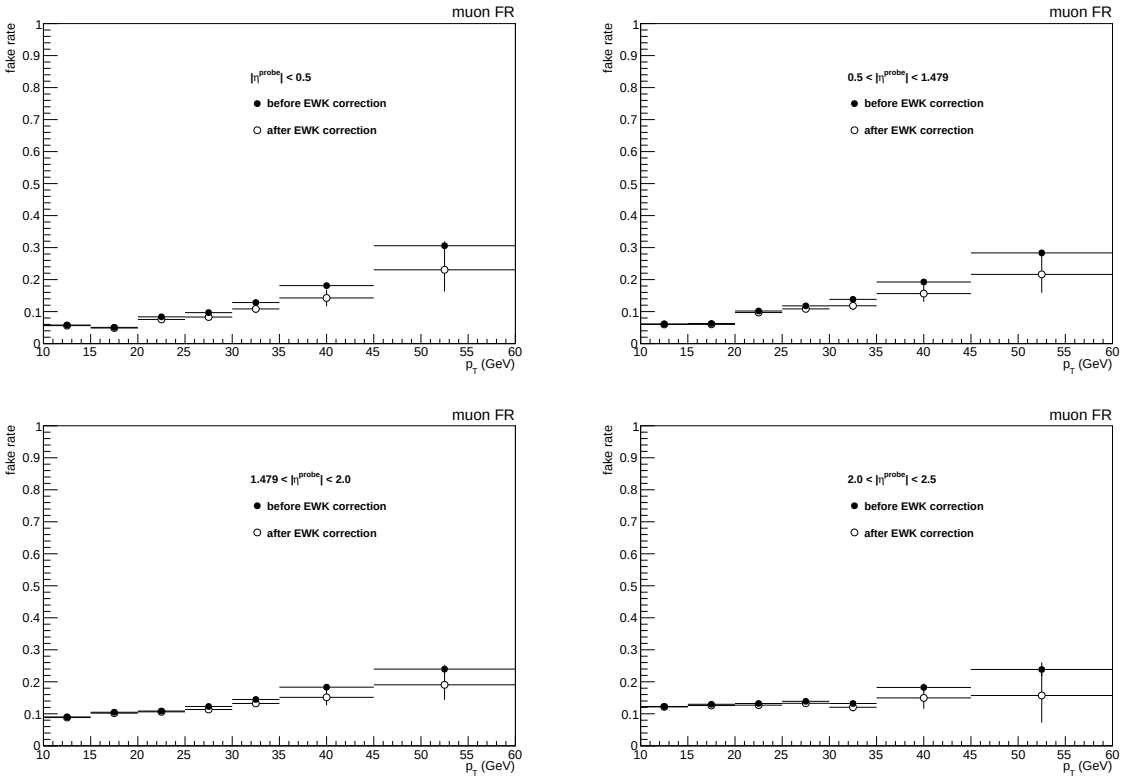

Figure 10: Measured **fake rate** for muons, for different η regions, before and after the corrections for the EWK contamination

estimated using the tand-and-probe method on Z→ll data events. Table 13 and 14 show the prompt rate for muon and electrons, respectively.

Muon prompt rate		
p_T range (GeV)	$0 < \|\eta\| \leq 1.479$	$1.479 < \|\eta\| \leq 2.5$
$10 < p_T \leq 15$	0.702 ± 0.003	0.762 ± 0.002
$15 < p_T \leq 20$	0.799 ± 0.001	0.805 ± 0.001
$20 < p_T \leq 25$	0.902 ± 0.001	0.893 ± 0.001
$25 < p_T$	0.974 ± 0.019	0.959 ± 0.001

Table 13: Measured **prompt rate** in bins of η and p_T of the loose **muon**. Statistical errors only.

Electron prompt rate			
p_T range (GeV)	$0 < \|\eta\| \leq 1.44$	$1.44 < \|\eta\| \leq 1.55$	$1.55 < \|\eta\| \leq 2.5$
$10 < p_T \leq 15$	0.588 ± 0.005	0.55 ± 0.02	0.296 ± 0.005
$15 < p_T \leq 20$	0.715 ± 0.002	0.57 ± 0.01	0.447 ± 0.004
$20 < p_T \leq 25$	0.732 ± 0.001	0.632 ± 0.007	0.629 ± 0.002
$25 < p_T \leq 50$	0.925 ± 0.001	0.850 ± 0.001	0.850 ± 0.001
$50 < p_T$	0.970 ± 0.001	0.941 ± 0.002	0.936 ± 0.001

Table 14: Measured **prompt rate** in bins of η and p_T of the loose electron. Statistical errors only.

If p is the prompt rate and f is the fake rate, N_{t0}, N_{t1}, N_{t2} and N_{t3} can be expressed in terms of the total number of events with three prompt leptons, N_{ppp}; with two prompt leptons and one

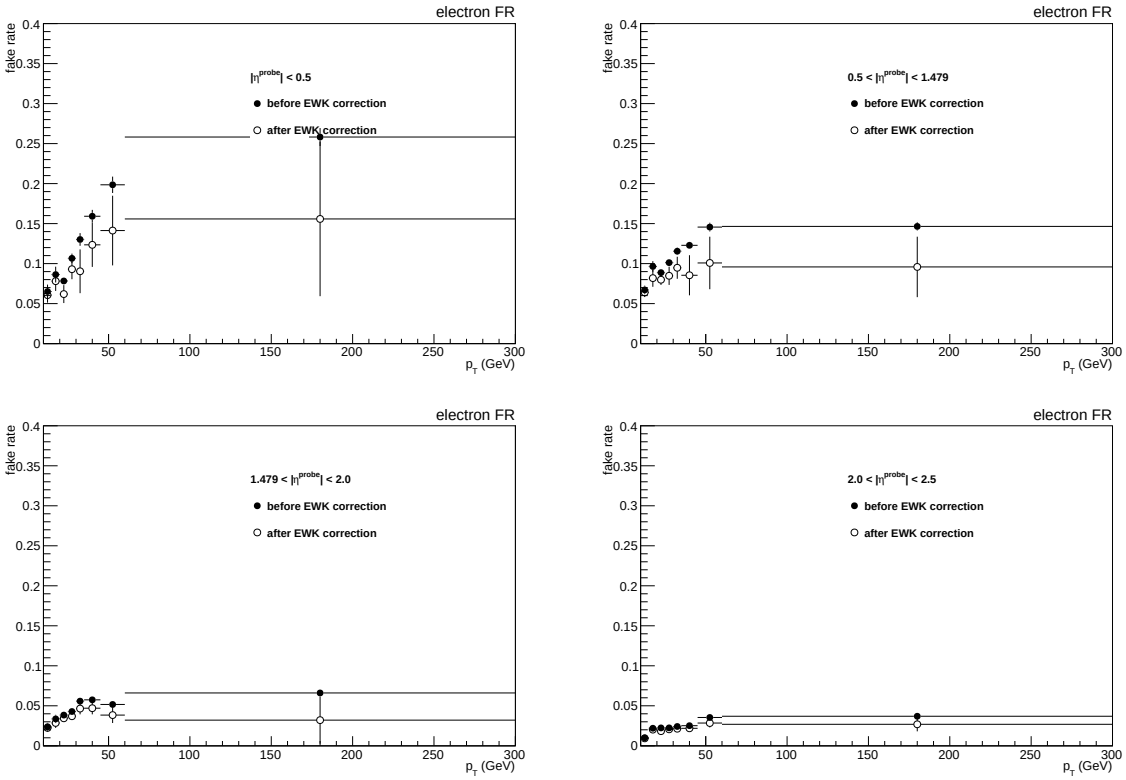

Figure 11: Measured **fake rate** for electrons, for different η regions, before and after the corrections for the EWK contamination

fake lepton, N_{ppf}; with one prompt lepton and two fake leptons, N_{pff}, and finally, with three fake leptons, N_{fff}:

$$
\begin{aligned}
N_{t0} &= (1-p)^3 N_{ppp} + (1-p)^2(1-f)N_{ppf} + (1-p)(1-f)^2 N_{pff} + (1-f)^3 N_{fff} \\
N_{t1} &= 3p(1-p)^2 N_{ppp} + [2p(1-p)(1-f) + f(1-p)^2]N_{ppf} + [2f(1-f)(1-p) \\
&\quad + p(1-f)^2]N_{pff} + 3f(1-f)^2 N_{fff} \\
N_{t2} &= 3p^2(1-p)N_{ppp} + [2pf(1-p) + p^2(1-f)]N_{ppf} + [2pf(1-f) + (1-p)f^2]N_{pff} \\
&\quad + 3f^2(1-f)N_{fff} \\
N_{t3} &= p^3 N_{ppp} + p^2 f N_{ppf} + pf^2 N_{pff} + f^3 N_{fff}
\end{aligned}
$$

For these equations the prompt rate and fake rate for different leptons are assumed to be independent of each other. Inverting the equations the number of events with one, two or three prompt leptons can be derived in terms of f and p:

$$
\begin{aligned}
N_{ppp} &= \frac{1}{(p-f)^3}[(1-f)^3 N_{t3} - f(1-f)^2 N_{t2} + f^2(1-f)N_{t1} - f^3 N_{t0}] \\
N_{ppf} &= \frac{1}{(p-f)^3}[3pf^2 N_{t0} - [f^2(1-p) + 2pf(1-f)]N_{t1} \\
&\quad + [2f(1-p)(1-f) + p(1-f)^2]N_{t2} - 3(1-p)(1-f)^2 N_{t3}] \\
N_{pff} &= \frac{1}{(p-f)^3}[-3p^2 f N_{t0} + [2pf(1-p) + p^2(1-f)]N_{t1} \\
&\quad - [f(1-p)^2 + 2p(1-p)(1-f)]N_{t2} + 3(1-p)^2(1-f)N_{t3}] \\
N_{fff} &= \frac{1}{(p-f)^3}[p^3 N_{t0} - p^2(1-p)N_{t1} + p(1-p)N_{t2} - (1-p)^3 N_{t3}]
\end{aligned}
$$

Failing leptons are weighted by f or p and passing leptons by (1-f) or (1-p). For example, as it is unknown for N_{t0} which lepton was prompt or which one fake, leptons are weighted alternatively with f and p, adding the weights (hence the factor 3).

The general weighting rule for going from (N_{t0},N_{t1},N_{t2},N_{t3}) to (N_{ppp},N_{ppf},N_{pff},N_{fff}), is:

- **ppp events**
  - Tight lepton estimated as prompt lepton: multiply by $p \cdot$ (1-f)
  - Fail lepton estimated as prompt lepton: multiply by $p \cdot f$
- **ppf and pff events**
  - Tight lepton estimated as prompt lepton: multiply by $p \cdot$ (1-f)
  - Tight lepton estimated as fake lepton: multiply by $f \cdot$ (1-p)
  - Fail lepton estimated as prompt lepton: multiply by $p \cdot f$
  - Fail lepton estimated as fake lepton: multiply by $f \cdot p$
- **fff events**
  - Tight lepton estimated as fake lepton: multiply by $f \cdot$ (1-p)
  - Fail lepton estimated as fake lepton: multiply by $f \cdot p$

A common weight of $\frac{1}{(p-f)}$ is applied per lepton.

The final background estimation in the signal region is computed as the sum of $p^2 f N_{ppf}$, $pf^2 N_{pff}$ and $f^3 N_{fff}$. The first is the one that more contributes, being much smaller the last two terms. Intuitively, $p^2 f N_{ppf}$ can be seen as the contribution from $t\bar{t}$ events.

5.1.3 Systematic uncertainties studies

In order to estimate the validity of the method, a closure test is performed using the fake rate estimated from QCD simulation. For completeness, the fake rate from $t\bar{t}$ MC events is also calculated, matching the loose leptons to b-quarks at generation level, and hence, knowing their origin. Figures 12 and 13 show the fake rate for muons and electrons, respectively, estimated from data and from QCD simulation (see Table 3).

In general, the behaviour of the fake rate estimated from data and the one from QCD simulation follow the same pattern, but in the case of QCD simulated events, the fake rate is lower for higher p_T bins.

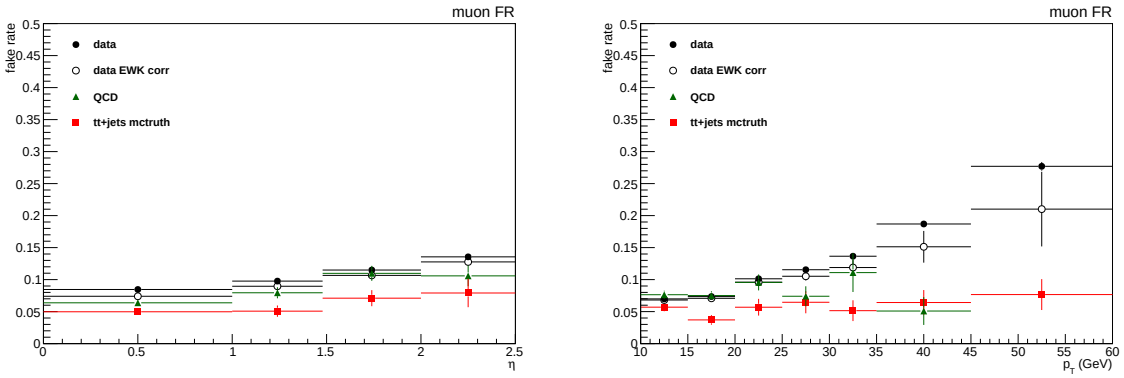

Figure 12: Estimated **fake rate** for muons from QCD events and MC-truth information, compared with measured data fake rate

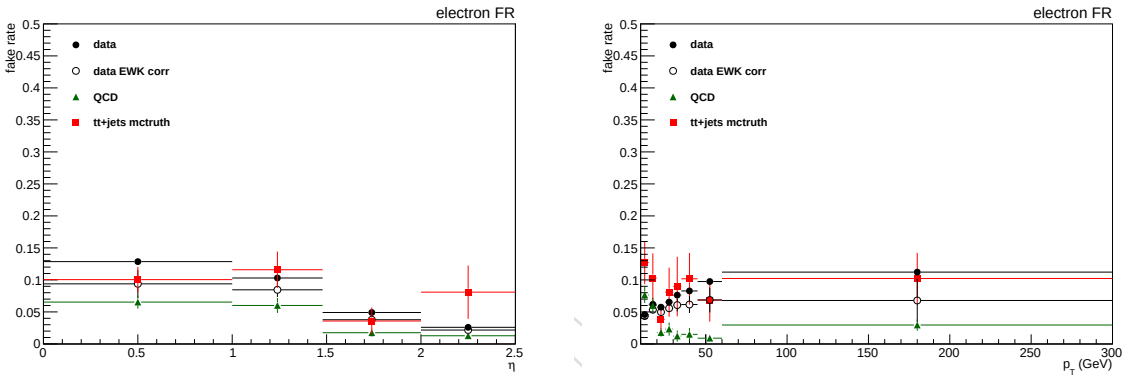

Figure 13: Estimated **fake rate** for electrons from QCD events and MC-truth information, compared with measured data fake rate

A closure test is performed in a $t\bar{t}$ enriched region, comparing the expected number of events for this process using the standard MC sample and applying cuts, with the estimated number of events considering the fake rate method on the same MC sample, but for a *loose +loose+loose* selection. In this latter case, the MC QCD fake rate is used to weight the events. The region where the test is performed is enriched in $t\bar{t}$ events with the following requirements:

- Invariant mass of the event, $m_{ll,OSSF} > 20$ GeV. The invariant mass used for this is the one that, among pairs of leptons with same-flavour and opposite charge (when more than one found, depending on the three lepton final state), has a closest mass to the Z boson mass value.
- Maximum p_T of the three leptons higher than 20 GeV
- No extra tight leptons selected in the event (i.e., only three tight leptons)
- PF MET > 30 GeV
- Number of b-tagged jets equal to one, with $p_{Tjet} > 50$ GeV
- Z-veto: $|m_{ll,OSSF} - m_Z| > 15$ GeV

The results are shown in Table 15. FR $t\bar{t}$ corresponds to the result of the closure test, when applying to a *loose+loose+loose* selection in $t\bar{t}$ MC event the fake rate technique, and MC $t\bar{t}$ is the MC estimate with the above selection for an inclusive $t\bar{t}$ MC sample. The difference is about 30%. This value is taken as a conservative estimation of the systematic error affecting the procedure. The expected number of events and data yield in this control region are summarized

Sample	Yield $\pm$ stat.
FR $t\bar{t}$	38.9 ± 1.1
MC $t\bar{t}$	26 ± 4

Table 15: Results for the fake rate data driven method closure test, for the **three lepton final state**. Statistical errors only

in Table 16. The fake rate data driven method is giving a good estimation of the $t\bar{t}$ contribution. In this region, with a purity of $t\bar{t}$ events higher than 85%, the agreement between the expected background and the observed data events is indeed within 7%.

Process	Yield $\pm$ stat.
Diboson	1.79 ± 0.08
WVV+Drell-Yan	0.2 ± 0.4
ttW	4.8 ± 0.3
ttZ	2.29 ± 0.17
Fakes ($t\bar{t}$)	59 ± 8
Total Background	69.5 ± 7.7
Data	64

Table 16: Data yields and expected background in 19.7 fb^{-1}, for a $t\bar{t}$ control region for the **three leptons final state**. Statistical errors only

In Figure 14 the distribution of some kinematic variables in this control regions is shown.

5.2 Fake Rate Method for Same-Sign Two Leptons Final State

The origin of the fake lepton background in the same-sign dilepton selection is mostly $t\bar{t}$ production, where one (prompt) lepton gets produced in the leptonic decay of one top, and the other gets produced in the hadronization of the b-quark from the decay of the other top, hence resulting in a pair of equal charge.

The estimation of this contribution is done using an equivalent method as described in the previous sections, and its exact implementation is taken over without change from the leptonic ttH analysis [32]. The main difference with respect to the three lepton fake rates described previously lies in the variable used in the loose-to-tight extrapolation. Whereas the three lepton implementation uses the lepton isolation and impact parameters, the two lepton estimation is based on the lepton MVA output described in Sec. 3.1. Note that some of the inputs of the lepton MVA are also isolation and impact parameter based observables.

The tight lepton selection corresponds to what was described in Sec. 3.1, whereas the loose selection requires a cut on the lepton MVA output of > 0.35. As previously, in the coming sections the fake rate is defined as the number of leptons passing the tight cuts, divided by the number of leptons passing the loose cuts, in any given bin:

$$f^i = \frac{N^i_{tight}}{N^i_{loose}},$$

i.e. the numerator is a subset of the denominator and the ratio is constrained to below unity. It can be directly interpreted as the probability of a loose lepton to pass the tight cuts.

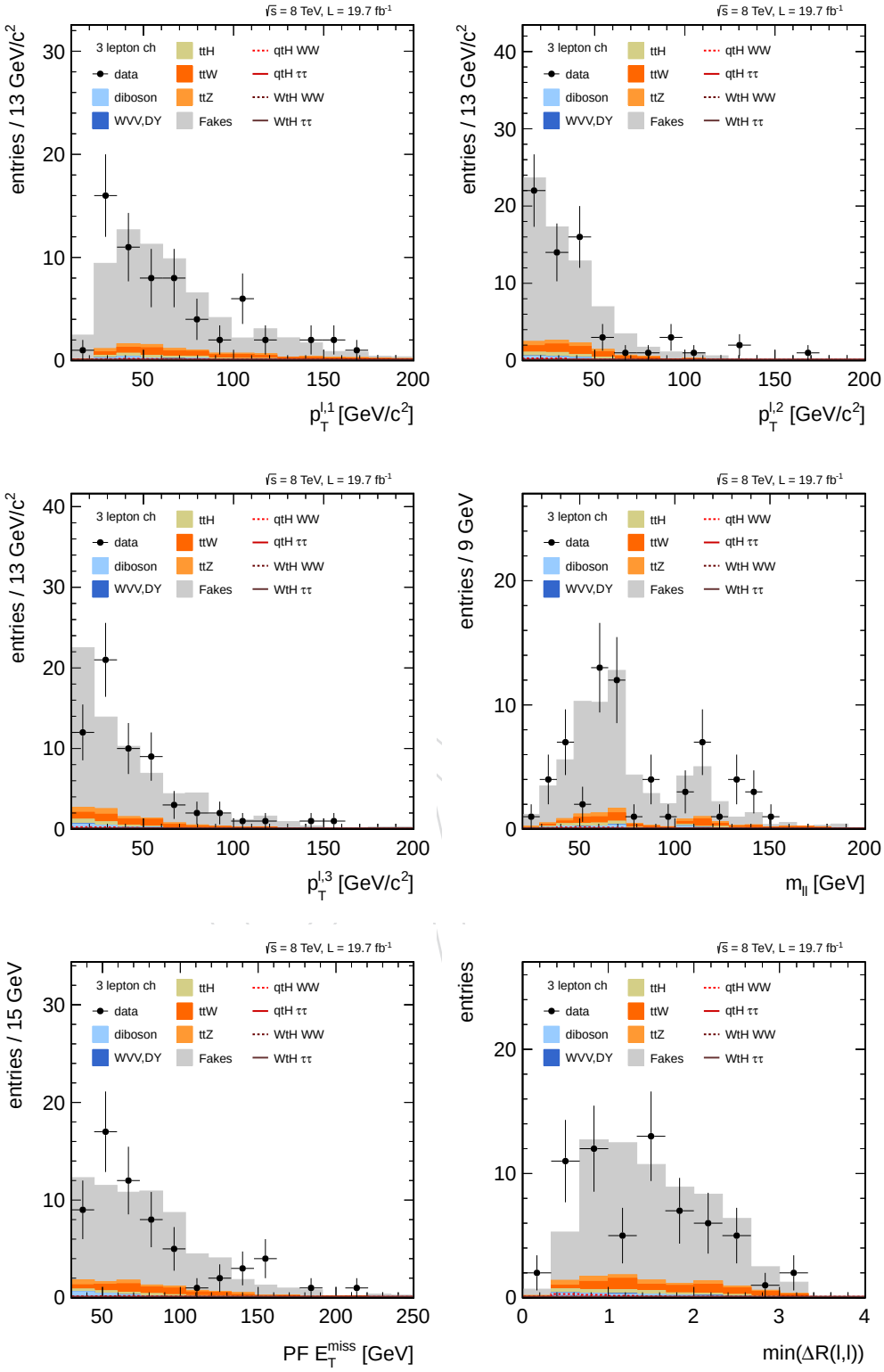

Figure 14: Kinematic variables distribution in the $t\bar{t}$ enriched region for the **three leptons final state**. From top to bottom, left to right: p_{T} of the three selected leptons, reconstructed invariant mass, PF MET and minimum ΔR between leptons

5.2.1 Fake Lepton Control Regions

The main ingredient in the fake rate method is the extraction of the ratio from an independent control sample, dominated by fake leptons. The base assumption is then that the ratio is the same in the control region and the signal region for which several effects have to be taken into account:

- Kinematic differences of the mother jets;
- A possible difference in fake composition (heavy flavor decays *vs.* light jet mis-reconstructions);
- Contamination of the control region by prompt leptons, driving up the ratio and therefore the prediction.

The extracted fake rate is binned in transverse momentum and pseudo-rapidity to account for the kinematic differences between regions.

Three different control selections are studied and then merged to extract the final ratios:

- QCD dilepton events selected using a non-isolated tag lepton with high impact parameter. To suppress contributions from electroweak processes with prompt leptons events are vetoed if they contain additional loose leptons, are not back-to-back ($|\Delta\phi| > 2.5$), or are not balanced in transverse momentum in favor of the probe lepton.
- QCD $b\bar{b}$ events selected using a tightly b-tagged jet as the tag, with event vetoes to suppress prompt leptons equivalent to the above.
- $Z \to \ell^+\ell^-$ with a third lepton, selected using two tight leptons with an invariant mass within 10 GeV of the Z mass and with additional cuts on $E_{\mathrm{T}}^{\mathrm{miss}}$, m_T, and additional leptons to suppress di-boson backgrounds.

5.2.2 Correction for Prompt Lepton Contamination

Each of the three control regions employs cuts to suppress the contribution from processes with prompt leptons, e.g. Z +jets events. A residual contamination from prompt leptons would bias the fake ratio towards higher values (since prompt leptons have a higher probability of passing the tight criteria) and therefore equally bias the fake background prediction.

The remaining prompt contamination is therefore estimated and subtracted from each control region. For the two QCD regions, the fake rate is measured separately for small ($[0-20]$ GeV) and large ($[45-80]$ GeV) values of $E_{\mathrm{T}}^{\mathrm{miss}}$. Both the ratios for fake leptons and for prompt leptons do not depend on $E_{\mathrm{T}}^{\mathrm{miss}}$, and only a difference in their relative contributions will result in different values of the measured rate. With this assumption and using simulated events to estimate the relative yields of events with prompt leptons in the two $E_{\mathrm{T}}^{\mathrm{miss}}$ regions it is possible to compute the tight/loose ratio for leptons from QCD alone:

$$f_{\mathrm{QCD}} = \frac{f_{\mathrm{S}} - r_{\mathrm{P}}^{\mathrm{SL}} f_{\mathrm{L}}}{1 - r_{\mathrm{P}}^{\mathrm{SL}}},$$

where f_{S} and f_{L} are the measured ratios for small (S) and large (L) values of $E_{\mathrm{T}}^{\mathrm{miss}}$, and $r_{\mathrm{P}}^{\mathrm{SL}}$ is the ratio of relative contributions of prompt leptons to the two bins. The full derivation of this formalism is given in [32].

5.2.3 Measured Fake Rates

The extracted fake rates for the different control regions separately can be found in the appendix of [32] and are not reported here. In general the measured rates are compatible with the expectation from non-prompt leptons in simulated $t\bar{t}$ events within their statistical uncertainties and an additional systematic uncertainty of $\pm 40\%$.

In order to reduce the sizable uncertainties of the individual measurements, the data from the three different control regions is combined, weighted by the inverse square of the uncertainties. The expected distributions from simulation are combined as well, using the same weights as the data. Combining the measurements improves the agreement both between data and simulation and between data and the true distribution from $t\bar{t}$ fakes. The resulting merged fake rate distributions for electrons and muons, in bins of p_T and η can be found in Fig. 15.

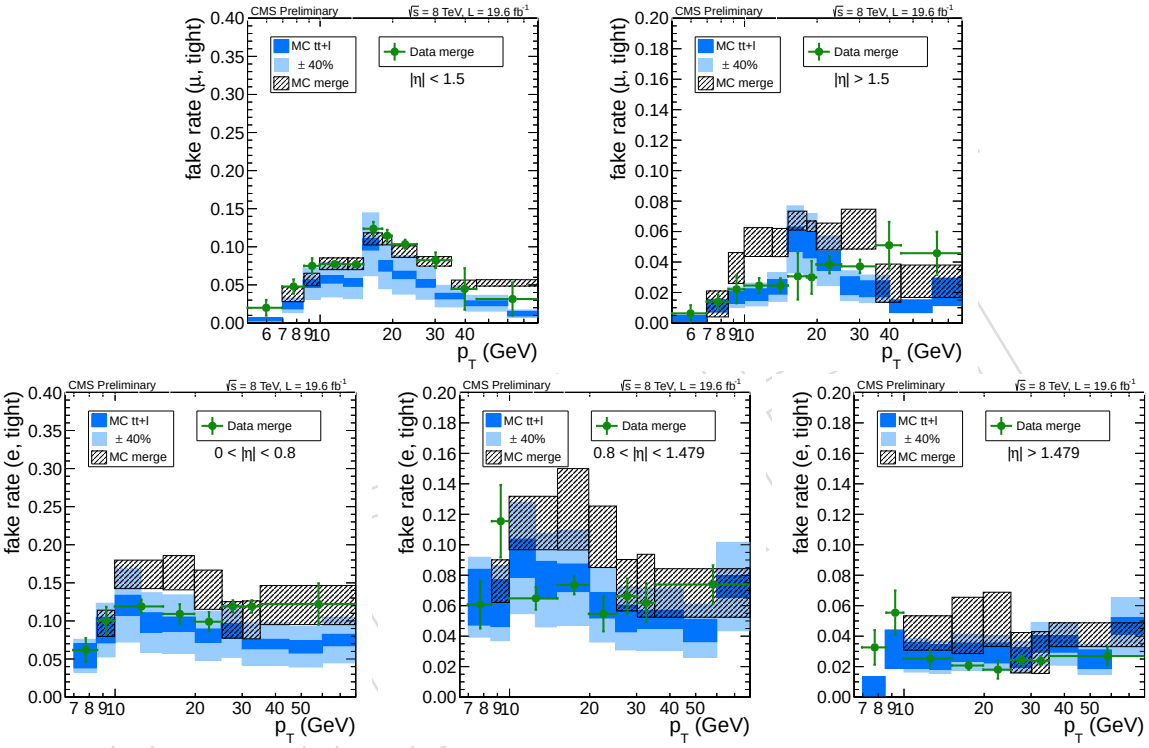

Figure 15: Measured fake rates for the tight lepton definition used in this analysis (MVA > 0.7) and for a loose definition with MVA > 0.35, after combining the different control regions. Combined values from data (green dots with error bars) are compared with the combination of the expectations from simulated events in control regions (black hatched boxes) and fake leptons in $t\bar{t}$ events in the signal region (blue filled boxes). The results in the top row are for muons, the ones at the bottom for electrons; the results in the different columns corresponds to different $|\eta$ ranges, increasing from left to right, as written in the plots. Data uncertainties for different p_T bins are not uncorrelated, and neither are the uncertainties on the values from control regions in simulations.

5.2.4 Fake Rate Application

The basic formalism of predicting event yields from loose sidebands to the signal region and tight-to-loose ratios has been described in the previous section and does not differ in the same-sign dilepton case. We neglect terms proportional to the probability of a prompt lepton failing the tight requirement after passing the loose (i.e. proportional to $(1 - \epsilon)$) such that the final

formula predicting the fake background reads:

$$N_{pp}^{\text{bkg}} = \frac{f_1}{1-f_1} N_{pf} \; + \; \frac{f_2}{1-f_2} N_{fp} \; - \; \frac{f_1 f_2}{(1-f_1)(1-f_2)} N_{ff}$$

where f_1 and f_2 are the corresponding fake rates for the first and second lepton in the event, N_{pf}, and N_{fp} [1] are the number of events with one lepton passing and one failing, once the first one passing and the second failing and once vice-versa, and N_{ff} is the number of events where both leptons fail the tight cut. We thus obtain an entirely data-driven prediction for the number of fake leptons from the tight/loose ratio – determined in control regions in data – and the number of events in the signal region where one or both leptons fail the tight cut.

5.2.5 Validation and Systematics

The systematic uncertainty of the fake prediction is estimated following the studies done in the original analysis [32]. They are briefly summarized here.

The measurement of the fake ratios themselves have an associated systematic uncertainty of about 40%, estimated from the level of agreement between control regions in simulation and the expected "true" fake rate from non-prompt leptons in $\bar{\text{t}}\text{t}$ MC.

The effect of this uncertainty on the overall normalization of the fake background prediction is estimated by scaling the rate within $\pm$40% and propagating the effect to the predictions. This overall rate uncertainty is assumed to be independent for electrons and muons, but otherwise correlated across the channels.

Furthermore, the effect of a possible systematic variation of the fake rates in different lepton kinematic ranges on the shape of final discriminating distributions is estimated. The fake rate is varied separately and oppositely in bins of p_{T} and η, and each variation is propagated to the prediction. The envelope for these variations is then taken as an uncertainty band for the shape of the fake background prediction. The considered variations are to shift the rate *up* by 40% for leptons with $p_{\text{T}} < 30$ GeV while simultaneously shifting it *down* by an equivalent amount for leptons with $p_{\text{T}} > 30$ GeV. Analogously, the rate is shifted separately for leptons in barrel and endcaps.

Generally, the overall uncertainty on the fake backgrounds is dominated by the overall normalization uncertainty, with the shape uncertainty band staying within about 10 to 20%.

A closure test on simulated events has been performed for the original analysis, finding generally a fair agreement between predictions from the data-driven fake rate method and the expected fake lepton yield in MC simulation.

5.3 Background from Charge Mis-Assignment

The same-sign dilepton final state is vulnerable to a contribution from processes yielding opposite-sign dilepton pairs ($\bar{\text{t}}\text{t}$, Z +jets) where one of the two legs has its charge wrongly assigned. This charge mis-identification probability is found to be negligible for muons in simulation, but is appreciable for electrons due to the possibility of strongly asymmetric conversions of hard bremsstrahlung. Predicting the background from charge mis-assignment is however straightforward. The mis-assignment probability can be extracted in the data from a sample where two opposite-sign electrons from a Z boson decay are reconstructed as having the same charge yet show a peak at the Z boson mass.

[1] Note that the repeating N_{pf} in [32] (v12) is a typo.

For electrons in the barrel, the charge mis-assignment probability is found to be about 0.03%, with no p_T dependence (Fig. 16, left). In the endcaps the probability increases with p_T from about 0.08% for $p_T < 20$ to about 0.28% for $p_T > 50$ GeV (Fig. 16, left). An excellent agreement between data and simulations is observed in the barrel, while in the endcaps the simulation appears to underestimate the charge mis-assignment probability by 10–20%.

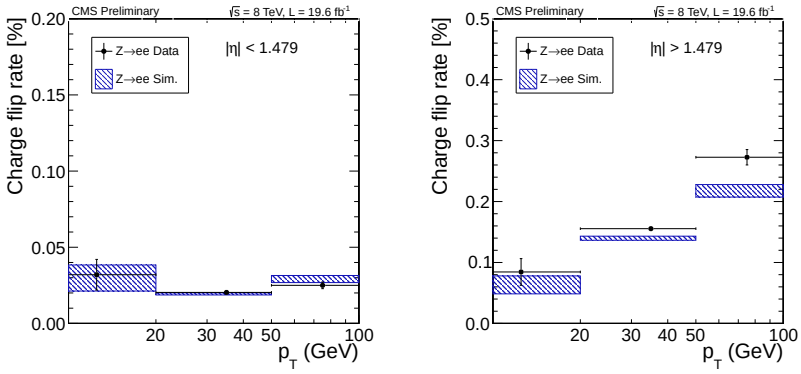

Figure 16: Measured charge mis-assignment probabilities in data for electrons in the barrel (left) and endcaps (right) as function of the electron p_T. The results from performing the same analysis on simulated $Z \to ee$ events is also shown. [32]

The contribution to the same-sign signal regions is then estimated by using an equivalent opposite-sign selection and extrapolating using the measured charge mis-assignment probability.

6 Shape Analysis

The signal extraction is performed using a multivariate likelihood discriminator to suppress the backgrounds. The production cross section for the signal is small, just a few fb, so the signal to background ratio, for a relatively loose selection, is fairly small (see for example Appendix A). A simple cut and count analysis is not enough to reach appreciable sensitivity to set limits on the signal production. Hence a multivariate method is used to build a discriminator to further reduce background events. We build a simple linear likelihood discriminator as the ratio of signal over signal + background likelihoods for a chosen set of discriminating observables:

$$L(x) = \frac{L_S(x)}{(L_S(x) + L_B(x))}. \tag{3}$$

For each event the signal and background likelihoods are calculated as the product of the respective signal and background probability density functions (PDFs), evaluated at the observed values:

$$L^i(x) = \prod_j PDF^i_j(x^j),$$

where the PDF's are normalized to unity.

The variables to enter the likelihood are chosen to be minimally correlated while providing good discrimination power.

6.1 Three Leptons Channel

6.1.1 Three Leptons Pre-Selection

The three lepton final state is dominated by the $t\bar{t}$ background contribution, so a search of good discriminant variables against this process is done. A tight pre-selection is applied on the events, in order to minimize the contribution from this background. The following requierements are used to select events for the final likelihood estimation:

- Exactly three tight leptons with $p_T > 20/10/10$ GeV
- Invariant mass of the event, $m_{ll,OSSF} > 20\,\text{GeV}$. The invariant mass used for this is the one that, among pairs of leptons with same-flavour and opposite charge (when more than one found, depending on the three lepton final state), has a closest mass to the Z boson mass value.
- Z-veto: $|m_{ll,OSSF} - m_Z| > 15\,\text{GeV}$
- $E_T^{miss} > 30\,\text{GeV}$
- Exactly one jet tagged as CSV medium
- At least one forward jet ($|\eta| > 1.5$)

After this selection, the expected number of signal and background events is summarized in Table 17.

Process	$\ell\ell\ell$
WZ/ZZ/WW	1.18 ± 0.06
WVV+Drell-Yan	0.11 ± 0.03
$t\bar{t}Z$	2.20 ± 0.18
$t\bar{t}W^{\pm}$	3.0 ± 0.3
$t\bar{t}H$	1.52 ± 0.06
Fakes	34 ± 6
Total Background	$\mathbf{41.9 \pm 5.8}$
$tH(\tau\tau)W$	0.07 ± 0.01
$tH(WW)W$	0.19 ± 0.01
$tH(\tau\tau)q$	0.31 ± 0.01
$tH(WW)q$	0.95 ± 0.02
Total Signal	$\mathbf{1.51 \pm 0.03}$
Data	**42**

Table 17: Data yields and expected number of signal and background events for the likelihood pre-selection for the **3 lepton final state**, in 19.7 fb^{-1}. Statistical errors only

6.1.2 Input Variables and Linear Likelihood Discriminant

Several variables were checked, resulting on a set of 5 variables that better discriminate the qtH production. These variables are:

- Number of central jets (as defined in Table 6)
- Number of forward jets with $|\eta| > 2.4$
- Total sum of the charges of the three leptons
- Minimum value of ΔR between the leptons in the event
- $\Delta\eta$ value between the b-tagged jet and the forward jet with highest η value

The correlation matrices for the qtH (H→WW) signal and the main background ($t\bar{t}$) are checked. For the latter, the fully leptonic MC sample is used. Small correlation between the variables is observed, as shown in Figure 17.

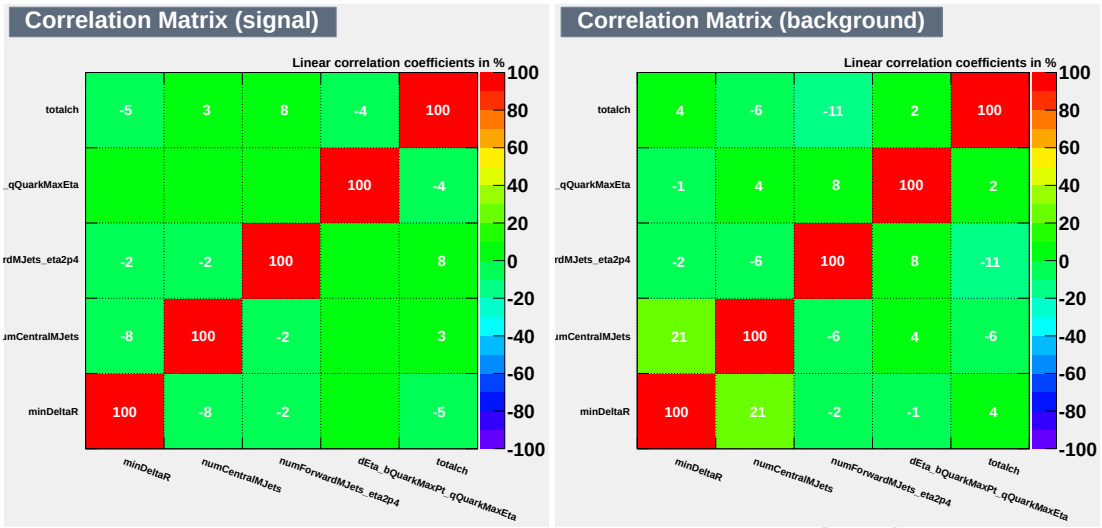

Figure 17: Correlation matrix for the variables used as inputs to the likelihood, for signal (left) and main background (right), for the **3 leptons final state**

These variables are used as PDFs for the final likelihood estimation. The qtH (H→WW) sample is used as signal and the weighted sum of fakes ($t\bar{t}$), $t\bar{t}W^{\pm}$ and $t\bar{t}Z$ is used as background. Adding qtH ($H \rightarrow \tau^{+}\tau^{-}$) or WtH production to the signal PDFs does not give any significant improvement.

Figures 18 show the comparison of the shapes for $t\bar{t}$ background when using the fake rate data driven method and a inclusive $t\bar{t}$ MC sample. For the latter, the statistics are pretty small, due to the size of the sample and the tight cuts. However, it can be seen that, for almost all the variables and within the statistical errors, the agreement between both sources is quite good. For this reason, not only the normalization from the data driven method is used for the estimation of the $t\bar{t}$ expected yield, but also the shapes are used for the likelihood discriminator construction.

In Figures 19 the likelihood PDFs are shown, normalized to unity. The same distributions, comparing all the processes and normalized to 19.7 fb^{-1}, are shown in Figure 20

The comparisons between the likelihood output for the main signal process qtH (H→WW) and each of the main backgrounds are shown in Figure 21 normalized to unity. In general the separation between signal and backgrounds is significant, with the signal distribution peaking at high likelihood values (close to 1) while for backgrounds the majority of events concentrate at low likelihood values (close to 0). The final likelihood comparison between data and simulation prediction is shown in Figure 22.

The discriminator gives a good separation between the main signal processes, the qtH ones, and the backgrounds. For a likelihood discriminator value higher than 0.5, the fraction of qtH (H→WW) and qtH ($H \rightarrow \tau\tau$) is nearly 79%, while only 33% of background events fullfill this requirement.

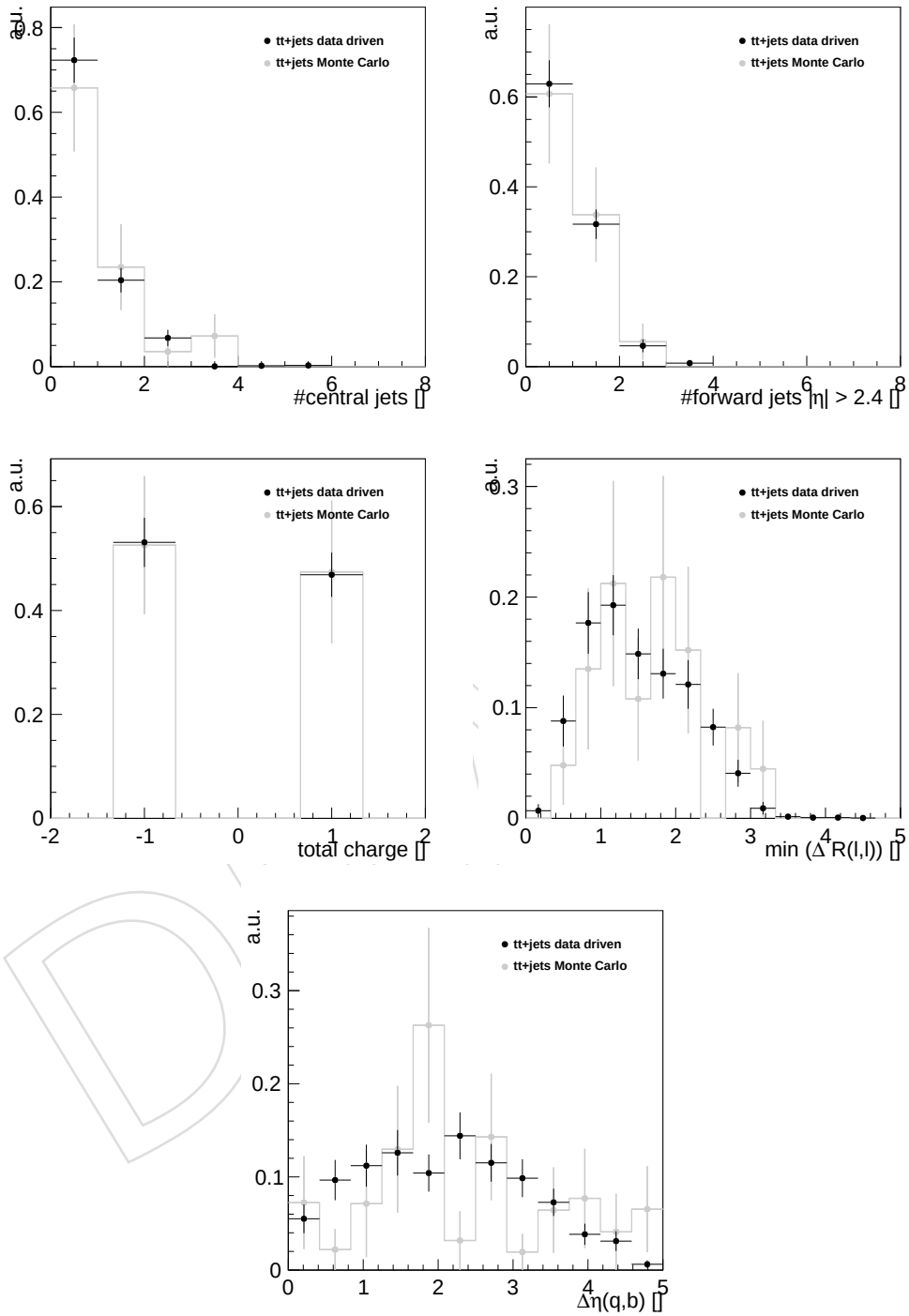

Figure 18: Comparison between the PDFs used for the likelihood discriminator for the **3 leptons final state**, comparing the $t\bar{t}$ background as estimated from the data driven method, or using a inclusive MC sample. From top to bottom, left to right: Number of central jets, number of forward jets with $|\eta| > 2.4$, total sum of the lepton charges, min (ΔR) between leptons, and $\Delta\eta$ between b-tagged jet and most forward jet. The distributions are normalized to unity

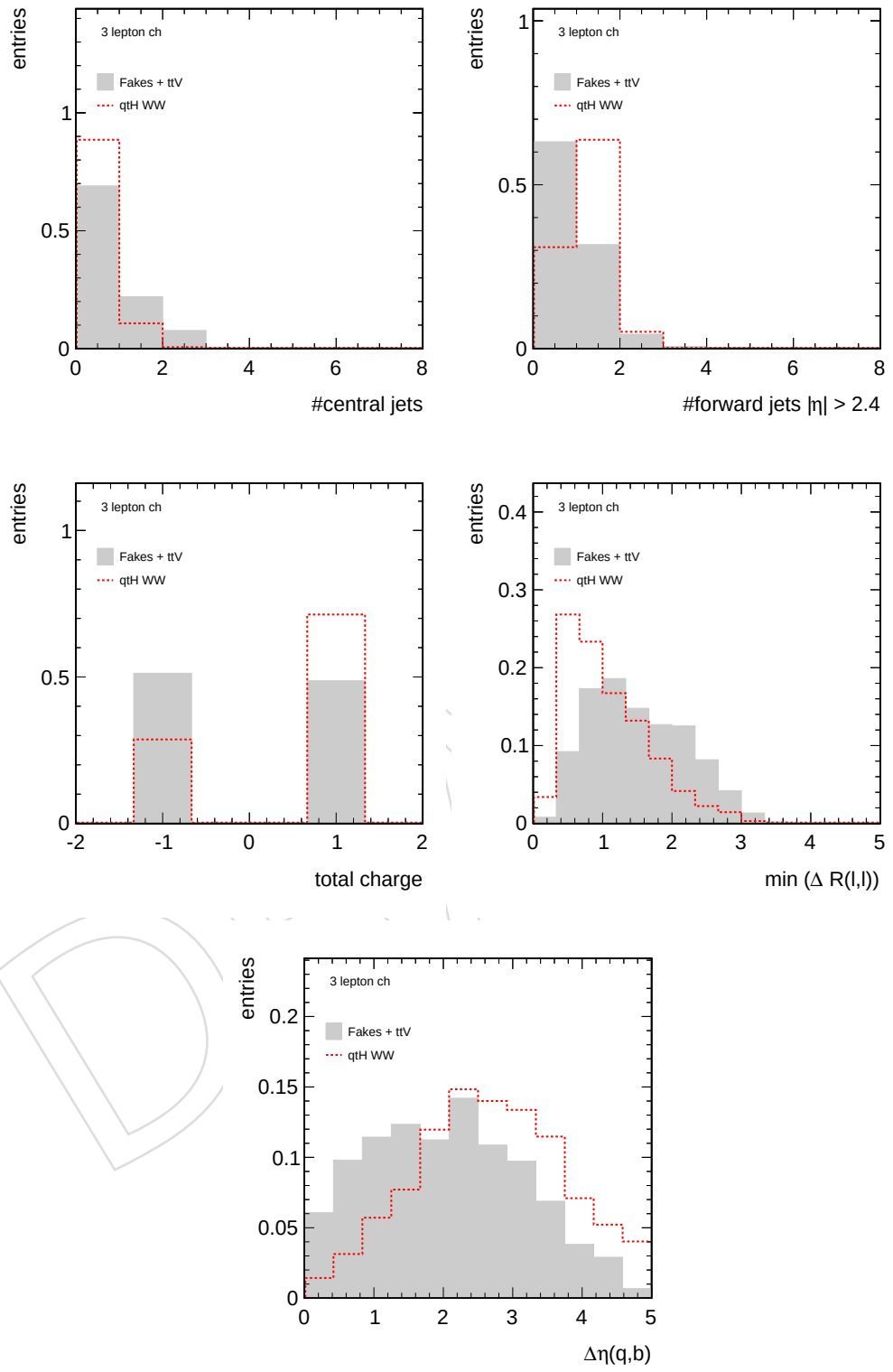

Figure 19: PDFs used for the likelihood discriminator for the **3 leptons final state**. From top to bottom, left to right: Number of central jets, number of forward jets with $|\eta| > 2.4$, total sum of the lepton charges, min (ΔR) between leptons, and $\Delta\eta$ between b-tagged jet and most forward jet. The distributions are normalized to unity

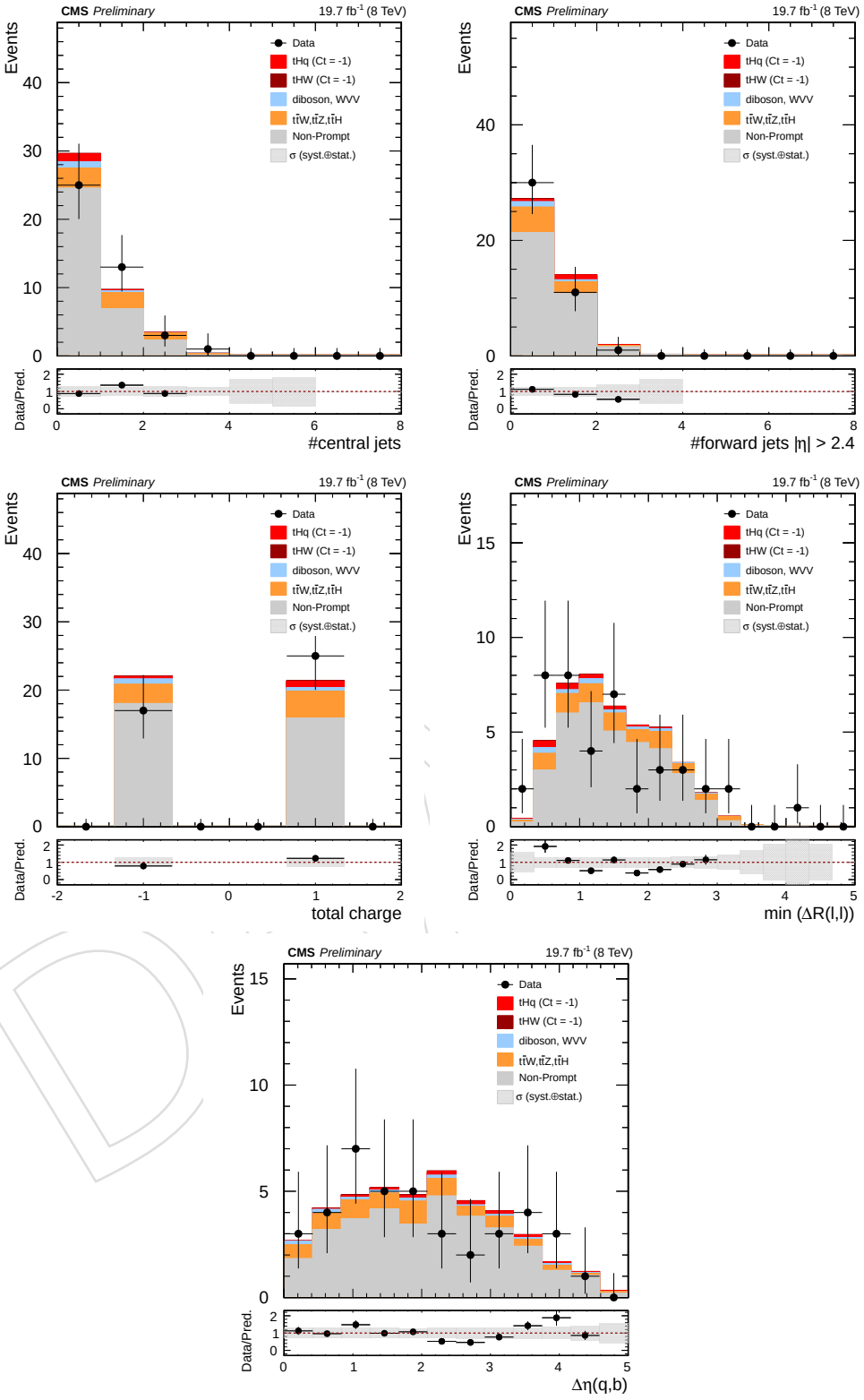

Figure 20: Discriminating variables distributions normalized to 19.7 fb^{-1} for the **3 leptons final state**. From top to bottom, left to right: Number of central jets, number of forward jets with $|\eta| > 2.4$, total sum of the lepton charges, min (ΔR) between leptons, and $\Delta\eta$ between b-tagged jet and most forward jet.

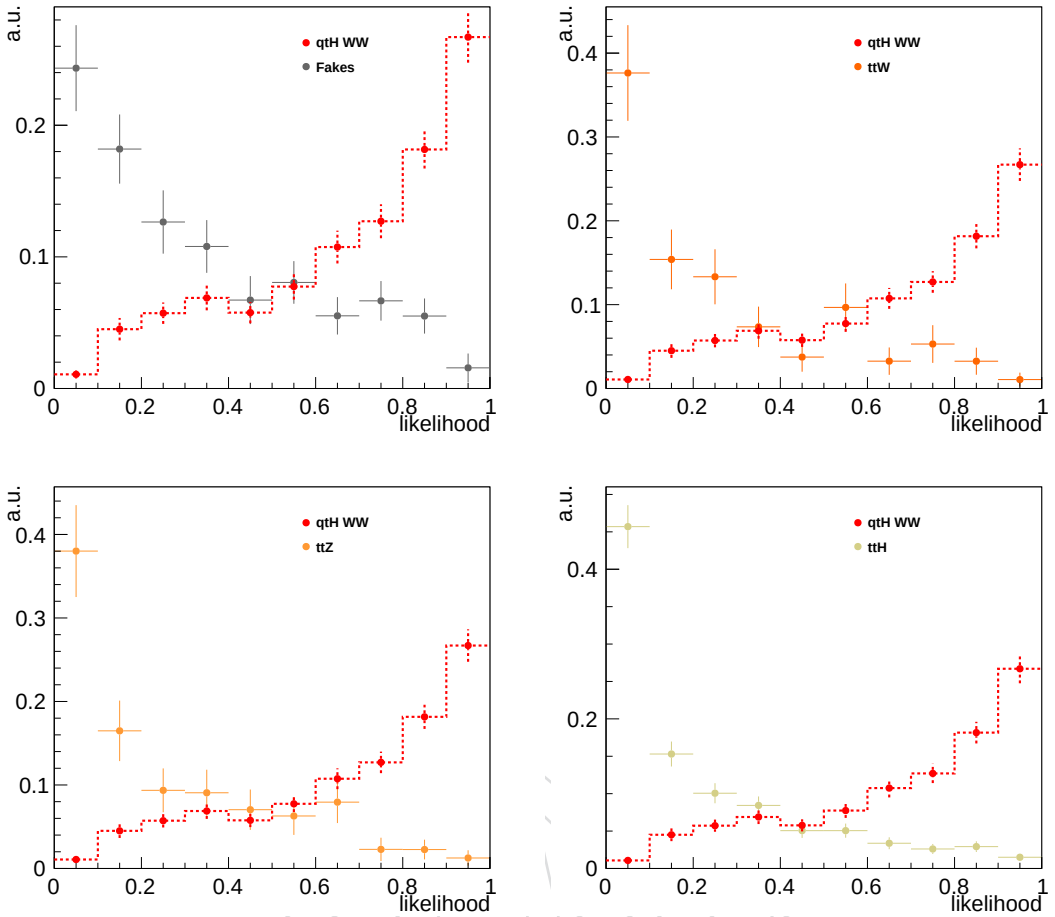

Figure 21: Likelihood discriminator comparison for signal events (qtH (H→WW)) vs the main backgrounds, for the **3 lepton final state**, from top to bottom, left to right: ($t\bar{t}$), ttW, ttZ, and ttH

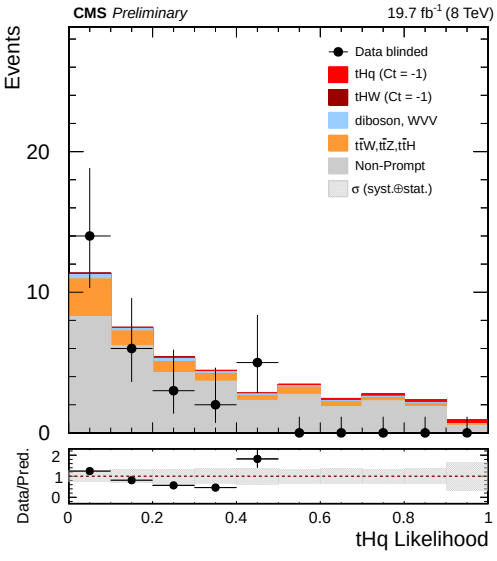

Figure 22: Likelihood discriminator output, for the **3 leptons final state**, for 19.7 fb^{-1}

6.2 Same-Sign Dilepton Channel

6.2.1 Same-Sign Dilepton Pre-Selection

The pre-selection criteria for the same-sign dilepton channel is defined as follows:

- Two leptons of equal charge, $p_T > 20\,\text{GeV}$
- No additional leptons (selected with lepton MVA > 0.35)
- Invariant mass of lepton pairs $> 12\,\text{GeV}$
- At least one central jet ($|\eta| < 1.0$)
- At least one forward jet ($|\eta| > 1.0$)
- At least one central jet tagged as CSV loose

It targets the final state of tHq, where the top quark decays leptonically, and the W boson from the $H \to W^+W^-$ decay with the same charge decays leptonically as well. The opposite charged W boson from the Higgs decay is expected to decay hadronically, producing two light jets. We therefore expect a signature of a lepton and b-jet balancing another lepton of the same charge and two light quark jets, with another light jet at high rapidity, see Fig. 6.

Applying this selection, and performing the background estimation methods described earlier, we obtain the yields shown in Tab. 18 below and in Fig. 23. As can be seen, the selection is very much dominated by the backgrounds, with fakes (and charge mis-identification) contributing 50% or more of the overall rate. The second most important background is from $t\bar{t}W^{\pm}$, followed by WZ and W^+W^+qq. The signal yield is split into contributions from tHq and tHW, and into the two different Higgs decay channels contributing: $H \to W^+W^-$ and $H \to \tau^+\tau^-$. A possible contribution from $H \to ZZ$ is negligible in this signature due to a veto on additional leptons. The combined signal yield is only a handful events for the three channels, giving a signal-to-background ratio of between 2.7% (ee) and 5.1% ($\mu\mu$).

The combined predicted yield and the observed number of events in data show an excess of about $20 - 25\%$ of data in the $\mu\mu$ and $e\mu$ channels, compatible within the uncertainties of the background predictions. Some additional cross-checks are being carried out and are being documented in appendix C (page 72).

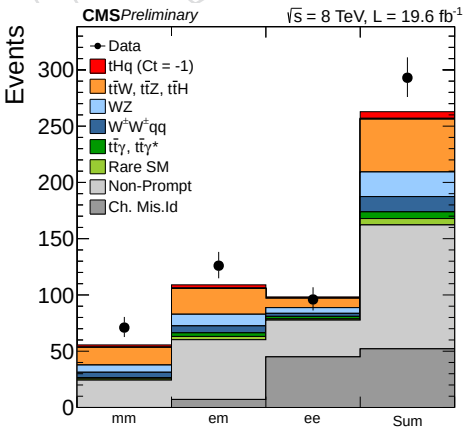

Figure 23: Predicted background and observed data yields in the three same-sign dilepton channels.

Process	$\mu\mu$	$e\mu$	ee
W^+W^+qq	4.72 ± 0.48	6.18 ± 0.53	2.68 ± 0.33
WZ	6.45 ± 0.27	10.33 ± 0.33	5.08 ± 0.23
$t\bar{t}\gamma^*$	0.52 ± 0.05	1.06 ± 0.08	0.61 ± 0.06
$t\bar{t}\gamma$	0.09 ± 0.05	2.15 ± 0.24	1.53 ± 0.21
$t\bar{t}Z$	2.36 ± 0.22	3.27 ± 0.25	1.48 ± 0.17
$t\bar{t}W^\pm$	10.78 ± 0.52	15.96 ± 0.60	5.56 ± 0.35
$t\bar{t}H$	2.53 ± 0.08	3.67 ± 0.10	1.32 ± 0.06
ZZ (Rare SM)	0.26 ± 0.03	0.72 ± 0.05	0.35 ± 0.03
VVV (Rare SM)	1.13 ± 0.09	1.86 ± 0.11	0.69 ± 0.07
$t\bar{t}WW$ (Rare SM)	0.22 ± 0.01	0.31 ± 0.01	0.13 ± 0.00
W^+W^+(DPI) (Rare SM)	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00
(Total rare SM bkg.)	1.61 ± 0.10	2.89 ± 0.12	1.17 ± 0.08
Charge Mis-ID	–	7.14 ± 0.10	44.97 ± 0.31
Non-Prompt	24.32 ± 1.50	53.09 ± 2.51	32.68 ± 2.03
Total Background	$\mathbf{53.39 \pm 1.70}$	$\mathbf{105.73 \pm 2.69}$	$\mathbf{97.07 \pm 2.14}$
$tH(\tau\tau)W$	0.07 ± 0.01	0.10 ± 0.01	0.03 ± 0.01
tH(WW)W	0.17 ± 0.02	0.27 ± 0.02	0.09 ± 0.01
$tH(\tau\tau)q$	0.49 ± 0.02	0.73 ± 0.03	0.29 ± 0.02
tH(WW)q	1.49 ± 0.04	2.15 ± 0.04	0.76 ± 0.02
Total Signal	$\mathbf{2.21 \pm 0.05}$	$\mathbf{3.25 \pm 0.05}$	$\mathbf{1.17 \pm 0.03}$
Data	**71**	**126**	**96**

Table 18: Data yields and expected background for the likelihood pre-selection for the **same-sign di-lepton final state**, in 19.7 fb^{-1}. Statistical errors only.

6.2.2 Input Variables

To extract a possible signal we build a multivariate analysis (MVA) on the pre-selection level and fit its discriminator output shapes to the data. The discriminating variables can be put into three categories:

- forward activity;
- jet and b-jet multiplicity;
- lepton kinematics and lepton charge.

None of the main background is expected to produce any forward activity, whereas the light quark jet produced in association with the top and Higgs of our tHq signal has a high chance of being produced at large rapidity. Diboson backgrounds generally are not expected to contain b-jets, whereas $t\bar{t} + X$ processes generally contain at least two hard b-quarks in the final state. Our signal contains exactly one hard b-quark from the top decay. Fake lepton backgrounds from $t\bar{t}$ do contain two b-quarks, however one of the two tagged jets is absorbed when producing the second isolated lepton; hence $t\bar{t}$ is expected to contribute one b-tagged jet only. Furthermore the two leptons in our signal are expected to be more back-to-back like compared to the main backgrounds which are more evenly distributed. The p_T of the trailing lepton will generally be harder in our signal, where it originates from the decay of an on-shell W boson, compared to fake leptons produced in the hadronization of a b-jet. Finally, our signal shows a strong asymmetry in lepton charge due to the prevalent charge in the incoming protons. The same is true for $t\bar{t} + X$ processes and WZ, but $t\bar{t}$ is a neutral state and will therefore produce leptons of each charge with the same probability.

We considered 12 variables to be included in the MVA, of which 4 were excluded due to strong correlations. They are ranked according to their separation power using the TMVA package of ROOT, and their correlations are computed using the same.

Rank	Variable	Separation Power
1	Jet multiplicity	0.1441
2	$\|\eta\|$ of hardest jet with $\|\eta\| > 1.0$	0.1433
3	η gap between most forward jet and next jet or lepton	0.1411
4	Scalar sum of jet p_{T}'s (H_{T})	0.1267
5	Medium b-tagged jet multiplicity	0.0416
6	Lepton ϕ difference ($\Delta\phi_{\ell\ell}$)	0.0254
7	Trailing lepton p_{T}	0.0101
8	Lepton charge	0.0033

Table 19: List of variables used for the signal MVA and their ranking according to TMVA.

The remaining four variables and their separation power are: the maximum $|\eta|$ of any jet (0.1107); the difference in η between the forward jet and the b-tagged jet (0.0991); the number of jets with $|\eta| > 2.0$ (0.0846); and the central jet multiplicity ($|\eta| < 1.0$, 0.0406).

The correlations matrices for signal and background of all 12 variables considered are shown in Fig. 24. As expected, all variables related to the forward jet show strong correlations, whereas lepton kinematics and b-jet multiplicity are almost uncorrelated.

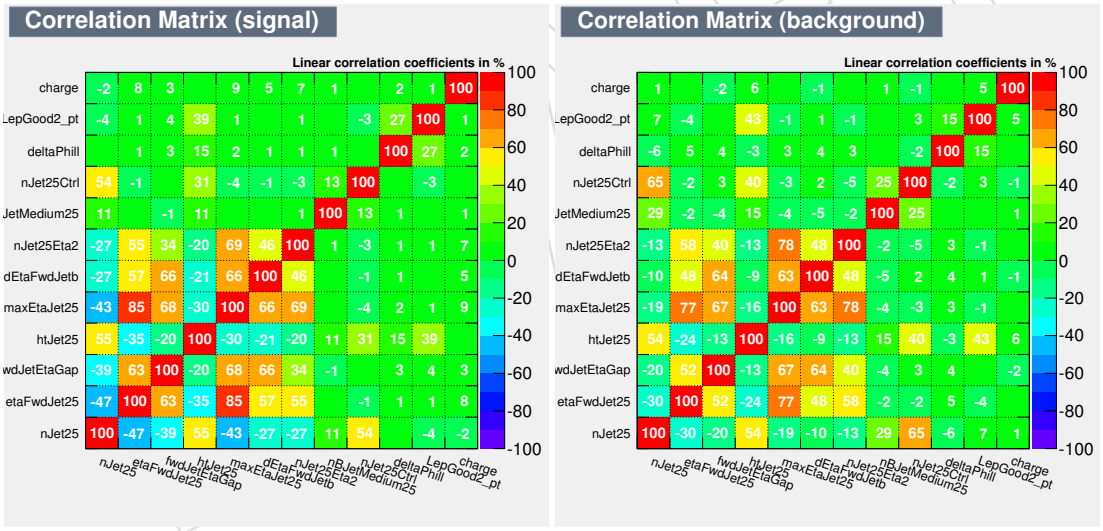

Figure 24: Correlation matrices (from TMVA output) for all 12 considered variables for the same-sign dilepton MVA for signal (left), and background (right).

Finally, the shapes for the eight input variables, for signal (composed of $\mathrm{tH(WW)q}$ and $\mathrm{tH}(\tau\tau)\mathrm{q}$) and background (composed of Fakes, WZ, $\mathrm{W^+W^+qq}$, $\mathrm{t\bar{t}W^{\pm}}$, and $\mathrm{t\bar{t}Z}$) processes, are shown in Fig. 25.

6.2.3 Linear Likelihood Discriminant

As for the three lepton channel, we build a simple linear likelihood discriminant from the eight input variables defined above in Eq. 3. The normalized likelihood discriminant distributions for all three channels are shown in Figure 26.

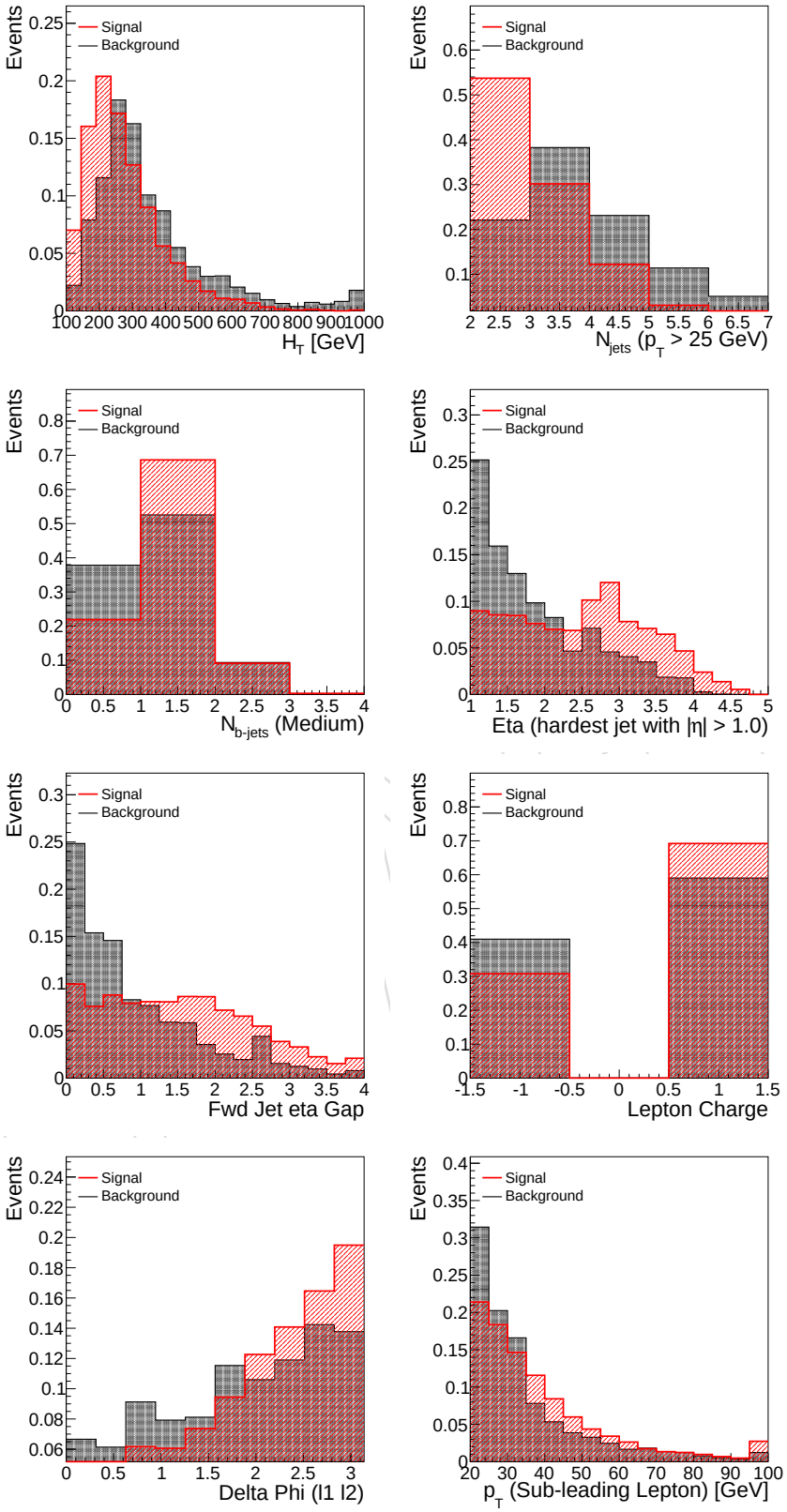

Figure 25: Signal and background shapes for same-sign dilepton likelihood discriminant input variables. Background is composed of (Fakes, WZ, W^+W^+qq, $t\bar{t}W^{\pm}$, and $t\bar{t}Z$).

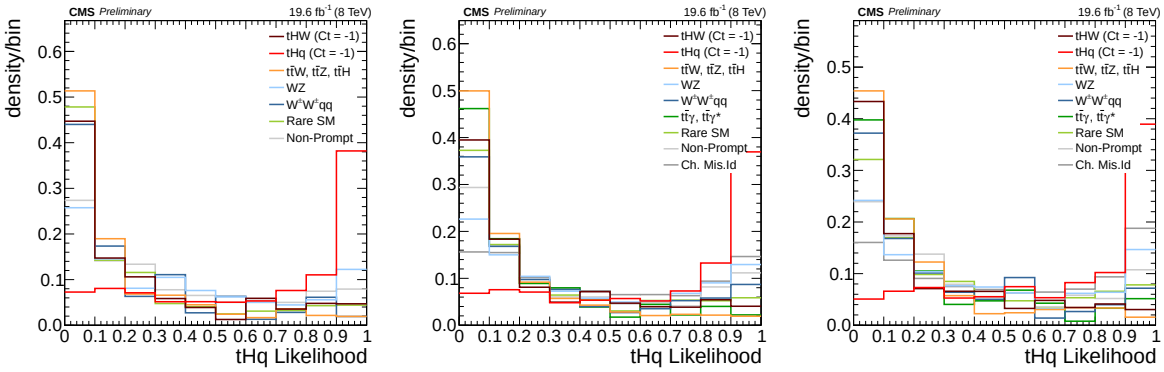

Figure 26: Linear likelihood discriminant output shapes for various signal and background processes, for the $\mu\mu$ (left), $e\mu$ (center), and ee (right) channels.

6.2.4 Performance and Results

Finally, the likelihood outputs for the three channels, normalized to the cross section is shown in Fig. 27.

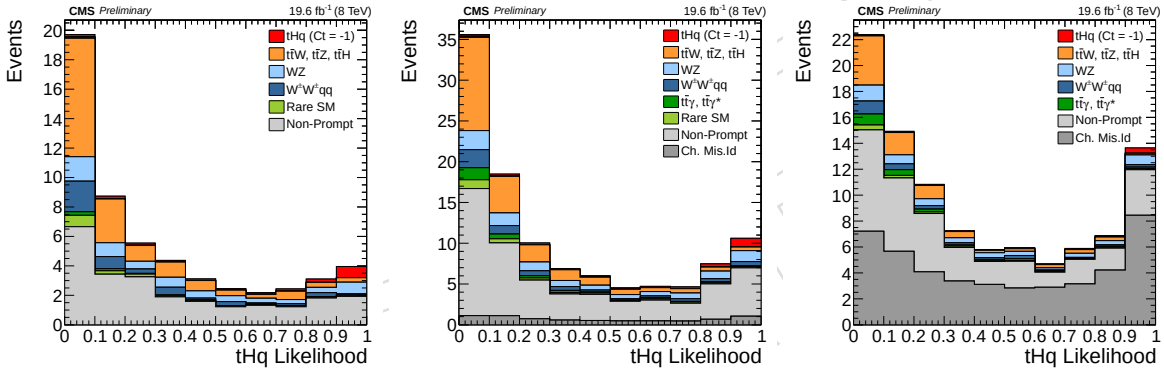

Figure 27: Linear likelihood discriminant output, normalized to cross section and integrated luminosity for signal and background processes, for the $\mu\mu$ (left), $e\mu$ (center), and ee (right) channels.

The comparison of observed data and predicted backgrounds for the eight likelihood input variables for all three channels can be found in Figures 28, 29, and 30.

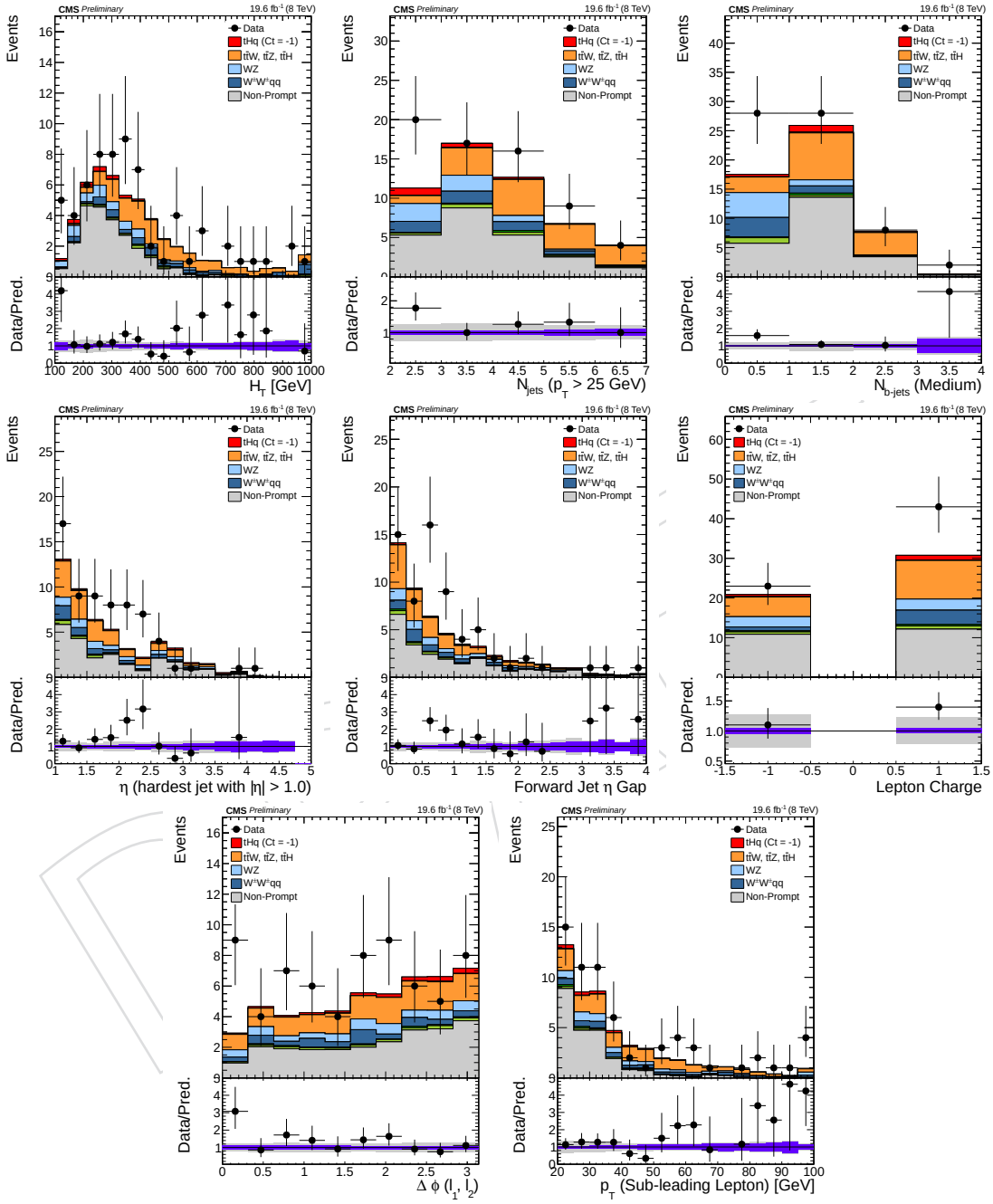

Figure 28: Observed data and predicted background distributions for the eight likelihood input variables for the $\mu\mu$ channel.

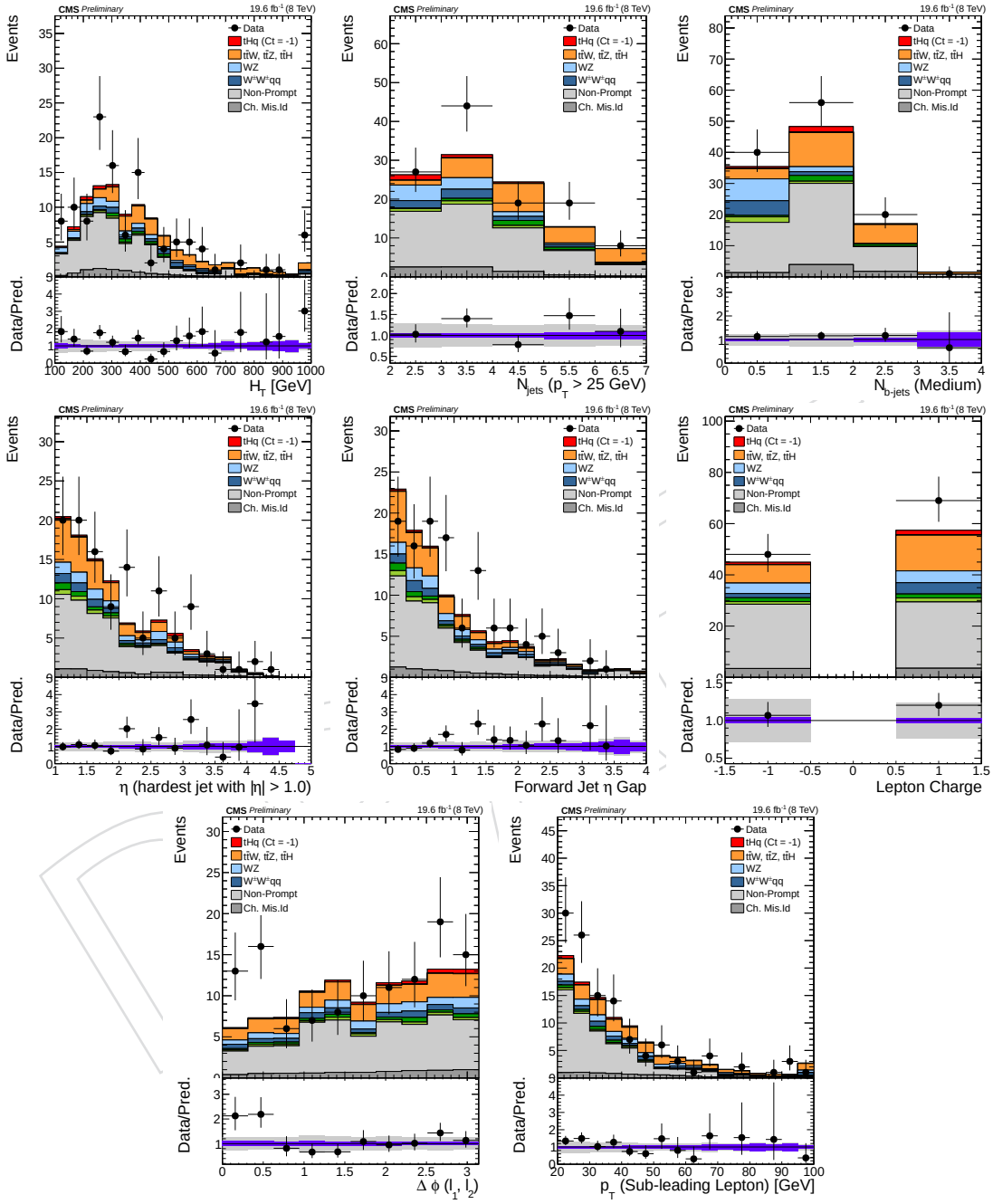

Figure 29: Observed data and predicted background distributions for the eight likelihood input variables for the $e\mu$ channel.

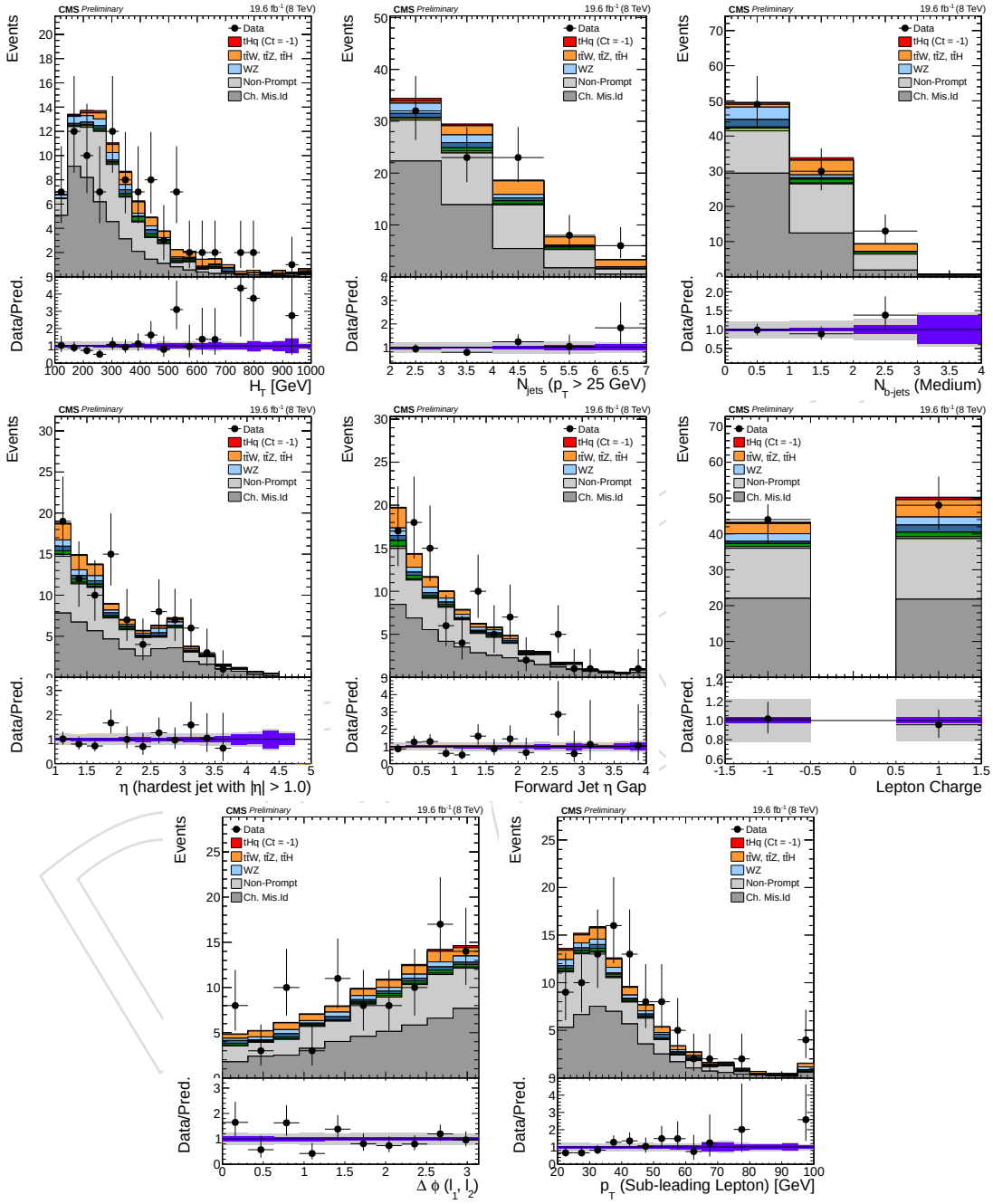

Figure 30: Observed data and predicted background distributions for the eight likelihood input variables for the *ee* channel.

6.3 Other possible Background contributions

Two processes that could also have a similar final state as tHq, for some specific combination of decays, would be $t\bar{t}WW$ and tbZ.

For $t\bar{t}WW$ a simple calculation is done to estimate the possible final expected yield for the likelihood preselection. Using Madgraph for Leading Order (LO) and taking the k-factor from plain $t\bar{t}$ production, the cross section for this processs is estimated to be 3 fb, about 1/10 of $\sigma(t\bar{t}H) \times$ BR($H \rightarrow W^+W^-$) = 130 $\times$ 0.23 = 30 fb. This means about a 1/10 of the $t\bar{t}H$ contribution can be expected to contribute to the final selection described in Sections 6.1 and 6.2, when assuming similar background-like kinematics. From results shown in Table 17 and Table 18 this leads to about 0.6 events for 2 leptons sign sign and less than 0.2 for 3 lepton final state (without taking into account leptonic decays). It is a negligible contribution. These calculations are confirmed by running on the simulation for the 2 same-sign lepton final state. The predicted yields for $\mu\mu$/eμ/ee channels in the same-sign preselection are 0.22/0.31/0.13 events respectively. I.e. consistent with the previous estimation.

In the case of tbZ, some calculations can be done to estimate the number of events at the final selection level (in which the final discriminator is built). The cross section for tHq (ignoring the tHW contribution) with C_t = -1 is 0.231 pb, while the tbZ process has an estimated cross section, for SM, of 0.0265 pb (10 times smaller than the signal). Assuming the same efficiency for this background as for tHq then, the expected number of events, at the likelihood selection level, would be about 10 times smaller than the signal (considering both processes have similar kinematics), i.e. about 0.13 events ((0.95+0.31)/10, from Table 17).

7 Systematic uncertainties

Different sources of systematic errors are taken into account.

- **Luminosity**: The uncertainty associated with the luminosity measurement in the CMS detector is 2.6 % [52].

- **Pile-up**: To estimate the uncertainty on the pile-up reweighting, the central value for the total inelastic cross section of 69.4mb for 2012 for 2012 data is varied $\pm$ 5% [53], see Figure 31. The difference on the expected number of events, at a given loose pre-selection, estimated with the central value compared to these variations is taken as a systematic. The effect on the signal is about 3%, and slightly lower for the backgrounds, around 1.3% (estimated from a WZ sample). This difference is due to the fact that the pile-up profiles on the MC samples are different, and thus the systematic. While for the backgrounds the pile-up profile used for the simulation is PU_S10_START53, for the signal is PU_2012_Startup_50ns_PoissonOOTPU.

- **Lepton efficiencies**: The uncertainties on the lepton efficiencies come from the method used to estimate the data-to-MC scale factors. From previous studies for muons, the recommendation is a 0.5% for identification, 0.2% for isolation [54], and 1.5% for the background shape [55]. In total, a 1.6% is considered per lepton (for muons and electrons).

- **Trigger efficiencies**: As explained in Section 3.2.3, no scale factors are used on MC to correct for trigger efficiencies. To account for possible inefficiencies, a systematic

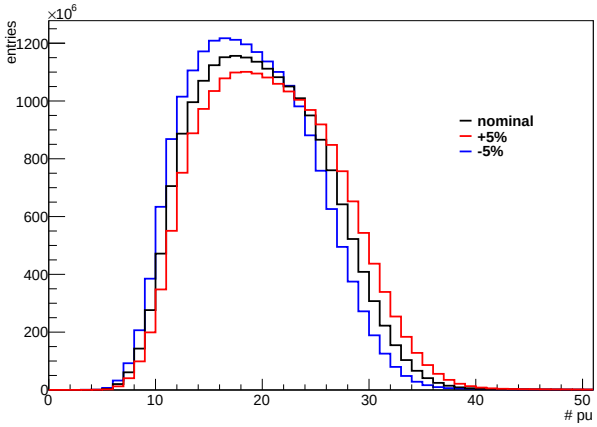

Figure 31: Number of pile-up interactions with different minimum bias cross-section values

error of 1% is considered.

- **Muon and Electron momentum scale**: To estimate the muon momentum scale the corrections computed with the Rochester Method are applied. This method provides a way to estimate the systematic uncertainty on the correction, smearing the central value of the correction by 1 sigma. The final effect on the event selection is negligible (<0.5%) and it is not considered. For electrons, the estimated momentum is varied by a 3% up and down, see Figure 32.

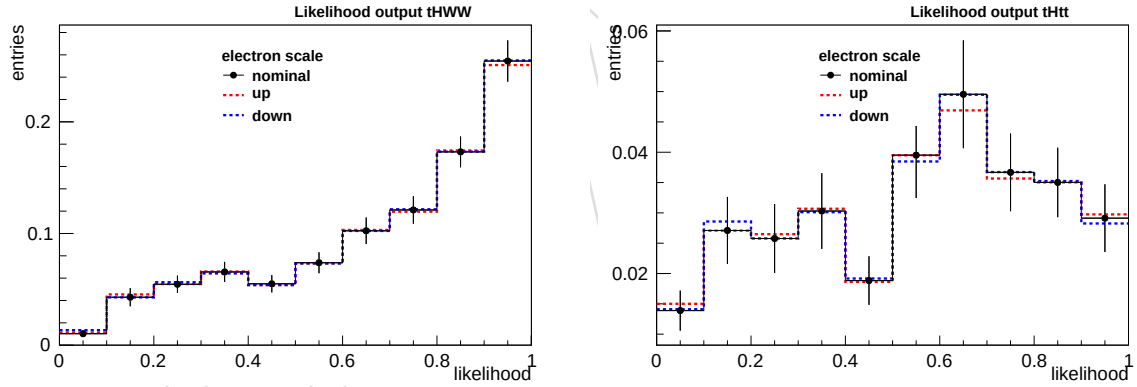

Figure 32: Nominal distribution of the likelihood output, for the **3 leptons final state**, for the tH (H→WW) (left) and tH (H→ $\tau\tau$) (right) signals, compared with the up/down variation corresponding to the electron energy scale uncertainty

- **Jet Energy Corrections - JEC**: The jet transverse energy is corrected using the official recipe. To account for the uncertainties on these corrections, the value for this energy is varied up and down within the given uncertainties. The difference of these variations with respect to the nominal value, propagated to the final discriminant output, is taken as a shape systematic, as shown in Figure 33. The corresponding figures for the same-sign dilepton channel can be found in the appendix (App. D, page 76)

- **b-Tagging efficiency**: The error assigned to the re-weighting by the b-tagging effi-

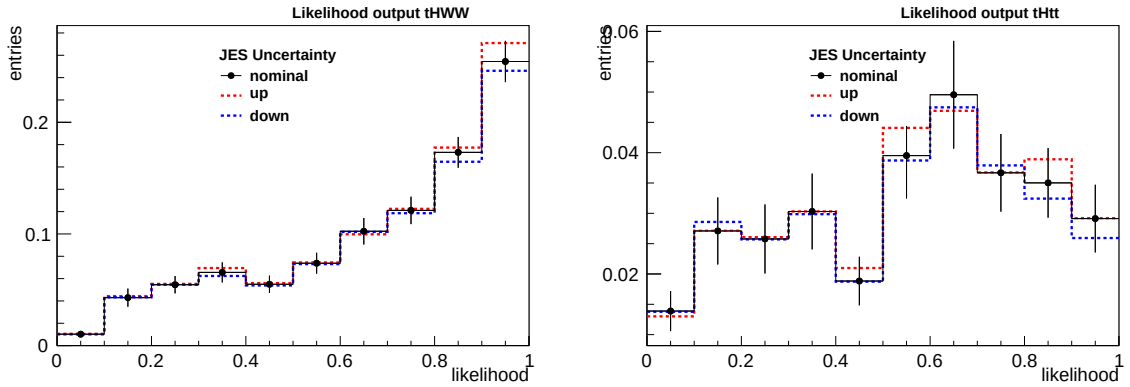

Figure 33: Nominal distribution of the likelihood output, for the **3 leptons final state**, for the tH (H→WW) (left) and tH (H→ $\tau\tau$) (right), signals compared with the up/down variation corresponding to the Jet Energy corrections uncertainty

ciencies is estimated using the official recipe. When the jet has c-quark flavour, or $p_{\mathrm{T}} < 20$ GeV or $p_{\mathrm{T}} > 800$ GeV, the assigned error to the scale factor is doubled.

For the final weight error estimation, the error on the scale factor is taken into account. Also, errors among b/c jets and light jets are taken as correlated, while as uncorrelated among b/c and light jets. The final systematic assigned to this b-tagging efficiency is assumed to be a shape uncertainty, varying the nominal likelihood output by the error on the b-tag weight, see Figure 34. The corresponding figures for the same-sign dilepton channel can be found in the appendix (App. D, page 76)

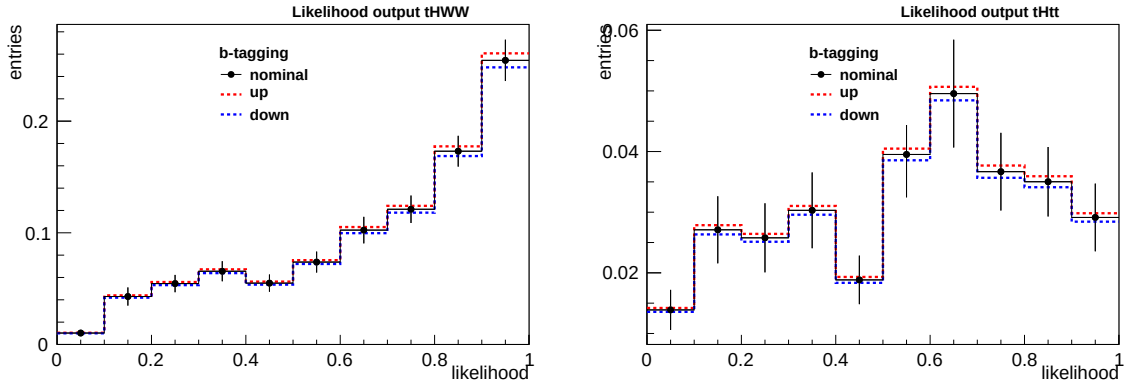

Figure 34: Nominal distribution of the likelihood output, for the **3 leptons final state**, for the tH (H→WW) (left) and tH (H→ $\tau\tau$) (right) signals, compared with the up/down variation corresponding to the b-tagging efficiency uncertainty

- **Data driven method for $\mathrm{t\bar{t}}$ estimation**: In the case of the three leptons final state, as explained in Section 5.1.3, a closure test is performed to check the validity of the method, yielding a 30% difference for a certain $\mathrm{t\bar{t}}$ enriched region. This uncertainty is taken as a scale systematic. Two other sources of systematics are considered, in this case, as shape uncertainties:

  1. Change on the fake rate as a result of applying a cut on the missing E_T in the QCD control region used to estimate it, changing this way the amount of EWK contribution to this region, see Figure 35 for the likelihood shape variation. The overall effect in the final prediction is about 10%.

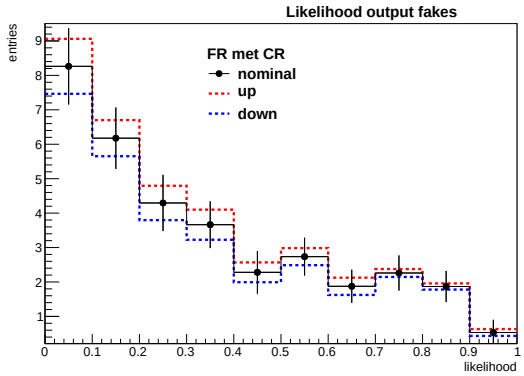

Figure 35: Nominal $t\bar{t}$ likelihood output distribution, for the **3 leptons final state**, compared with the up/down variation corresponding to the change on the fake rate estimated varying the control region definition

2. Change on the fake rate varying up and down within the statistical errors shown in Tables 11 and 12, properly propagated to the final weight estimation. This effect is shown bin by bin for the likelihood output in Figure 36, while the total effect on the expected number of events is about 14%.

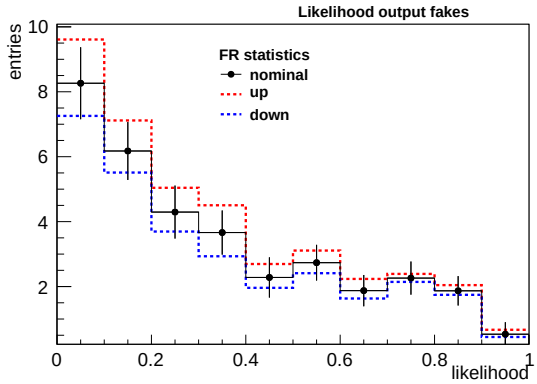

Figure 36: Nominal $t\bar{t}$ likelihood output distribution, for the **3 leptons final state**, compared with the up/down variation corresponding to statistical error on the fake rate estimated .

In case of the same-sign dilepton channels, the rate and shape uncertainties are estimated as described in Sec 5.2.5, following the studies in [32]. The overall rate uncertainty amounts to 50% for the $\mu\mu$ channel and 20% + 40% for the $e\mu$ channel (from the muon and electron fake rates respectively). The shape uncertainty envelopes are shown in Fig. 49 and are between 10 and 20%.

- **Data driven method for charge mis-identification**: The systematic uncertainty of the charge mis-identification prediction is based on the uncertainty on the measure mis-id. probabilities, propagated to the prediction, and amounts to about 30%.

- **Flavor Scheme**: The effect of employing the five-flavor scheme (5FS) for modeling the signal process, as opposed to the four-flavor scheme (4FS), is studied by comparing the signal selection efficiency at parton level for each flavor scheme. While the 5FS assumes that the b-quark can be present in the colliding protons, the 4FS does not, and hence, the initial b-quark in the diagram comes from a gluon splitting. A difference on the efficiency of 8% and 7%, respectively, for the same-sign

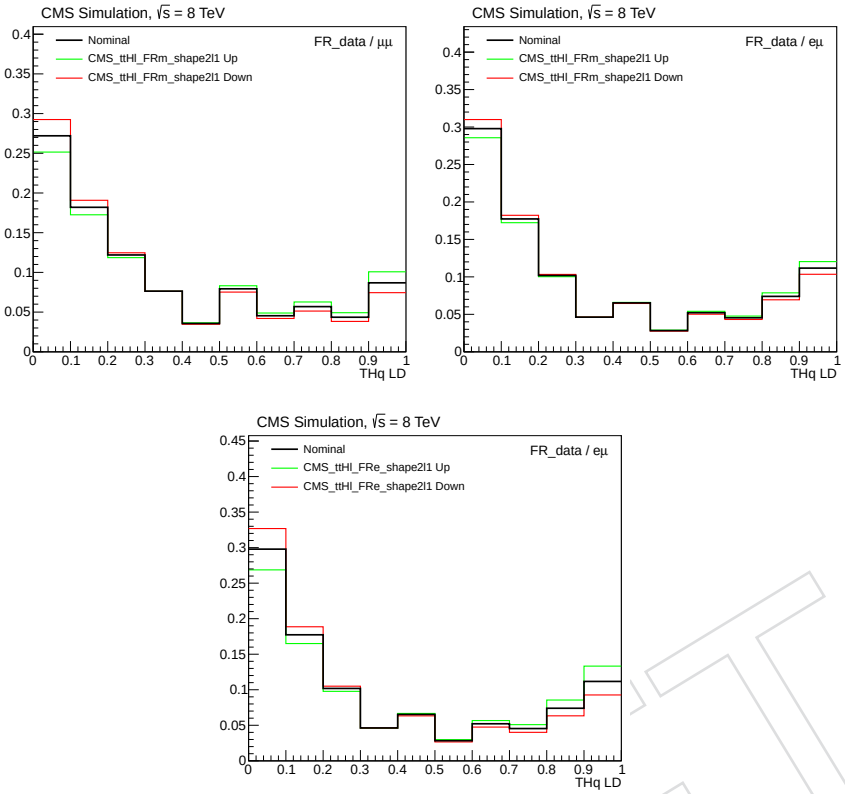

Figure 37: Uncertainty on the linear likelihood discriminant shape from the fake lepton estimation in the **same-sign dilepton** channels, for the $\mu\mu$ channel (left), the muon fake rate in the $e\mu$ channel (center), and the electron fake rate in the $e\mu$ channel (right).

dilepton and three leptons final state is observed at parton level, when applying a similar selection to the events as applied in Sec. 6. The difference in the estimated cross section, as given by MADGRAPH 5, is 18% and 11%. Taking into account both the differences in the efficiency and the cross section, the final assigned systematic uncertainty is 10% for same-sign dilepton channels, and 16% for three leptons final state.

- **PDF, Scale and Higgs decay BR**: The estimated theoretical error for the signal corresponding to the Particle Density Functions (PDF) is +1.9%/−0.2%. This error is estimated at QCD NLO using as nominal set of PDF the CTEQ6.6 one, evaluated with `Madgraph5_aMC@NLO`. The renormalization and factorization scales, μ_R and μ_F respectively, are fixed to $(m_H + m_t)/2$. To estimate the effect of the variation of these scales, they are varied independently by a factor 2, estimating the effect on the signal efficiency. This effect is found to be +4.8%/−4.3%.

  The uncertainty associated to the estimation of the Higgs decay branching ratio to a W pair or a τ pair is obtained from the LHC Higgs Cross Section Working Group [56]. The uncertainty on BR($H \to W^+W^-$) is 4.26%, while for BR($H \to \tau^+\tau^-$) is 5.71%.

- **Background expected cross-sections**: The uncertainties associated to the calculation of the $t\bar{t}W^{\pm}$, $t\bar{t}Z$, and $t\bar{t}H$ cross sections at 8 TeV at NLO are 11%, 13%, and 6%, respectively [57, 58].

A summary of all the systematic uncertainties considered is shown in Table 20. Additional plots for systematic variations can be found in Appendix D.

Systematic	Rate or shape	Description, applied to
Luminosity	rate	all predicted from MC
Pile-up	rate	all predicted from MC
Lepton id-iso efficiencies	rate	all predicted from MC
Trigger efficiencies	rate	all predicted from MC
Electron momentum scale	shape	all predicted from MC
JEC/JER	shape	all predicted from MC
b-tagging efficiency	shape	qtH, WtH
Fakes lepton mis-id	rate/shape	Lepton mis-id background
Fakes charge mis-id	shape	Charge mis-id background
PDF and Scale	rate	qtH, WtH
Higgs decay BR	rate	qtH, ttH
Theoretical cross-section	rate	ttW, ttZ, ttH

Table 20: Summary of systematic uncertainties

8 Results and Interpretation

Results shown here have been updated after the tau veto is applied for the same-sign channels, to remove the overlap with the qtH, H $\rightarrow \tau\tau$ hadronic analysis [59]. See Appendix F for more details on this selection, and also a comparison with the results without the tau veto application on SS channels.

After the event pre-selection is applied to the data events, we observe 117 events in the same-sign $e\mu$ channel and 66 in the $\mu\mu$ channel, and 42 events in the trilepton analysis (3ℓ). A simultaneous fit is performed on the likelihood output for the channels eμ, $\mu\mu$, and 3ℓ in order to estimate the upper limit on the Higgs cross section for C_t = -1. The same-sign *ee* channel is not included in the final combination. The signal is scanned in these three statistically independent samples, and the results are combined by properly taking into account the correlations of the systematics.

This search has an expected 95% confidence level (C.L.) upper limit of 8.1 and 9.3 times the expected cross section for C_t = -1, for the $\mu\mu$ and eμ same sign leptons final states, respectively, and 8.6 times the expected cross section for the three leptons final state. The combination of these three channels gives a final result of 5.0 times the expected cross section. The observed upper limits are 9.3 for $\mu\mu$ channel, 11.4 for $e\mu$ channel and 11.5 for the three leptons final state channel, while 6.7 times the expected cross section for the combination. In this evaluation, all extra Higgs yields arising from the C_t = -1 condition are considered as signal. Both the significance and the upper limit have been computed with the bayesian-frequentist hybrid approach CLs [60–62]. In Table 21 a summary of these results is shown, separately for the eμ and $\mu\mu$ same-sign (SS) channels, and for the three leptons final state. The upper-limit for the ee SS channel gives a result of about 22 times $\sigma_{C_t = -1}$. No gain is found in the final combination by adding this channel, so it is not included.

A maximum likelihood fit is performed to compare pre-fit and post-fit nuisance parameters given by the *combine* tool (provided in the Higgs analysis group). This check is done for every channel separately and for the combination of the three, both for using the expectation only (Asimov dataset) and the observed Data.

Channel	Observed	Expected	68% prob. band	95% prob. band
SS eμ	11.4	9.3	[7.0, 13.5]	[5.4, 18.8]
SS $\mu\mu$	9.3	8.1	[6.0, 11.8]	[4.7, 16.7]
3ℓ	11.5	8.6	[6.6, 12.4]	[5.7, 18.0]
combined	6.7	5.0	[3.6, 7.1]	[2.9, 10.3]

Table 21: Expected and Observed 95% C.L. upper limit on Higgs cross section for C_t = -1 on qtH production. The $\pm 1\sigma$ and $\pm 2\sigma$ probability bands are also shown.

960 Tables 22 and 23 show the best fit signal strength, r, from the maximum likelihood fit when using both the Asimov and the observed datasets, respectively.

Channel	r	68% prob. band	n_{LL} s+b	n_{LL} b
SS eμ	0.010	[-0.010, +3.87]	2.3 $\times 10^{-6}$	-5.7 $\times 10^{-14}$
SS $\mu\mu$	0.009	[-0.009, +3.10]	2.3 $\times 10^{-6}$	0
3ℓ	0.009	[-0.009, +3.38]	2.9 $\times 10^{-6}$	2.8 $\times 10^{-14}$
combined	0.003	[-0.003, +1.96]	7.6 $\times 10^{-7}$	-5.7 $\times 10^{-14}$

Table 22: Best fit signal strength, r, from the maximum likelihood fit and n_{LL} results, using **Asimov data**.

Channel	r	68% prob. band	n_{LL} s+b	n_{LL} b
SS eμ	1.85	[-1.85, +4.60]	-1.85	-1.77
SS $\mu\mu$	0.20	[-0.20, +3.93]	-2.85	-2.85
3ℓ	3.08	[-2.98, +3.89]	-1.20	-0.67
combined	1.81	[-1.81, +2.34]	-5.78	-5.40

Table 23: Best fit signal strength, r, from the maximum likelihood fit and n_{LL} results, using **Observed data**.

961

962 Tables 24, 25, and 26 show the pre and the post-fit expected yields and uncertainties for the
963 signals and the background processes. In the case of the post-fit, both the results for the back-
964 ground only fit and signal plus background fit cases are shown. In Figures 38, 39 and 40, the
965 likelihood distributions before the maximum likelihood fit, and after the fit for the background-
966 only and signal plus background cases, are shown, for the SS $\mu\mu$, SS eμ and 3ℓ channels, respec-
967 tively.

	$\mu\mu$		
	pre-fit	b-only fit	s+b fit
W^+W^+qq	4.62 ± 0.68	4.60 ± 0.68	4.60 ± 0.68
WZ	5.38 ± 1.88	5.56 ± 2.27	5.47 ± 2.10
$t\bar{t}\gamma^*$	0.51 ± 0.23	0.50 ± 0.25	0.50 ± 0.20
$t\bar{t}\gamma$	0.09 ± 0.03	0.09 ± 0.03	0.09 ± 0.03
$t\bar{t}Z$	2.23 ± 0.46	2.23 ± 0.44	2.23 ± 0.41
$t\bar{t}W^{\pm}$	10.23 ± 2.18	10.09 ± 2.26	10.18 ± 2.24
$t\bar{t}H$	2.26 ± 0.34	2.26 ± 0.31	2.26 ± 0.34
Rare SM	1.41 ± 0.61	1.40 ± 0.63	1.40 ± 0.68
Non-Prompt	23.01 ± 10.88	36.32 ± 8.72	33.34 ± 8.34
$tH(\tau\tau)W$	0.06 ± 0.02		0.10 ± 0.12
$tH(WW)W$	0.16 ± 0.03		0.28 ± 0.29
$tH(\tau\tau)q$	0.33 ± 0.05		0.59 ± 0.61
$tH(WW)q$	1.41 ± 0.21		2.55 ± 2.62
Total Background	$\mathbf{49.73 \pm 11.31}$	$\mathbf{63.05 \pm 9.36}$	$\mathbf{60.07 \pm 8.95}$
Total Signal	$\mathbf{1.95 \pm 0.22}$		$\mathbf{3.53 \pm 2.71}$
Data		**66**	

Table 24: Pre and post-fit expected yields and uncertainties for background and signal processes for both the background only (b-only) and signal+background (s+b) fits for the **same-sign $\mu\mu$ final state**.

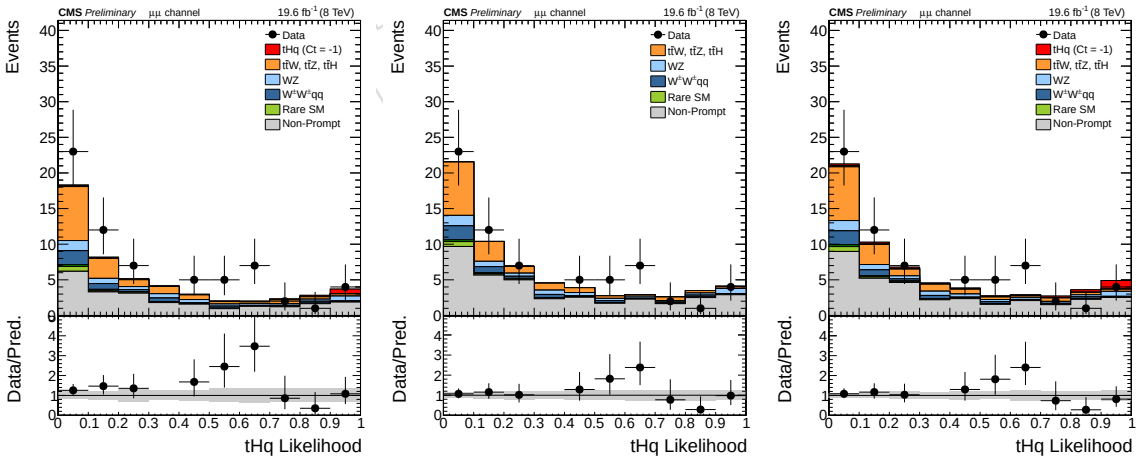

Figure 38: Linear likelihood discriminant output, normalized to cross section and integrated luminosity for signal and background processes for the **SS $\mu\mu$ channel**, **before** the maximum likelihood fit (left), for **background-only fit** (center), and for **signal plus background fit** (right).

	$e\mu$		
	pre-fit	b-only fit	s+b fit
W^+W^+qq	6.05 ± 0.85	6.03 ± 0.85	6.03 ± 0.85
WZ	8.71 ± 3.00	8.97 ± 3.54	8.83 ± 3.25
$t\bar{t}\gamma^*$	1.04 ± 0.48	1.04 ± 0.52	1.04 ± 0.42
$t\bar{t}\gamma$	2.02 ± 0.59	2.02 ± 0.64	2.02 ± 0.60
$t\bar{t}Z$	2.88 ± 0.57	2.88 ± 0.54	2.87 ± 0.50
$t\bar{t}W^{\pm}$	14.95 ± 3.21	14.69 ± 3.35	14.85 ± 3.32
$t\bar{t}H$	3.24 ± 0.47	3.24 ± 0.44	3.24 ± 0.47
Rare SM	2.58 ± 1.11	2.57 ± 1.16	2.57 ± 1.23
Charge Mis-ID	6.99 ± 1.78	6.99 ± 1.81	6.96 ± 1.76
Non-Prompt	50.96 ± 25.27	68.59 ± 11.13	63.74 ± 12.46
$tH(\tau\tau)W$	0.07 ± 0.01		0.13 ± 0.14
tH(WW)W	0.26 ± 0.04		0.47 ± 0.48
$tH(\tau\tau)q$	0.50 ± 0.07		0.90 ± 0.91
tH(WW)q	2.06 ± 0.31		3.73 ± 3.84
Total Background	$\mathbf{99.41 \pm 25.77}$	$\mathbf{117.00 \pm 12.41}$	$\mathbf{112.13 \pm 13.53}$
Total Signal	$\mathbf{2.89 \pm 0.32}$		$\mathbf{5.22 \pm 3.98}$
Data		**117**	

Table 25: Pre and post-fit expected yields and uncertainties for background and signal processes for both the background only (b-only) and signal+background (s+b) fits for the **same-sign $e\mu$ final state**.

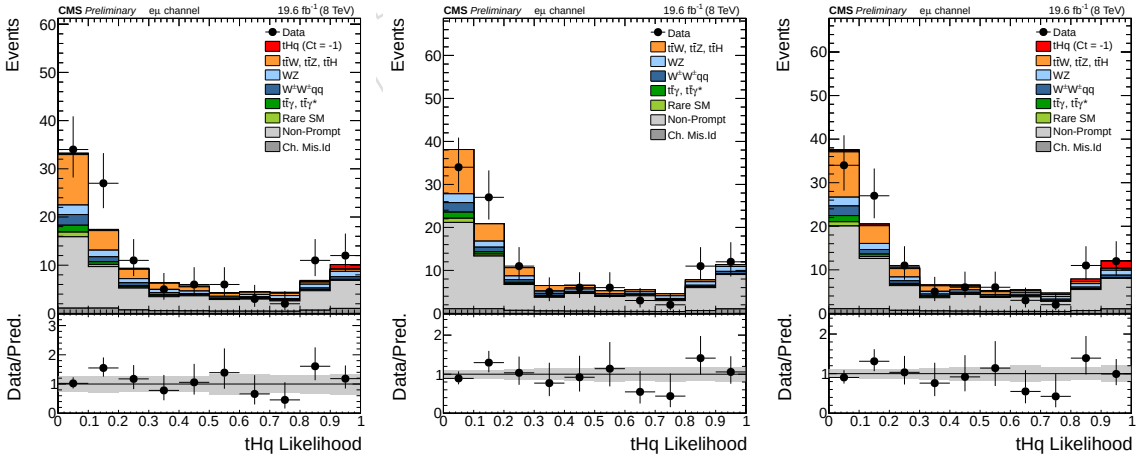

Figure 39: Linear likelihood discriminant output, normalized to cross section and integrated luminosity for signal and background processes for the **SS $e\mu$ channel**, **before** the maximum likelihood fit (left), for **background-only fit** (center), and for **signal plus background fit** (right).

	3ℓ		
	pre-fit	b-only fit	s+b fit
WZ/ZZ/WW	1.18 ± 0.14	1.19 ± 0.13	1.19 ± 0.14
VVV	0.11 ± 0.03	0.11 ± 0.03	0.11 ± 0.03
$t\bar{t}Z$	2.20 ± 0.39	2.21 ± 0.35	2.21 ± 0.36
$t\bar{t}W^{\pm}$	3.01 ± 0.47	3.03 ± 0.48	3.03 ± 0.51
$t\bar{t}H$	1.52 ± 0.19	1.52 ± 0.18	1.52 ± 0.18
Non-Prompt	33.95 ± 12.28	33.70 ± 5.58	31.44 ± 6.52
$tH(\tau\tau)W$	0.07 ± 0.02		0.12 ± 0.12
tH(WW)W	0.19 ± 0.04		0.35 ± 0.35
$tH(\tau\tau)q$	0.31 ± 0.06		0.56 ± 0.58
tH(WW)q	0.95 ± 0.17		1.73 ± 1.80
Total Background	$\mathbf{41.97 \pm 12.30}$	$\mathbf{41.77 \pm 5.61}$	$\mathbf{39.50 \pm 6.55}$
Total Signal	$\mathbf{1.52 \pm 0.18}$		$\mathbf{2.76 \pm 1.93}$
Data		**42**	

Table 26: Pre and post-fit expected yields and uncertainties for background and signal processes for both the background only (b-only) and signal+background (s+b) fits for the **three-lepton final state**.

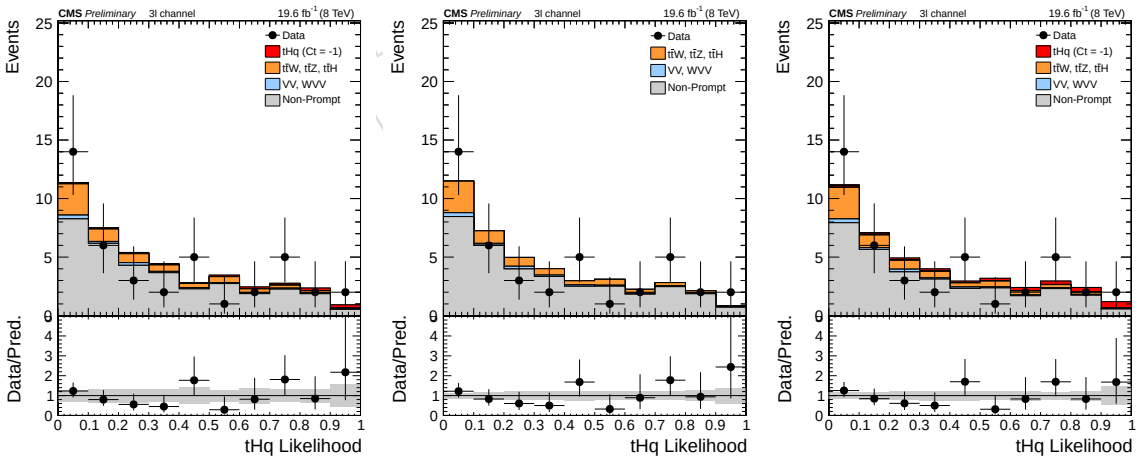

Figure 40: Linear likelihood discriminant output, normalized to cross section and integrated luminosity for signal and background processes for the **3ℓ channel**, **before** the maximum likelihood fit (left), for **background-only fit** (center), and for **signal plus background fit** (right).

8.1 Checks on the effect of single systematic sources

In this section, a systematic study is performed to check the effect of every single source of systematc in the final combined result. Two approaches are followed. The first approach consists on estimating the variation with respect to the nominal value, estimated without systematics, when adding every single source of systematics. The upper limits are estimated using the Asymptotic method. For this first case, it is found to be 3.0 at 95% C.L., with a 68% probability band equal to [2.1, 4.4]. The second approach shows the variation of the upper limit, when removing one single systematic source, and only for the expected value. This value is 4.8 at 95% C.L., with a 68% probability band equal to [3.4, 7.2]. These results are shown in Table 27.

Systematic source	Type	Impact as exclusive source on final limit (%)	Improvement of final limit after removal (%)
Luminosity	rate	< 1	< 1
Pileup	rate	< 1	< 1
Lepton trigger efficiency	rate	< 1	< 1
Lepton ID, Iso, IP, efficiencies (SS)	rate	< 1	< 1
Lepton ID, Iso, IP, efficiencies (3ℓ)	rate	< 1	< 1
Electron energy scale	shape	< 1	< 1
JEC	shape	< 1	< 1
JER	shape	< 1	< 1
b-tagging efficiencies	shape	< 1	< 1
Flavour Scheme	rate	2	1
Higgs BR	rate	< 1	< 1
QCD scale	rate	< 1	< 1
PDFs	rate	< 1	< 1
Rare SM normalization	rate	< 1	< 1
$t\bar{t}\gamma$ normalization	rate	< 1	< 1
$t\bar{t}\gamma^*$ normalization	rate	< 1	< 1
μ FR normalization (SS)	rate	26	19
e FR normalization (SS)	rate	12	5
μ FR leptons shape (SS)	shape	< 1	1
e FR leptons shape (SS)	shape	< 1	2
Closure test FR normalization (3ℓ)	rate	3	3
QCD Control Region for FR (3ℓ)	shape	1	< 1
FR variation within stats (3ℓ)	shape	1	< 1
Charge-misID (SS)	rate	< 1	< 1
Bin-by-bin stats for Fakes (3ℓ)	shape	2	3
Bin-by-bin stats for Fakes (SS)	shape	4	3

Table 27: Improvement when omitting single systematic uncertainties on the final combined limit and impact when using a systematic source exclusively. Values denoted with < 1 have a negligible impact compared to the significant figures listed in the table.

8.2 Checks on the post-fit nuisance parameters

In this section two sets of results are shown per channel and the combination: first, tables with the post-fit parameters considering background only and signal+background fits, for the Asimov data. After, the same tables when considering the observed data. Only the nuissance parameters in which a larger variation is observed are shown.

In Tables 28 and 29 the results for the variation of the parameters are shown for Asimov and Observed data, respectively, for the case of the combination of the channels. For each parameter, we show the shift in value and the post-fit uncertainty, both normalized to the input values, and the linear correlation between the parameter and the signal strength (as given by the tool). In Tables 30 and 31 a summary of the yields after the fit is presented.

The same Tables can be found for the $\mu\mu$ SS channel (Tables 32, 33, 34, 35), the emu SS channel (Tables 36, 37, 38, 39) and 3ℓ channel (Tables 40, 41, 42, 43)

A brief description of the nuissance parameters that appear in these tables is summarized:

- CMS_FakesMCTestBounding: rate systematic to account for the mc closure test for the estimation of ttbar (fakes) background in the 3 lepton final state.
- FakesStatFR: Shape systematic for the 3l final state, to account for the statistical errors in the fake rate estimation.
- CMS_tHql_FRe_norm and CMS_tHql_FRm_norm: rate systematic for the normalization of the fake rates for electrons and muons, respectively, for the 2SS leptons final state.
- CMS_tHql_FRe_shape2l1 and CMS_tHql_FRe_shape2l2: shape systematics (per legs, l1 and l2) for the fake rate estimatino, for the 2SS leptons final state.
- CMS_scale_j: Jet Energy Scale systematic.
- StatBounding_Fakes_3l_binXX and StatBounding_Fakes_ss2l_binYY bin by bin statistical uncertainty (sample size).

Asimov Data	b-only fit	$s+b$ fit	
name	$\Delta x/\sigma_{\text{in}}, \sigma_{\text{out}}/\sigma_{\text{in}}$	$\Delta x/\sigma_{\text{in}}, \sigma_{\text{out}}/\sigma_{\text{in}}$	$\rho(\theta,\mu)$
CMS_FakesMCTestBounding	-0.00, 0.70	-0.00, 0.71	-0.17
CMS_tHql_FRe_norm	-0.00, 0.57	-0.00, 0.58	-0.17
CMS_tHql_FRe_shape2l1	+0.00, 0.89	-0.00, 0.89	-0.12
CMS_tHql_FRe_shape2l2	+0.00, 0.89	+0.00, 0.89	+0.03
CMS_tHql_FRm_norm	+0.00, 0.61	-0.00, 0.65	-0.35

Table 28: Post-fit nuisance parameters variation from the maximum likelihood fit, for background only and signal plus background fits, using **Asimov Data**, for the **combination of the three channels**.

Observed Data	b-only fit	$s+b$ fit	
systematic source	$\Delta x/\sigma_{\text{in}}$, $\sigma_{\text{out}}/\sigma_{\text{in}}$	$\Delta x/\sigma_{\text{in}}$, $\sigma_{\text{out}}/\sigma_{\text{in}}$	$\rho(\theta,\mu)$
CMS_FakesMCTestBounding	+0.05, 0.71	-0.13, 0.75	-0.25
CMS_res_j	-0.07, 0.99	-0.05, 0.87	+0.03
CMS_scale_j	+0.21, 1.41	+0.14, 1.26	-0.07
CMS_tHql_FRe_norm	+0.24, 0.49	+0.13, 0.54	-0.24
CMS_tHql_FRe_shape2l1	+0.06, 0.87	-0.08, 0.89	-0.18
CMS_tHql_FRe_shape2l2	-0.24, 0.83	-0.20, 0.84	+0.06
CMS_tHql_FRm_norm	+1.09, 0.51	+0.87, 0.60	-0.40
StatBounding_Fakes_3l_bin10	+0.54, 0.85	+0.24, 0.96	-0.26
StatBounding_Fakes_3l_bin3	-0.31, 0.96	-0.29, 0.97	+0.02
StatBounding_Fakes_3l_bin4	-0.33, 0.97	-0.31, 0.98	+0.02
StatBounding_Fakes_3l_bin5	+0.43, 0.91	+0.41, 0.92	-0.03
StatBounding_Fakes_3l_bin6	-0.36, 0.98	-0.35, 0.98	+0.02
StatBounding_Fakes_3l_bin8	+0.41, 0.93	+0.34, 0.94	-0.08
StatBounding_Fakes_ss2l_bin2	+0.57, 0.93	+0.56, 0.94	-0.01
StatBounding_Fakes_ss2l_bin4	-0.78, 0.96	-0.73, 0.97	+0.06
StatBounding_Fakes_ss2l_bin6	+0.50, 0.90	+0.46, 0.91	-0.05
StatBounding_Fakes_ss2l_bin7	+0.53, 0.88	+0.50, 0.90	-0.04
StatBounding_Fakes_ss2l_bin8	-0.44, 0.95	-0.43, 0.96	+0.01

Table 29: Post-fit nuisance parameters variation from the maximum likelihood fit, for background only and signal plus background fits, using **Observed Data**, for the **combination of the three channels**.

channel	process	s+b fit	b-only fit
SS eμ	Charge Mis-ID	6.987	6.987
SS eμ	$t\bar{t}\gamma$	2.023	2.023
SS eμ	$t\bar{t}\gamma^*$	1.040	1.040
SS eμ	Fakes	50.949	50.957
SS eμ	Rare SM	2.578	2.578
SS eμ	W^+W^+(DPI)	6.051	6.051
SS eμ	WZ	8.707	8.707
SS eμ	$t\bar{t}H$	3.243	3.243
SS eμ	$t\bar{t}W^{\pm}$	14.953	14.953
SS eμ	$t\bar{t}Z$	2.875	2.875
SS eμ	tH(WW)W	0.001	0.000
SS eμ	tH($\tau\tau$)W	0.000	0.000
SS eμ	tH(WW)q	0.004	0.000
SS eμ	tH($\tau\tau$)q	0.001	0.000
3ℓ	Fakes	33.947	33.950
3ℓ	Others	0.107	0.107
3ℓ	VV	1.184	1.184
3ℓ	$t\bar{t}H$	1.517	1.517
3ℓ	$t\bar{t}W^{\pm}$	3.013	3.013
3ℓ	$t\bar{t}Z$	2.203	2.203
3ℓ	tH(WW)W	0.000	0.000
3ℓ	tH($\tau\tau$)W	0.000	0.000
3ℓ	tH(WW)q	0.002	0.000
3ℓ	tH($\tau\tau$)q	0.001	0.000
SS $\mu\mu$	$t\bar{t}\gamma$	0.092	0.092
SS $\mu\mu$	$t\bar{t}\gamma^*$	0.505	0.505
SS $\mu\mu$	Fakes	23.006	23.010
SS $\mu\mu$	Rare SM	1.409	1.409
SS $\mu\mu$	W^+W^+(DPI)	4.615	4.615
SS $\mu\mu$	WZ	5.385	5.385
SS $\mu\mu$	$t\bar{t}H$	2.257	2.257
SS $\mu\mu$	$t\bar{t}W^{\pm}$	10.231	10.231
SS $\mu\mu$	$t\bar{t}Z$	2.229	2.229
SS $\mu\mu$	tH(WW)W	0.000	0.000
SS $\mu\mu$	tH($\tau\tau$)W	0.000	0.000
SS $\mu\mu$	tH(WW)q	0.003	0.000
SS $\mu\mu$	tH($\tau\tau$)q	0.001	0.000

Table 30: Post-fit nuisance parameters variation from the maximum likelihood fit, for background only and signal plus background fits, using **Asimov Data**, for the **combination of the three channels**.

channel	process	s+b fit	b-only fit
SS eμ	Charge Mis-ID	6.957	6.987
SS eμ	Fakes	63.735	68.593
SS eμ	Rare SM	2.567	2.567
SS eμ	W^+W^+(DPI)	6.026	6.035
SS eμ	WZ	8.825	8.969
SS eμ	$t\bar{t}\gamma$	2.015	2.015
SS eμ	$t\bar{t}\gamma^*$	1.038	1.037
SS eμ	$t\bar{t}H$	3.239	3.239
SS eμ	$t\bar{t}W^{\pm}$	14.851	14.688
SS eμ	$t\bar{t}Z$	2.873	2.877
SS eμ	tH(WW)W	0.467	0.000
SS eμ	tH($\tau\tau$)W	0.126	0.000
SS eμ	tH(WW)q	3.726	0.000
SS eμ	tH($\tau\tau$)q	0.900	0.000
3ℓ	Fakes	31.445	33.705
3ℓ	Others	0.107	0.107
3ℓ	VV	1.189	1.190
3ℓ	$t\bar{t}H$	1.524	1.525
3ℓ	$t\bar{t}W^{\pm}$	3.033	3.034
3ℓ	$t\bar{t}Z$	2.207	2.207
3ℓ	tH(WW)W	0.350	0.000
3ℓ	tH($\tau\tau$)W	0.122	0.000
3ℓ	tH(WW)q	1.730	0.000
3ℓ	tH($\tau\tau$)q	0.556	0.000
SS $\mu\mu$	Fakes	33.336	36.316
SS $\mu\mu$	Rare SM	1.403	1.403
SS $\mu\mu$	W^+W^+(DPI)	4.597	4.597
SS $\mu\mu$	WZ	5.468	5.564
SS $\mu\mu$	$t\bar{t}\gamma$	0.092	0.092
SS $\mu\mu$	$t\bar{t}\gamma^*$	0.504	0.504
SS $\mu\mu$	$t\bar{t}H$	2.258	2.260
SS $\mu\mu$	$t\bar{t}W^{\pm}$	10.182	10.087
SS $\mu\mu$	$t\bar{t}Z$	2.229	2.230
SS $\mu\mu$	tH(WW)W	0.285	0.000
SS $\mu\mu$	tH($\tau\tau$)W	0.101	0.000
SS $\mu\mu$	tH(WW)q	2.552	0.000
SS $\mu\mu$	tH($\tau\tau$)q	0.595	0.000

Table 31: Expected yields variation from the maximum likelihood fit, for background only and signal plus background fits, using **Observed Data**, for the **combination of the three channels**.

Asimov Data	b-only fit	$s+b$ fit	
systematic source	$\Delta x/\sigma_{\text{in}}$, $\sigma_{\text{out}}/\sigma_{\text{in}}$	$\Delta x/\sigma_{\text{in}}$, $\sigma_{\text{out}}/\sigma_{\text{in}}$	$\rho(\theta,\mu)$
CMS_tHql_FRm_norm	+0.00, 0.63	-0.00, 0.70	-0.45

Table 32: Post-fit nuisance parameters variation from the maximum likelihood fit, for background only and signal plus background fits, using **Asimov Data**, for the **SS $\mu\mu$ channel**.

Observed Data	b-only fit	$s+b$ fit	
systematic source	$\Delta x/\sigma_{\text{in}}$, $\sigma_{\text{out}}/\sigma_{\text{in}}$	$\Delta x/\sigma_{\text{in}}$, $\sigma_{\text{out}}/\sigma_{\text{in}}$	$\rho(\theta,\mu)$
CMS_scale_j	+0.17, 1.52	+0.16, 1.52	-0.09
CMS_tHql_FRm_norm	+1.05, 0.52	+1.02, 0.67	-0.63
StatBounding_Fakes_ss2l_bin4	-0.61, 0.99	-0.60, 1.00	+0.10
StatBounding_Fakes_ss2l_bin6	+0.40, 0.94	+0.40, 0.94	-0.08
StatBounding_Fakes_ss2l_bin7	+0.76, 0.91	+0.76, 0.92	-0.10
StatBounding_Fakes_ss2l_bin9	-0.40, 0.98	-0.40, 0.98	+0.05

Table 33: Post-fit nuisance parameters variation from the maximum likelihood fit, for background only and signal plus background fits, using **Observed Data**, for the **SS $\mu\mu$ channel**.

channel	process	s+b fit	b-only fit
SS $\mu\mu$	$t\bar{t}\gamma$	0.092	0.092
SS $\mu\mu$	$t\bar{t}\gamma^*$	0.505	0.505
SS $\mu\mu$	Fakes	22.999	23.010
SS $\mu\mu$	Rare SM	1.409	1.409
SS $\mu\mu$	W^+W^+(DPI)	4.615	4.615
SS $\mu\mu$	WZ	5.385	5.385
SS $\mu\mu$	$t\bar{t}H$	2.257	2.257
SS $\mu\mu$	$t\bar{t}W^{\pm}$	10.230	10.231
SS $\mu\mu$	$t\bar{t}Z$	2.229	2.229
SS $\mu\mu$	tH(WW)W	0.001	0.000
SS $\mu\mu$	tH($\tau\tau$)W	0.000	0.000
SS $\mu\mu$	tH(WW)q	0.008	0.000
SS $\mu\mu$	tH($\tau\tau$)q	0.002	0.000

Table 34: Expected yields variation from the maximum likelihood fit, for background only and signal plus background fits, using **Asimov Data**, for the **SS $\mu\mu$ channel**.

channel	process	s+b fit	b-only fit
SS $\mu\mu$	$t\bar{t}\gamma$	0.093	0.093
SS $\mu\mu$	$t\bar{t}\gamma^*$	0.510	0.510
SS $\mu\mu$	Fakes	35.365	35.717
SS $\mu\mu$	Rare SM	1.432	1.432
SS $\mu\mu$	W^+W^+(DPI)	4.640	4.639
SS $\mu\mu$	WZ	5.542	5.552
SS $\mu\mu$	$t\bar{t}H$	2.280	2.280
SS $\mu\mu$	$t\bar{t}W^{\pm}$	10.280	10.269
SS $\mu\mu$	$t\bar{t}Z$	2.256	2.256
SS $\mu\mu$	tH(WW)W	0.031	0.000
SS $\mu\mu$	tH($\tau\tau$)W	0.011	0.000
SS $\mu\mu$	tH(WW)q	0.279	0.000
SS $\mu\mu$	tH($\tau\tau$)q	0.065	0.000

Table 35: Expected yields variation from the maximum likelihood fit, for background only and signal plus background fits, using **Observed Data**, for the **SS $\mu\mu$ channel**.

Asimov Data	b-only fit	$s+b$ fit	
systematic source	$\Delta x/\sigma_{\text{in}}, \sigma_{\text{out}}/\sigma_{\text{in}}$	$\Delta x/\sigma_{\text{in}}, \sigma_{\text{out}}/\sigma_{\text{in}}$	$\rho(\theta,\mu)$
CMS_tHql_FRe_norm	-0.00, 0.65	-0.00, 0.73	-0.46
CMS_tHql_FRe_shape2l1	+0.00, 0.89	-0.00, 0.92	-0.26
CMS_tHql_FRe_shape2l2	+0.00, 0.89	-0.00, 0.89	+0.08

Table 36: Post-fit nuisance parameters variation from the maximum likelihood fit, for background only and signal plus background fits, using **Asimov Data**, for the **SS eμ channel**.

Observed Data	b-only fit	$s+b$ fit	
systematic source	$\Delta x/\sigma_{\text{in}}, \sigma_{\text{out}}/\sigma_{\text{in}}$	$\Delta x/\sigma_{\text{in}}, \sigma_{\text{out}}/\sigma_{\text{in}}$	$\rho(\theta,\mu)$
CMS_scale_j	-0.01, 0.93	+0.00, 0.86	+0.03
CMS_tHql_FRe_norm	+0.59, 0.61	+0.42, 0.76	-0.57
CMS_tHql_FRe_shape2l1	-0.01, 0.89	-0.16, 0.94	-0.35
CMS_tHql_FRe_shape2l2	-0.23, 0.84	-0.18, 0.86	+0.12
CMS_tHql_FRm_norm	+0.32, 0.90	+0.22, 0.93	-0.25
StatBounding_Fakes_ss2l_bin2	+0.44, 0.95	+0.43, 0.96	-0.03
StatBounding_Fakes_ss2l_bin8	-0.33, 0.98	-0.32, 0.98	+0.04
StatBounding_Fakes_ss2l_bin9	+0.34, 0.94	+0.30, 0.95	-0.10

Table 37: Post-fit nuisance parameters variation from the maximum likelihood fit, for background only and signal plus background fits, using **Observed Data**, for the **SS eμ channel**.

channel	process	s+b fit	b-only fit
SS eμ	Charge Mis-ID	6.987	6.987
SS eμ	$t\bar{t}\gamma$	2.023	2.023
SS eμ	$t\bar{t}\gamma^*$	1.040	1.040
SS eμ	Fakes	50.932	50.957
SS eμ	Rare SM	2.578	2.578
SS eμ	W^+W^+(DPI)	6.051	6.051
SS eμ	WZ	8.707	8.707
SS eμ	$t\bar{t}H$	3.243	3.243
SS eμ	$t\bar{t}W^{\pm}$	14.953	14.953
SS eμ	$t\bar{t}Z$	2.875	2.875
SS eμ	tH(WW)W	0.002	0.000
SS eμ	tH($\tau\tau$)W	0.001	0.000
SS eμ	tH(WW)q	0.016	0.000
SS eμ	tH($\tau\tau$)q	0.004	0.000

Table 38: Expected yields variation from the maximum likelihood fit, for background only and signal plus background fits, using **Asimov Data**, for the **SS eμ channel**.

channel	process	s+b fit	b-only fit
SS eμ	Charge Mis-ID	6.994	7.035
SS eμ	$t\bar{t}\gamma$	2.013	2.011
SS eμ	$t\bar{t}\gamma^*$	1.039	1.037
SS eμ	Fakes	62.137	66.924
SS eμ	Rare SM	2.560	2.559
SS eμ	W^+W^+(DPI)	6.044	6.044
SS eμ	WZ	8.706	8.701
SS eμ	$t\bar{t}H$	3.237	3.234
SS eμ	$t\bar{t}W^{\pm}$	14.955	14.919
SS eμ	$t\bar{t}Z$	2.865	2.861
SS eμ	tH(WW)W	0.478	0.000
SS eμ	tH($\tau\tau$)W	0.129	0.000
SS eμ	tH(WW)q	3.807	0.000
SS eμ	tH($\tau\tau$)q	0.922	0.000

Table 39: Expected yields variation from the maximum likelihood fit, for background only and signal plus background fits, using **Observed Data**, for the **SS eμ channel**.

Asimov Data	b-only fit	$s+b$ fit	
systematic source	$\Delta x/\sigma_{\text{in}}$, $\sigma_{\text{out}}/\sigma_{\text{in}}$	$\Delta x/\sigma_{\text{in}}$, $\sigma_{\text{out}}/\sigma_{\text{in}}$	$\rho(\theta,\mu)$
CMS_FakesMCTestBounding	-0.00, 0.70	-0.00, 0.73	-0.26

Table 40: Post-fit nuisance parameters variation from the maximum likelihood fit, for background only and signal plus background fits, using **Asimov Data**, for the **3ℓ channel**.

Observed Data	b-only fit	$s+b$ fit	
systematic source	$\Delta x/\sigma_{\text{in}}$, $\sigma_{\text{out}}/\sigma_{\text{in}}$	$\Delta x/\sigma_{\text{in}}$, $\sigma_{\text{out}}/\sigma_{\text{in}}$	$\rho(\theta,\mu)$
CMS_FakesMCTestBounding	+0.06, 0.71	-0.23, 0.78	-0.34
FakesStatFR	-0.20, 0.98	-0.31, 0.96	-0.10
StatBounding_Fakes_3l_bin10	+0.54, 0.85	+0.12, 0.98	-0.26
StatBounding_Fakes_3l_bin3	-0.31, 0.96	-0.28, 0.97	+0.02
StatBounding_Fakes_3l_bin4	-0.33, 0.97	-0.30, 0.98	+0.02
StatBounding_Fakes_3l_bin5	+0.43, 0.91	+0.39, 0.92	-0.05
StatBounding_Fakes_3l_bin6	-0.36, 0.98	-0.34, 0.99	+0.02
StatBounding_Fakes_3l_bin8	+0.41, 0.93	+0.29, 0.95	-0.11

Table 41: Post-fit nuisance parameters variation from the maximum likelihood fit, for background only and signal plus background fits, using **Observed Data**, for the **3ℓ channel channel**.

channel	process	s+b fit	b-only fit
3ℓ	Fakes	33.938	33.950
3ℓ	Others	0.107	0.107
3ℓ	VV	1.184	1.184
3ℓ	$t\bar{t}H$	1.517	1.517
3ℓ	$t\bar{t}W^{\pm}$	3.013	3.013
3ℓ	$t\bar{t}Z$	2.203	2.203
3ℓ	tH(WW)W	0.001	0.000
3ℓ	$tH(\tau\tau)W$	0.000	0.000
3ℓ	tH(WW)q	0.006	0.000
3ℓ	$tH(\tau\tau)q$	0.002	0.000

Table 42: Expected yields variation from the maximum likelihood fit, for background only and signal plus background fits, using **Asimov Data**, for the **3ℓ channel channel**.

channel	process	s+b fit	b-only fit
3ℓ	Fakes	30.246	33.739
3ℓ	Others	0.107	0.107
3ℓ	VV	1.184	1.186
3ℓ	$t\bar{t}H$	1.518	1.519
3ℓ	$t\bar{t}W^{\pm}$	3.010	3.014
3ℓ	$t\bar{t}Z$	2.204	2.206
3ℓ	tH(WW)W	0.594	0.000
3ℓ	$tH(\tau\tau)W$	0.207	0.000
3ℓ	tH(WW)q	2.928	0.000
3ℓ	$tH(\tau\tau)q$	0.941	0.000

Table 43: Expected yields variation from the maximum likelihood fit, for background only and signal plus background fits, using **Observed Data**, for the **3ℓ channel channel**.

9 Conclusions

We presented the first search for the associated production of a single top quark and a Higgs boson in the multilepton final state. Due to the destructive interference in the t-channel for the standard model couplings, this process is very sensitive to both the magnitude and the sign of a non-standard top-Higgs coupling.

When analyzing the total of $19.7 \pm 0.5\,\text{fb}^{-1}$ of data collected by CMS at 8 TeV, 225 events have been found which pass the final event selection criteria, corresponding to the eμ and $\mu\mu$ SS channels and 3ℓ channel. A statistical analysis of the results led to a 95% confidence level expected upper limit on the production of events containing a Higgs boson with inverted top coupling (C_t = -1) of 5.0 times the cross section, and a observed upper limit of 6.7.

References

[1] CMS Collaboration, "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC", *Phys.Lett.* **B716** (2012) 30–61, doi:10.1016/j.physletb.2012.08.021, arXiv:1207.7235.

[2] ATLAS Collaboration, "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC", *Phys.Lett.* **B716** (2012) 1–29, doi:10.1016/j.physletb.2012.08.020, arXiv:1207.7214.

[3] CDF and D0 Collaboration, "Evidence for a particle produced in association with weak bosons and decaying to a bottom-antibottom quark pair in Higgs boson searches at the Tevatron", *Phys.Rev.Lett.* **109** (2012) 071804, doi:10.1103/PhysRevLett.109.071804, arXiv:1207.6436.

[4] CMS Collaboration Collaboration, "Evidence for the direct decay of the 125 GeV Higgs boson to fermions", *Nature Phys.* **10** (2014) doi:10.1038/nphys3005, arXiv:1401.6527.

[5] T. A. collaboration, "Evidence for Higgs Boson Decays to the $\tau^+\tau^-$ Final State with the ATLAS Detector",.

[6] Tevatron Electroweak Working Group Collaboration, "Combination of CDF and D0 results on the mass of the top quark using up to 9.7 fb^{-1} at the Tevatron", arXiv:1407.2682.

[7] ATLAS Collaboration, CDF Collaboration, CMS Collaboration, D0 Collaboration Collaboration, "First combination of Tevatron and LHC measurements of the top-quark mass", arXiv:1403.4427.

[8] CMS Collaboration, "Search for the associated production of the Higgs boson with a top-quark pair", *JHEP* **09** (2014) 087, doi:10.1007/JHEP09(2014)087, arXiv:1408.1682.

[9] ATLAS collaboration Collaboration, "Search for H to photons produced in association with top quarks and constraints on the top quark-Higgs boson Yukawa coupling using data taken at 7 TeV and 8 TeV with the ATLAS detector", *ATLAS Conference Note* **ATLAS-CONF-2014-043** (2014).

[10] ATLAS and CMS Collaboration, "Combination of single top-quark cross-sections measurements in the t-channel at sqrt[s]=8 TeV with the ATLAS and CMS experiments", *ATLAS and CMS Conference Note* **CMS-PAS-TOP-12-002, ATLAS-COM-CONF-2013-061, ATLAS-CONF-2013-098** (2013).

[11] CDF Collaboration, D0 Collaboration Collaboration, "Observation of s-channel production of single top quarks at the Tevatron", *Phys.Rev.Lett.* **112** (2014) 231803, doi:10.1103/PhysRevLett.112.231803, arXiv:1402.5126.

[12] CMS Collaboration, "Observation of the associated production of a single top quark and a W boson in pp collisions at sqrt(s) = 8 TeV", *Phys.Rev.Lett.* **112** (2014) 231802, doi:10.1103/PhysRevLett.112.231802, arXiv:1401.2942.

[13] M. Farina et al., "Lifting degeneracies in Higgs couplings using single top production in association with a Higgs boson", *JHEP* **05** (2013) 022, doi:10.1007/JHEP05(2013)022, arXiv:1211.3736.

[14] R. Frederix, S. Frixione, F. Maltoni, and T. Stelzer, "Automation of next-to-leading order computations in QCD: The FKS subtraction", *JHEP* **0910** (2009) 003, doi:10.1088/1126-6708/2009/10/003, arXiv:0908.4272.

[15] V. Hirschi et al., "Automation of one-loop QCD corrections", *JHEP* **05** (2011) 044, doi:10.1007/JHEP05(2011)044, arXiv:1103.0621.

[16] R. Frederix et al., "Four-lepton production at hadron colliders: aMC@NLO predictions with theoretical uncertainties", *JHEP* **1202** (2012) 099, doi:10.1007/JHEP02(2012)099, arXiv:1110.4738.

[17] A. Djouadi and G. Moreau, "The couplings of the Higgs boson and its CP properties from fits of the signal strengths and their ratios at the 7+8 TeV LHC", arXiv:1303.6591.

[18] S. Biswas, E. Gabrielli, and B. Mele, "Single top and Higgs associated production as a probe of the Htt coupling sign at the LHC", *JHEP* **01** (2013) 088, doi:10.1007/JHEP01(2013)088, arXiv:1211.0499.

[19] S. Biswas, E. Gabrielli, F. Margaroli, and B. Mele, "Direct constraints on the top-Higgs coupling from the 8 TeV LHC data", *JHEP* **07** (2013) 073, doi:10.1007/JHEP07(2013)073, arXiv:1304.1822.

[20] J. Aguilar-Saavedra, "Top flavor-changing neutral interactions: Theoretical expectations and experimental detection", *Acta Phys.Polon.* **B35** (2004) 2695–2710, arXiv:hep-ph/0409342.

[21] ATLAS Collaboration, "Search for flavour changing neutral currents in top quark decays top to charm and Higgs, with Higgs to photons, and limit on the tcH coupling with the ATLAS detector at the LHC", *ATLAS Conference Note* **ATLAS-CONF-2013-081** (2012).

[22] CMS Collaboration, "A search for anomalous production of events with three or more leptons using 8 TeV data", *CMS Conference Note* **CMS-SUS-13-002** (2013).

[23] J. Aguilar-Saavedra, R. Benbrik, S. Heinemeyer, and M. Perez-Victoria, "A handbook of vector-like quarks: mixing and single production", *Phys.Rev.* **D88** (2013) 094010, doi:10.1103/PhysRevD.88.094010, arXiv:1306.0572.

[24] ATLAS Collaboration, "Search for heavy top-like quarks decaying to a Higgs boson and a top quark in the lepton plus jets final state in pp collisions at sqrt(s) = 8 TeV with the ATLAS detector", *ATLAS Conference Note* **ATLAS-CONF-2013-018** (2012).

[25] CMS Collaboration, "Inclusive search for a vector-like T quark by CMS", *CMS Conference Note* **CMS-B2G-12-015** (2012).

[26] "https://twiki.cern.ch/twiki/bin/view/CMS/PdmV2012Analysis#22_JanReReco",.

[27] "https://twiki.cern.ch/twiki/bin/view/CMS/LumiCalc#lumiCalc2_py_and_pixel_LumiCalc",.

[28] J. Aldwall et al., "MadGraph 5: going beyond", *JHEP* **06** (2011) 128, doi:10.1007/JHEP06(2011)128, arXiv:hep-ph/1106.0522.

[29] T. Sjöstrand, S. Mrenna, and P. Skands, "PYTHIA 6.4 physics and manual", *JHEP* **05** (2006) 026, doi:10.1088/1126-6708/2006/05/026, arXiv:hep-ph/0603175.

[30] CMS Collaboration, "Search for associated production of a single top quark and a Higgs boson in events where the Higgs boson decays to two photons at $\sqrt{s}$ = 8 TeV", CMS Physics Analysis Summary CMS-PAS-HIG-14-001, 2014.

[31] "https://twiki.cern.ch/twiki/bin/view/CMS/PileupInformation",.

[32] C. Botta et al., "Search for the standard model Higgs boson produced in association with top quarks and decaying to leptons", *CMS AN* **2013/159** (2013).

[33] N. Amapane et al., "Search for the standard model Higgs boson in the decay channel $H \to ZZ \to 4\ell$ in pp collisions", *CMS AN* **2012/141** (2012).

[34] N. Amapane et al., "Updated results on the new boson discovered in the search for the standard model Higgs boson in the $H \to ZZ \to 4\ell$ channel in pp collisions", *CMS AN* **2012/367** (2012).

[35] CMS Collaboration, "Commissioning of the particle-flow event reconstruction with leptons from J/Psi and W decays at 7 TeV", *CMS Physics Analysis Summary* **CMS-PAS-PFT-10-003** (2010).

[36] A. Bodek and J. Han, "Improved Rochester Misalignment and Muon Scale Corrections Extracted for 2011A, 2011B CMS Data", CMS Analysis Note CMS-AN-2012-298, 2012.

[37] J. Brochero et al., "Search for the Higgs Boson in the Fully Leptonic WW Final State", CMS Analysis Note CMS-AN-2011-201, 2011.

[38] J. Brochero et al., "Search for the Higgs Boson Decaying to WW in the Fully Leptonic Final State and A First Measurement of the WW Cross-Section at 8 TeV", CMS Analysis Note CMS-AN-2012-194, 2012.

[39] J. Brochero et al., "Higgs Boson Decaying to WW in the Fully Leptonic Final State", CMS Analysis Note CMS-AN-2012-378, 2012.

[40] J. Brochero et al., "Higgs Boson Decaying to WW in the Fully Leptonic Final State using 2011 and 2012 Data", CMS Analysis Note CMS-AN-2013-022, 2013.

[41] CMS Collaboration, "CMS TDR. Vol. I", Technical Report CERN/LHCC 2006-001.

[42] CMS Collaboration, "CMS TDR. Vol. II", *J. Phys. G: Nucl. Part.Phys.* **34** (2007) 995, doi:10.1088/0954-3899/34/6/S01.

[43] "http://www-cdf.fnal.gov/ jyhan/cms_momscl/cms_rochcor_manual.html",.

[44] "https://twiki.cern.ch/twiki/bin/view/CMS/ElectronMomentumRegression",.

[45] G. Bauer et al., "Update on the associate Higgs (WH) production search with H$\to W^+W^-$ in the 3 leptons final state at 7 and 8 TeV", CMS Analysis Note CMS-AN-2013-019, 2013.

[46] "https://twiki.cern.ch/twiki/bin/view/CMS/MultivariateElectronIdentification",.

[47] "https://twiki.cern.ch/twiki/bin/viewauth/CMS/TagAndProbeForHIG",.

[48] "https://twiki.cern.ch/twiki/bin/view/CMSPublic/WorkBookMetAnalysis",.

[49] P. Harris et al., "Pileup Jet Identification", CMS Analysis Note CMS-AN-2013-186, 2013.

[50] "https://twiki.cern.ch/twiki/bin/viewauth/CMS/BTagSFMethods",.

[51] H. Bakhshian et al., "Computing the contamination from fakes in leptonic final states", CMS Analysis Note CMS-AN-2010-261, 2010.

[52] CMS Collaboration, "CMS Luminosity Based on Pixel Cluster Counting - Summer 2013 Update", CMS Physics Analysis Summary CMS-PAS-LUM-13-001, 2013.

[53] "https://twiki.cern.ch/twiki/bin/view/CMS/PileupSystematicErrors",.

[54] "https://twiki.cern.ch/twiki/bin/viewauth/CMS/MuonReferenceEffs",.

[55] "https://indico.cern.ch/event/259303/contribution/1/material/slides/0.pdf",.

[56] "https://twiki.cern.ch/twiki/bin/view/LHCPhysics/CrossSections",.

[57] J. M. Campbell and R. K. Ellis, "$t\bar{t}W^{\pm}$ production and decay at NLO", *JHEP* **07** (2012) 052, doi:doi:10.1007/JHEP07(2012)052, arXiv:1204.5678.

[58] M. Garzelli, A. Kardos, C. Papadopoulos, and Z. Trocsanyi, "$t\bar{t}W^{\pm}$ and $t\bar{t}Z$ Hadroproduction at NLO accuracy in QCD with Parton Shower and Hadronization effect", *JHEP* **11** (2012) 056, doi:doi:10.1007/JHEP11(2012)056, arXiv:1208.2665.

[59] Y. Takahashi et al., "earch for associated production of a single top quark and a Higgs boson in events where the Higgs boson decays into tau lepton pairs", CMS Analysis Note CMS-AN-2014-179, 2011.

[60] "https://twiki.cern.ch/twiki/bin/view/CMS/SWGuideHiggsAnalysisCombinedLimit",.

[61] A. Read, " Presentation of search results: the CL_s technique ", *Journal of Physics G.: Nuclear and Particle Physics* **28** (2002) 2693, doi:doi:10.1088/0954-3899/28/10/313.

[62] T. Junk, "Confidence level computation for combining searches with small statistics", *Nucl.Instrum.Meth.* **A434** (1999) 435–443, doi:10.1016/S0168-9002(99)00498-2, arXiv:hep-ex/9902006.

[63] Y. Takahashi, J. Steggemann, and R. Manzoni, "Search for associated production of a single top quark and a Higgs boson in events where the Higgs boson decays into tau lepton pairs", *CMS AN* **2014/147** (2014).

[64] CMS Collaboration, "Search for the standard model Higgs boson produced in association with top quarks in multilepton final states", *CMS Public Analysis Summary* **CMS-PAS-HIG-13-020** (2013).

A Additional Three-Lepton Plots

In the following, some distribution for the most characteristic variables used in the analysis are shown, at some loose pre-selection of the events:

- Invariant mass of the event, $m_{ll} > 20$ GeV. The invariant mass used is the one that, reconstructed with two opposite sign leptons, has a closer value to the Z boson mass
- Maximum p_T of the three leptons higher than 20 GeV. This cut is applied to match the trigger threshold of the double lepton triggers used to select data events
- No extra tight leptons selected in the event (i.e., only three tight leptons)
- Missing E_T, with PF algorithm, higher than 30 GeV

The expected number of event and the data measurement for these requirements are summarized in Table 44. Diboson corresponds to the sum of WZ, WW, and ZZ backgrounds. WVV+Drell-Yan is the sum of WWW, WWZ, WZZ (triboson), and Drell-Yan backgrounds. The fakes background is mainly $t\bar{t}$, and its estimation from a data-driven method is explained in Section 5.

Sample	all channels
qtH→WW	2.92 ± 0.04
qtH→ $\tau\tau$	0.94 ± 0.02
WtH→WW	0.76 ± 0.02
WtH→ $\tau\tau$	0.28 ± 0.02
Diboson	1496 ± 3
WVV+Drell-Yan	110 ± 15
$t\bar{t}W^{\pm}$	16.7 ± 0.6
$t\bar{t}Z$	29.9 ± 0.6
Fakes ($t\bar{t}$)	512 ± 23
Total Background	2173 ± 25
Data	2256

Table 44: Data yields and expected background in 19.7 fb^{-1} for the **3 lepton final state** preselection. Statistical errors only.

In Figures 41 and 42 some distribution for the main variables are shown.

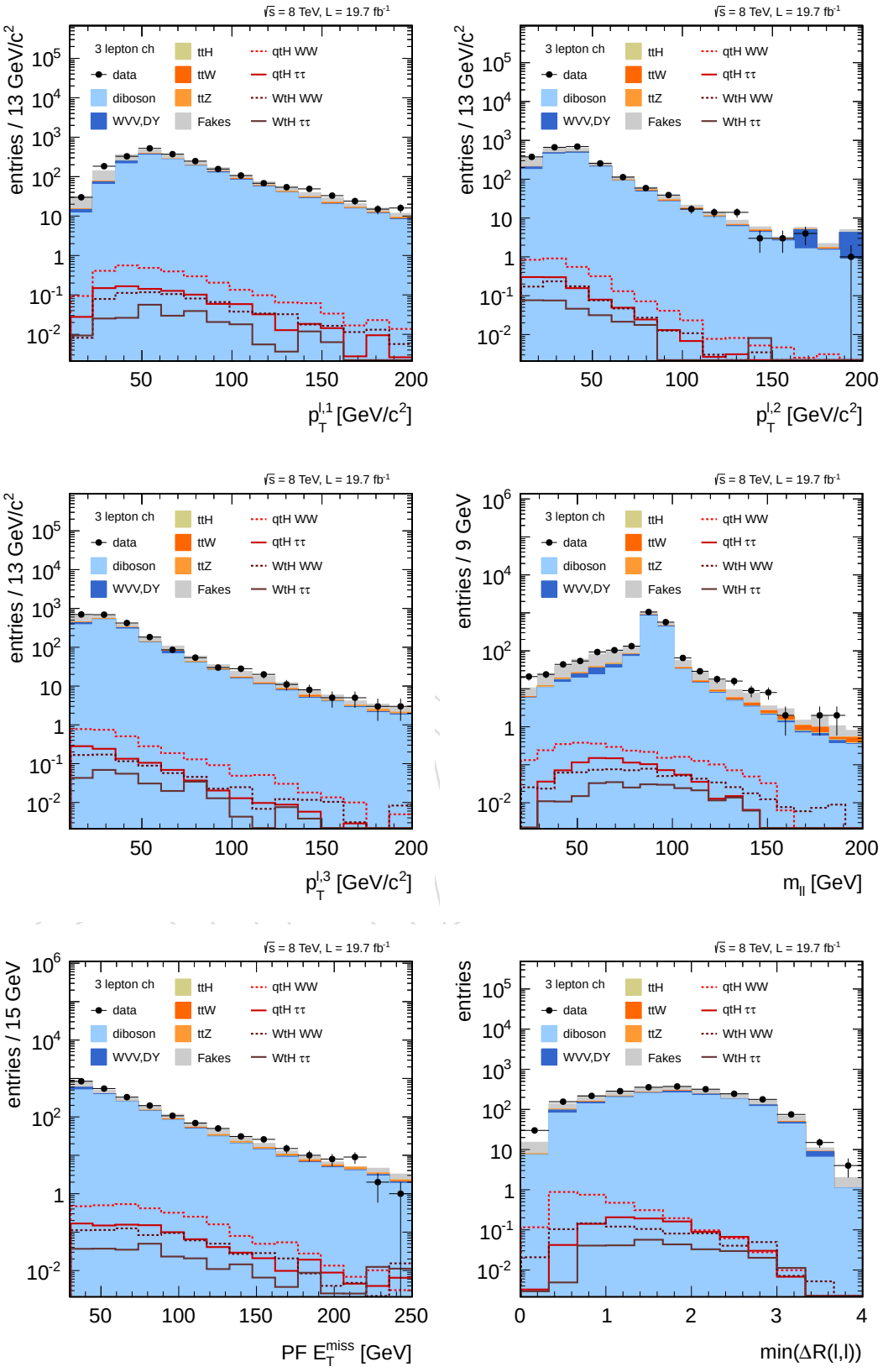

Figure 41: Distribution at a loose event pre-selection for the **3 lepton final state**.

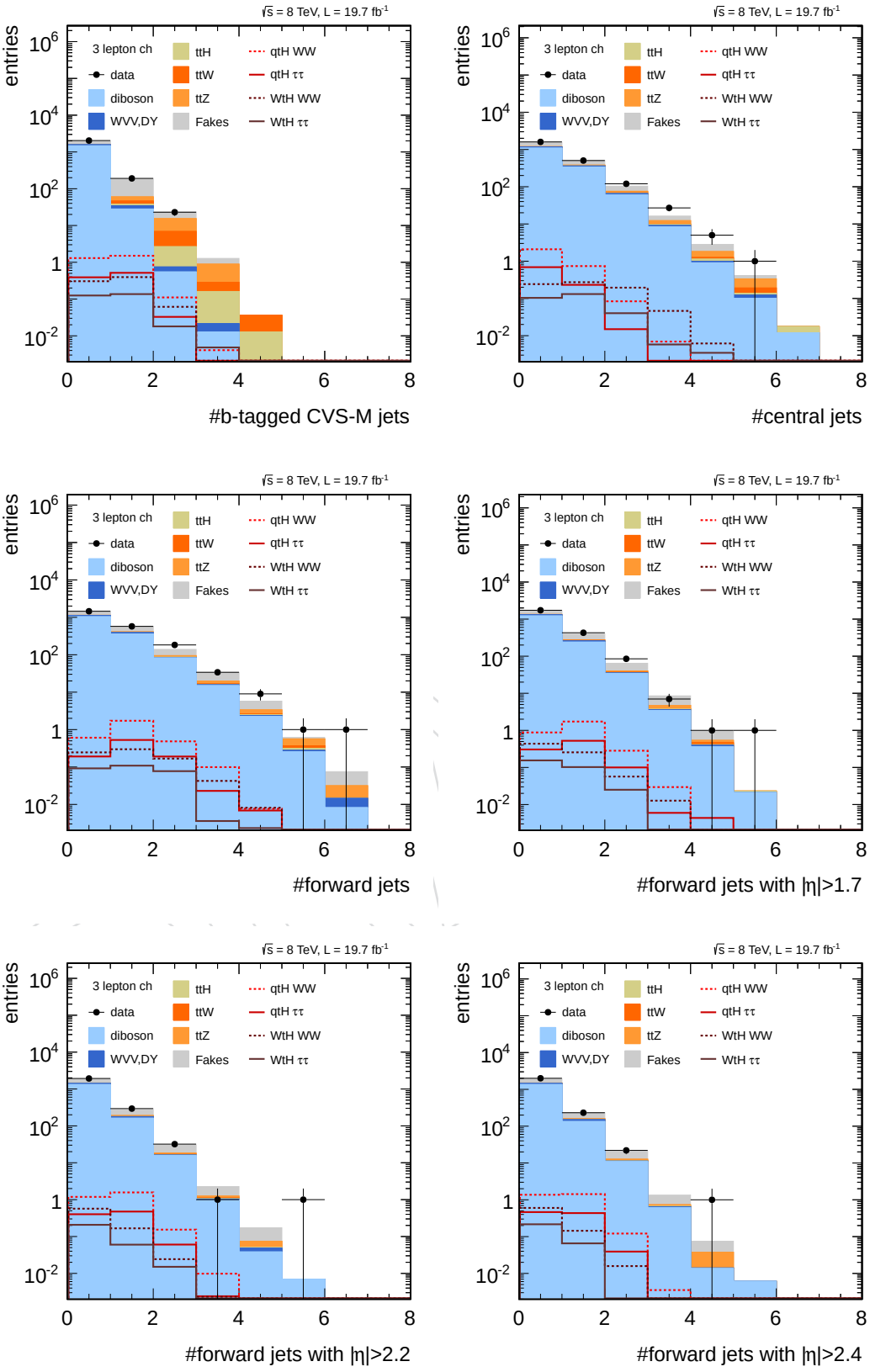

Figure 42: Distribution at a loose event pre-selection for the **3 lepton final state**.

B Additional Three-Lepton Tables

In this Appendix a more detailed table with the expected yields and observed data by sub-channels is shown, in the case of the three lepton final state, for the tight pre-selection used to estimate the final likelihood discriminator. Table 45 shows the data yields and expected number of events for signal and backgrounds processes, for the four channels considered for the three lepton final state, separately: $\mu\mu\mu$, $\mu\mu$e, eeμ, and eee. The total number of events, i.e. the sum of these four channels, is shown in Table 17.

Process	$\mu\mu\mu$	$\mu\mu$e	eeμ	eee
WZ/ZZ/WW	0.53 ± 0.04	0.38 ± 0.04	0.16 ± 0.02	0.113 ± 0.016
WVV+Drell-Yan	0.019 ± 0.011	0.05 ± 0.02	0.041 ± 0.014	0.002 ± 0.002
$t\bar{t}Z$	0.64 ± 0.10	0.62 ± 0.09	0.67 ± 0.10	0.28 ± 0.06
$t\bar{t}W^{\pm}$	0.52 ± 0.11	1.65 ± 0.19	0.70 ± 0.12	0.16 ± 0.06
$t\bar{t}H$	1.52 ± 0.06	0.63 ± 0.04	0.43 ± 0.03	0.120 ± 0.016
Fakes	8.0 ± 2.8	16.1 ± 4.0	9.2 ± 3.0	0.55 ± 0.74
Total Background	$\mathbf{10.1 \pm 2.8}$	$\mathbf{19.5 \pm 4.0}$	$\mathbf{11.2 \pm 3.0}$	$\mathbf{1.2 \pm 0.7}$
tH($\tau\tau$)W	0.008 ± 0.002	0.024 ± 0.004	0.034 ± 0.007	0.001 ± 0.001
tH(WW)W	0.040 ± 0.005	0.066 ± 0.006	0.074 ± 0.009	0.014 ± 0.004
tH($\tau\tau$)q	0.061 ± 0.005	0.141 ± 0.008	0.084 ± 0.006	0.020 ± 0.003
tH(WW)q	0.168 ± 0.009	0.413 ± 0.014	0.323 ± 0.012	0.048 ± 0.004
Total Signal	$\mathbf{0.277 \pm 0.012}$	$\mathbf{0.644 \pm 0.018}$	$\mathbf{0.514 \pm 0.017}$	$\mathbf{0.083 \pm 0.006}$
Data	**9**	**18**	**9**	**6**

Table 45: Data yields and expected number of signal and background events for the likelihood pre-selection for the separated 3 lepton final state channels: $\mu\mu\mu$, $\mu\mu$e, eeμ, and eee, in 19.7 fb^{-1}. Statistical errors only

C Additional Same-Sign Dilepton Plots

C.1 Background Control Region Plots

We present some additional plots in a background dominated control region, selected by inverting the pre-selection requirement of a forward jet:

- Two leptons of equal charge, $p_{\mathrm{T}} > 20\,\mathrm{GeV}$
- No additional leptons (selected with lepton MVA > 0.35)
- Invariant mass of lepton pairs $> 12\,\mathrm{GeV}$
- At least one central jet ($|\eta| < 1.0$)
- No forwards jets ($|\eta| > 1.0$)
- At least one central jet tagged as CSV loose

Figures 43 and 44 show jet multiplicities, H_{T}, b-tagged jet multiplicities and $E_{\mathrm{T}}^{\mathrm{miss}}$ for the $\mu\mu$ and $e\mu$ channel in this selection. As can be seen, when the selection approaches that of the leptonic $t\bar{t}H$ analysis (at least two medium b-tagged jets and at least four jets), the excess in data goes down in the $e\mu$ channel, but remains in the $\mu\mu$ channel, consistent with what was seen in that analysis [32].

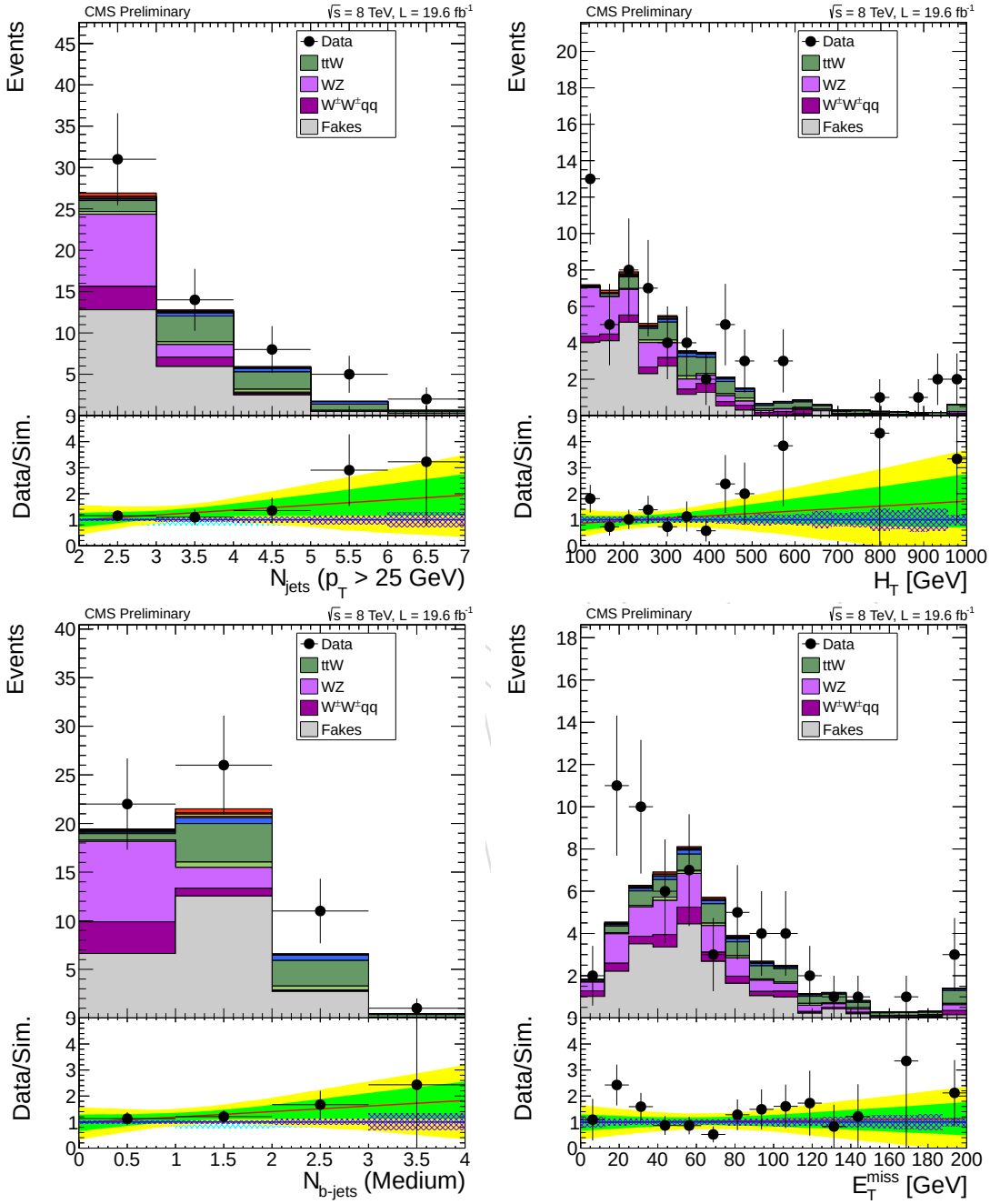

Figure 43: Additional plots for the **same-sign dilepton** channel; pre-selection with veto on forward jets. Observed data vs. predicted background for the $\mu\mu$ channel. Jet multiplicity (top left), scalar sum of jet p_T's (H_T) (top right), medium b-tagged jet multiplicity (bottom left), missing transverse energy (bottom right).

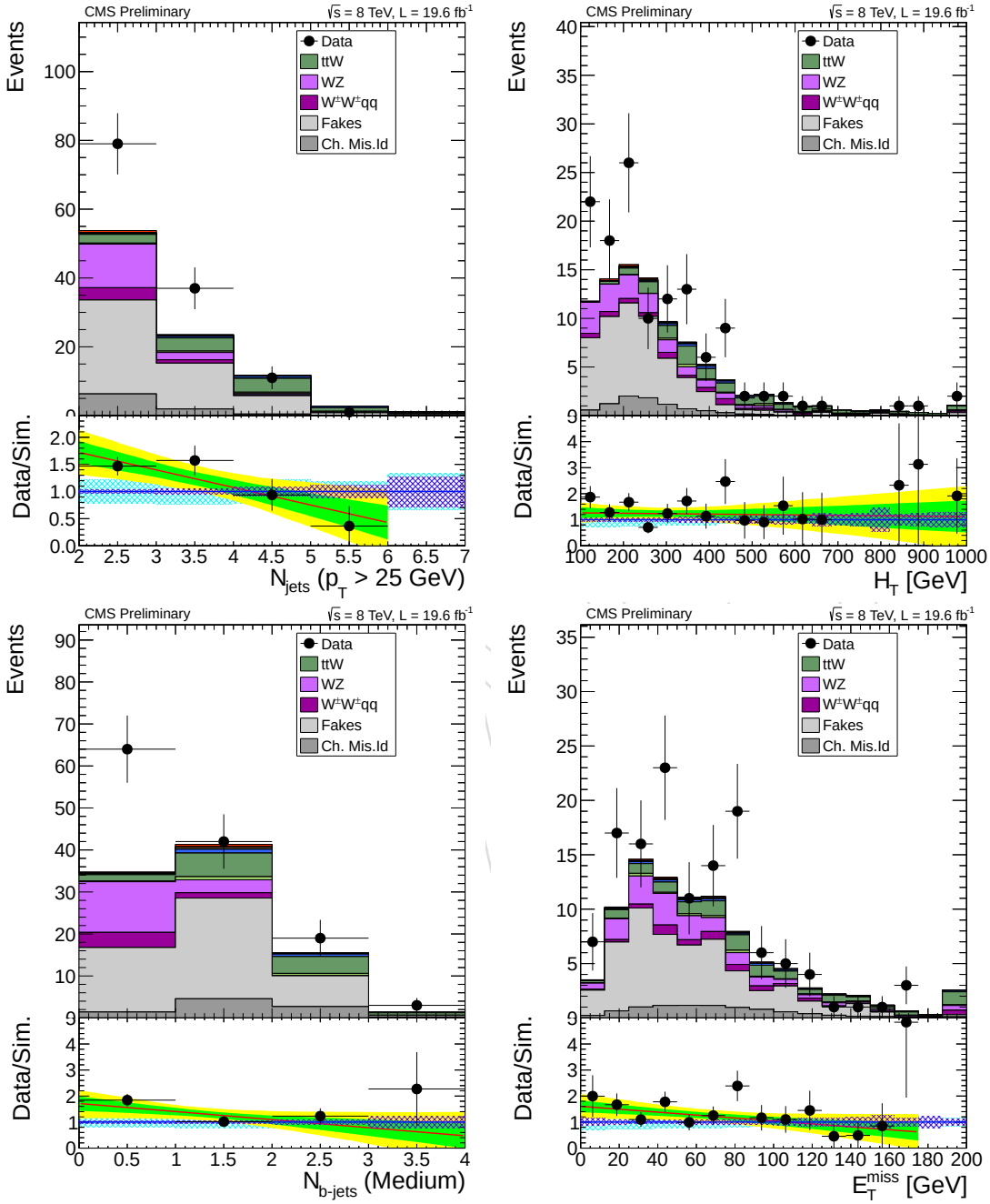

Figure 44: Additional plots for the **same-sign dilepton** channel; pre-selection with veto on forward jets. Observed data vs. predicted background for the $e\mu$ channel. Jet multiplicity (top left), scalar sum of jet p_T's (H_T) (top right), medium b-tagged jet multiplicity (bottom left), missing transverse energy (bottom right).

C.2 Cross-Check with Official POG ID

An additional quick cross-check on the excess in the $\mu\mu$ channel is provided by observing the number of muons in our pre-selection failing the official POG tight identification criteria. The lepton MVA used in this analysis is not tuned to discriminate fakes from light jets as well as the official POG ID. Hence if the excess would be predominantly in events where one of the two leptons fails the POG ID it would indicate a problem with light jet fakes. This is not the case, as can be seen in Fig. 45.

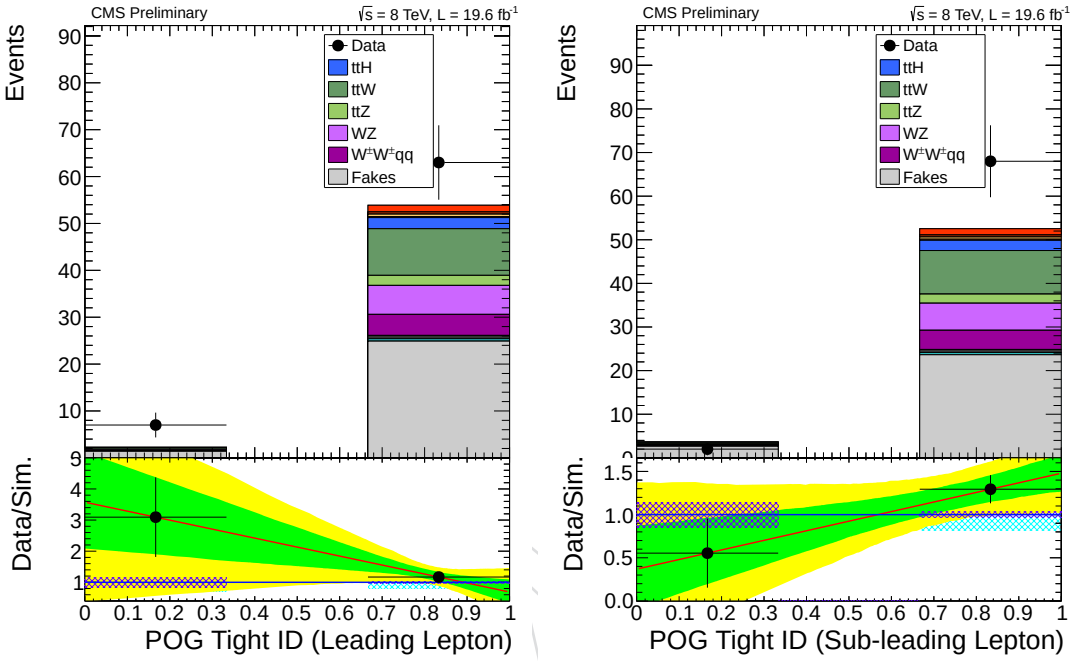

Figure 45: Passing or failing the official POG tight ID for events in our pre-selection ($\mu\mu$ channel). More than 90% of muons *pass* the tight ID, and the excess is mainly coming from events where the leptons passed the ID.

D Systematic Shape Plots

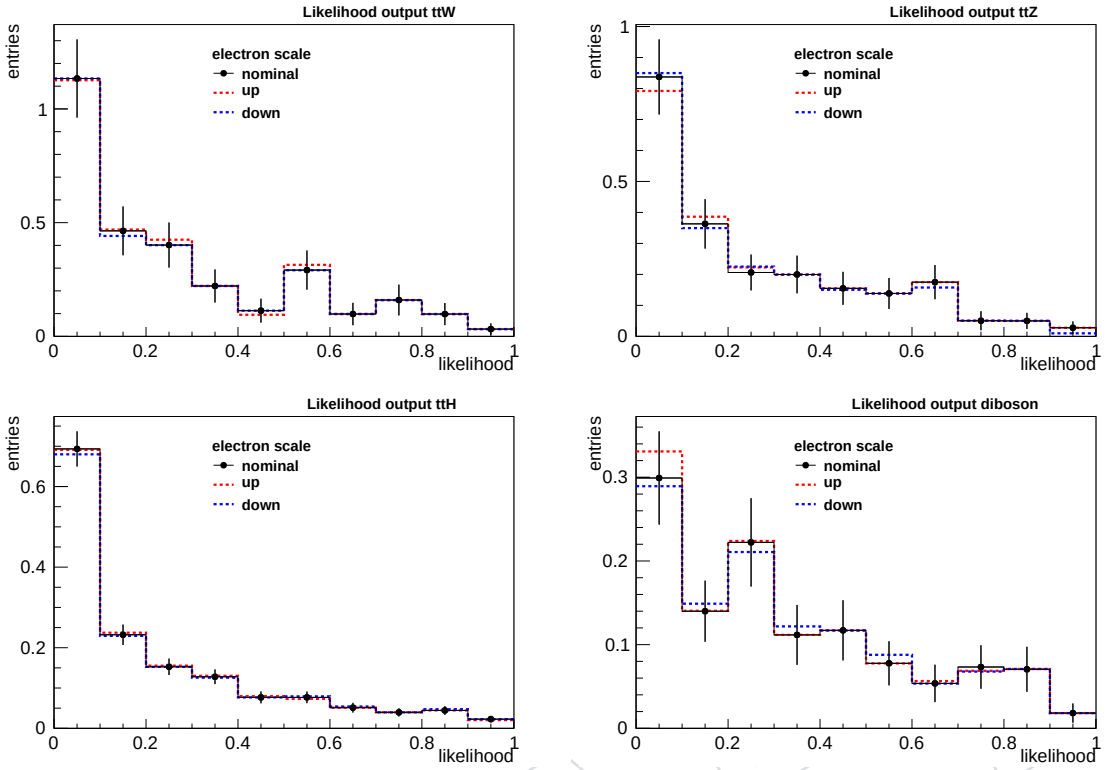

Figure 46: Nominal distribution for the likelihood, for the **3 lepton final state**, compared with the up/down variation due to electron energy scale. From left to right, up to bottom, for ttW, ttZ, ttH, and diboson backgrounds

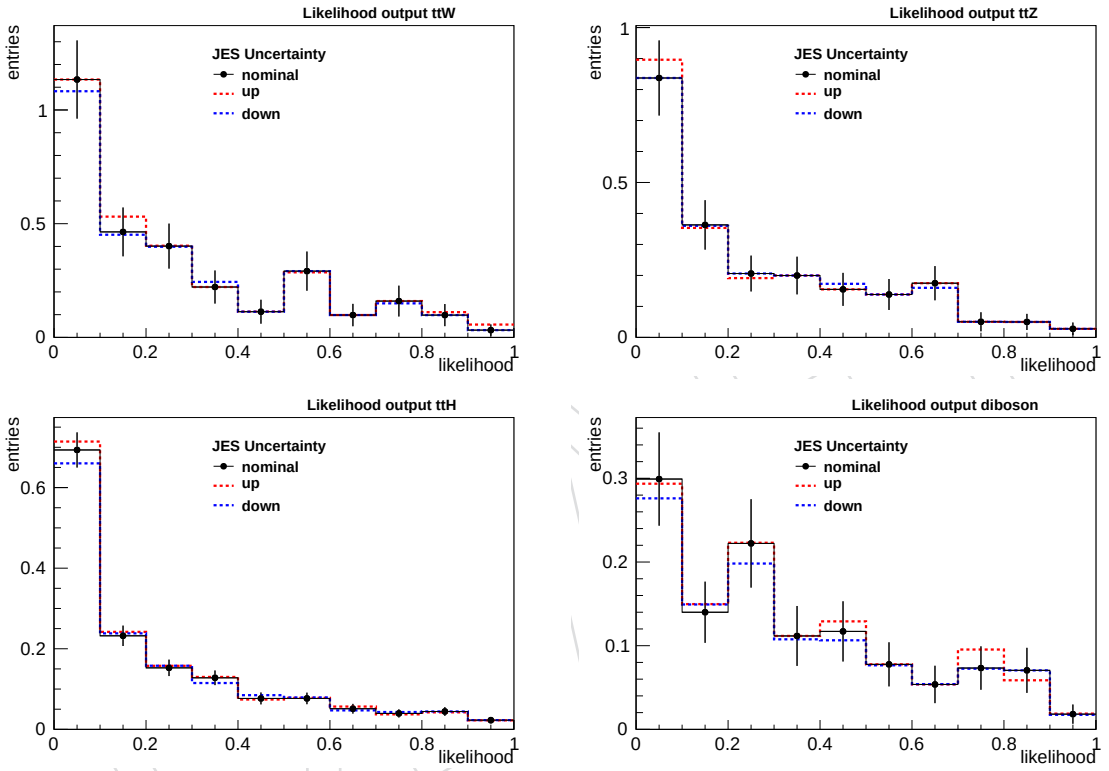

Figure 47: Nominal distribution for the likelihood, for the **3 lepton final state**, compared with the up/down variation due to the Jet Energy Corrections. From left to right, up to bottom, for ttW, ttZ, ttH, and diboson backgrounds

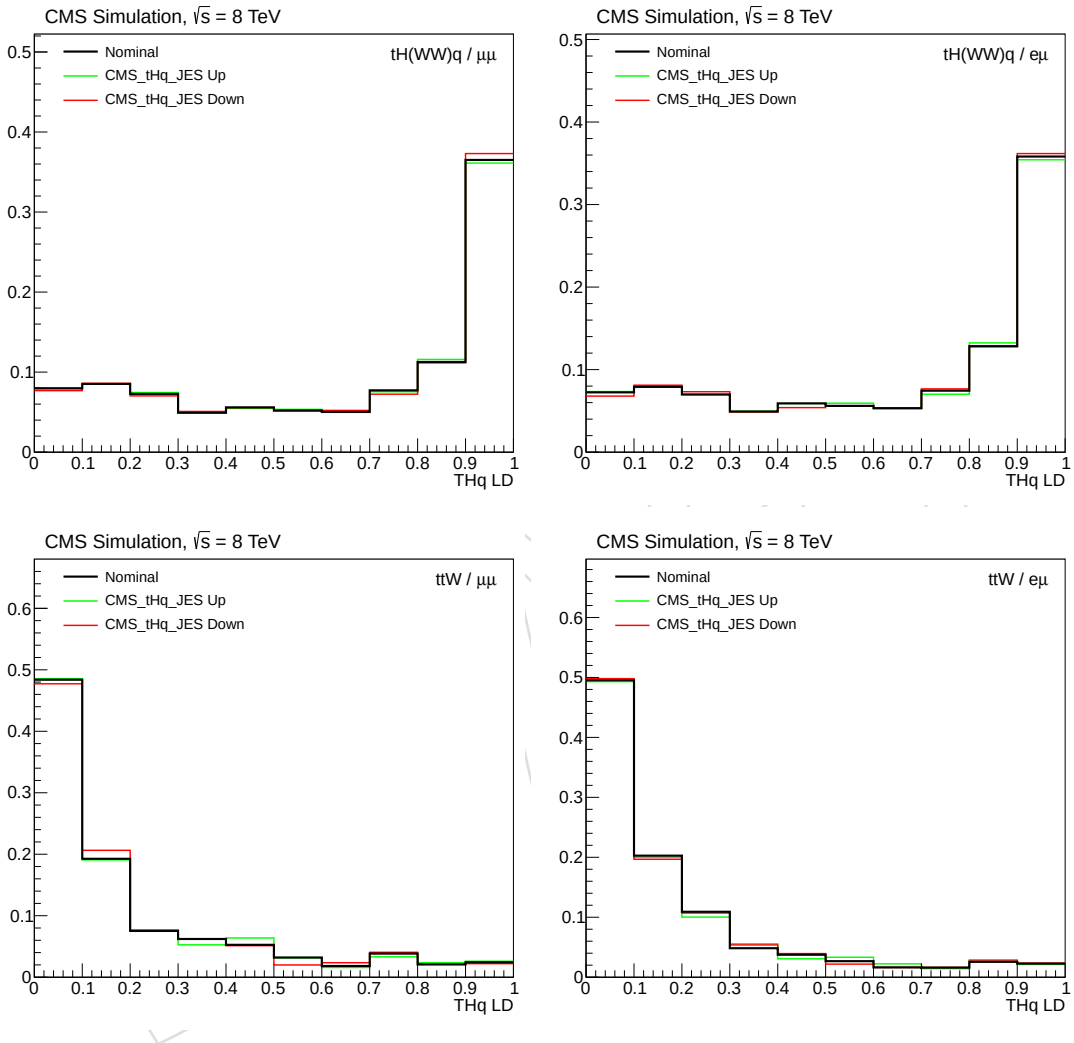

Figure 48: Uncertainty on the linear likelihood discriminant shape from jet energy scale uncertainties in the **same-sign dilepton** channels, for the $\mu\mu$ (left) and $e\mu$ (right) channels, for tH(WW)q signal (top) and $t\bar{t}W^{\pm}$ background (bottom).

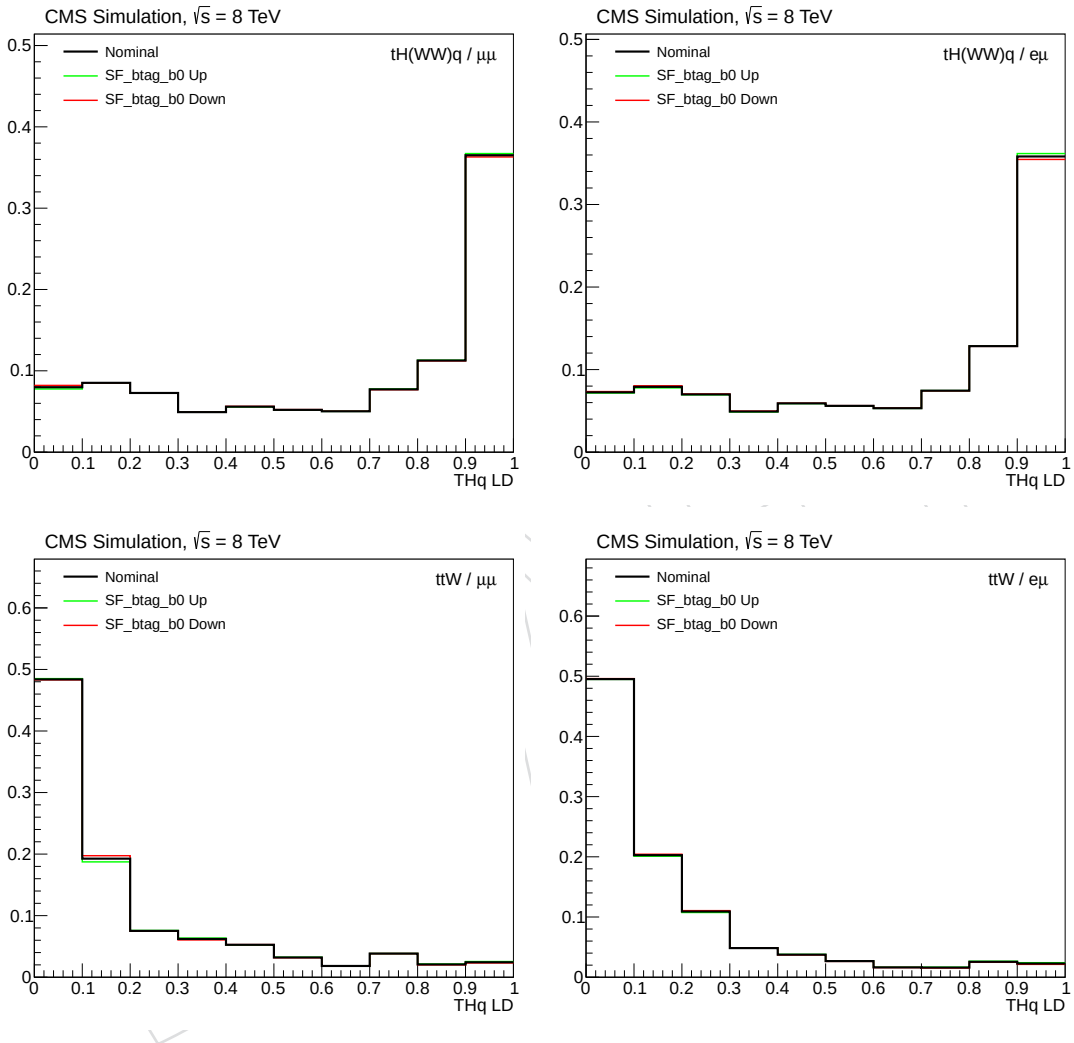

Figure 49: Uncertainty on the linear likelihood discriminant shape from b-tagging scale factor uncertainties in the **same-sign dilepton** channels, for the $\mu\mu$ (left) and $e\mu$ (right) channels, for tH(WW)q signal (top) and $t\bar{t}W^{\pm}$ background (bottom).

E Likelihood in Background Dominated Control Selection

We unblind the likelihood outputs in sub-regions of our pre-selection that are dominated by signal: at least four hadronic jets in the same-sign dilepton channel, no jets with pseudo-rapidity greather than 2.4 in the trilepton channel. The cuts have signal efficiencies of about 10%, and 20% in the same-sign and trilepton channel, respectively. As expected, the likelihood output is strongly peaking towards the background-only side. A good agreement between data and the predicted yields and shape is observed, see Figures 50 and 51

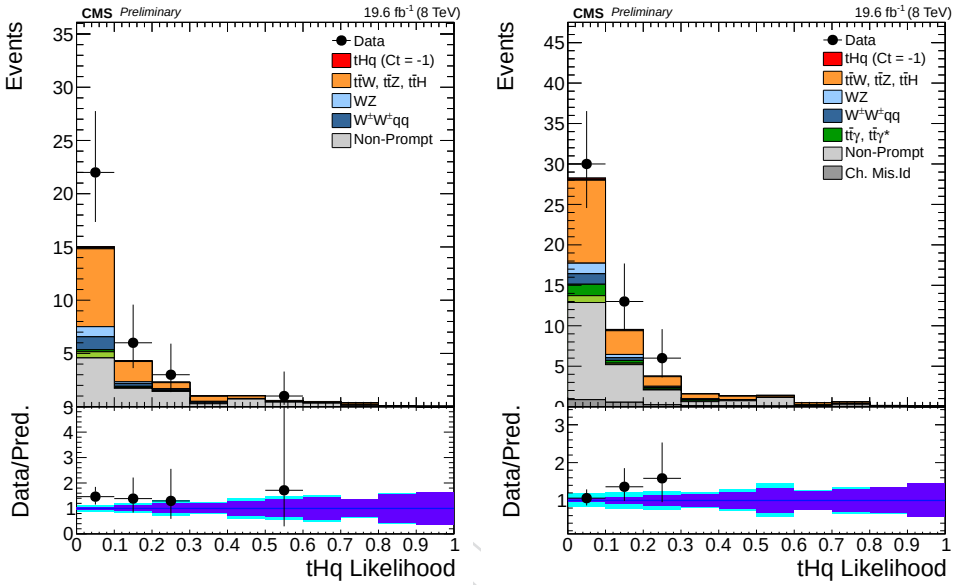

Figure 50: Likelihood outputs in a sub-region of the **same-sign dilepton** pre-selection with at least four hadronic jets (about 10% efficient for signal events). Same-sign $\mu\mu$ channel (left), same-sign $e\mu$ channel (right).

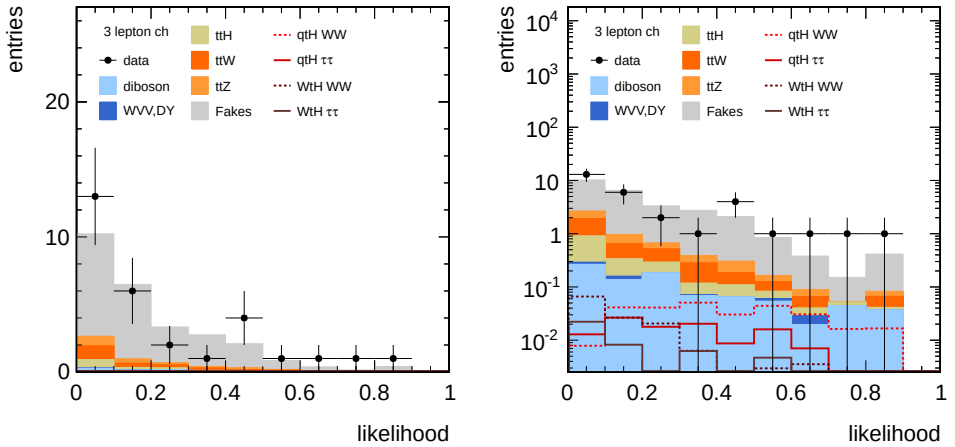

Figure 51: Likelihood outputs in a sub-region of the **three lepton** pre-selection with at no hadronic jet with $|\eta| > 2.4$ (about 20% efficient for signal events). Linear y-axis scale (left), logarithmic scale (right).

F Vetoing Hadronic Taus

To expedite a future combination with a dedicated analysis targeting $tH(\tau\tau)q$ in final states with same-sign dileptons and a hadronic tau [63], we modify our pre-selection to reject events falling into that selection. We therefore veto events that have one or more tau candidates passing the selection described below, following closely what is applied in [63].

F.1 Tau Selection and Efficiencies

Tau candidates are selected if they have $p_T > 20\,\text{GeV}$, $|\eta| < 2.3$, $d_z < 0.2\,\text{cm}$, and are not within $\Delta R < 0.5$ of any pre-selected lepton (i.e. passing lepton selection before the lepton MVA cut.) Furthermore, tau candidates are required to pass the following algorithms:

- `"againstMuLoose"`
- `"againstElectronLooseMVA3"`
- `"byLooseCombinedIsolationDeltaBetaCorr"`

Vetoing events with one or more such selected tau candidate rejects about 33% of $tH(\tau\tau)q$ events passing the pre-selection, equally for the $e\mu$ and $\mu\mu$ channels. Of the larger $tH(WW)q$ signal, about 4.4% are rejected, resulting in an overall signal efficiency of about 88.1%. In data, about 7% of events are rejected.

The remaining overlap of the event selection with respect to the signal events selected in [63] has been checked and found to be 0.15% of our selection (corresponding to 0.7% of their selection). Before applying the veto, the overlap was about 6% of our selection and 33% of their selection.

The resulting yields for all processes in all three channels can be found in Tab. 46 below (to be compared with Tab. 18, on page 38).

Process	$\mu\mu$	$e\mu$	ee
W^+W^+qq	4.62 ± 0.48	6.05 ± 0.53	2.57 ± 0.33
WZ	5.38 ± 0.24	8.71 ± 0.30	4.38 ± 0.21
$t\bar{t}\gamma^*$	0.50 ± 0.05	1.04 ± 0.08	0.58 ± 0.06
$t\bar{t}\gamma$	0.09 ± 0.05	2.02 ± 0.24	1.49 ± 0.21
$t\bar{t}Z$	2.23 ± 0.21	2.87 ± 0.23	1.29 ± 0.15
$t\bar{t}W^\pm$	10.23 ± 0.50	14.95 ± 0.59	5.22 ± 0.34
$t\bar{t}H$	2.26 ± 0.08	3.24 ± 0.09	1.15 ± 0.05
Rare SM bkg.	1.41 ± 0.09	2.58 ± 0.12	1.08 ± 0.07
Charge Mis-ID	–	6.99 ± 0.10	43.83 ± 0.31
Non-Prompt	23.01 ± 1.46	50.96 ± 2.47	30.83 ± 1.99
Total Background	$\mathbf{49.73 \pm 1.66}$	$\mathbf{99.41 \pm 2.64}$	$\mathbf{92.42 \pm 2.09}$
$tH(\tau\tau)W$	0.06 ± 0.01	0.07 ± 0.01	0.02 ± 0.00
$tH(WW)W$	0.16 ± 0.02	0.26 ± 0.02	0.08 ± 0.01
$tH(\tau\tau)q$	0.33 ± 0.02	0.50 ± 0.02	0.20 ± 0.01
$tH(WW)q$	1.41 ± 0.04	2.06 ± 0.04	0.72 ± 0.02
Total Signal	$\mathbf{1.95 \pm 0.04}$	$\mathbf{2.89 \pm 0.05}$	$\mathbf{1.03 \pm 0.03}$
Data	**66**	**117**	**92**

Table 46: Data yields and expected background for the likelihood pre-selection for the **same-sign di-lepton final state**, in 19.7 fb^{-1}. Statistical errors only.

F.2 Results with Tau Veto Applied

Results on the upper limit when applying the tau veto in the event selection are summarized in Table 47. Results for the three lepton final state remain the same, as shown in Table 47.

Channel	Observed	Expected	68% prob. band	95% prob. band
SS eμ	11.4	9.3	[7.0, 13.5]	[5.4, 18.8]
SS $\mu\mu$	9.3	8.1	[6.0, 11.8]	[4.7, 16.7]
3ℓ	11.5	8.6	[6.6, 12.4]	[5.7, 18.0]
combined	6.7	5.0	[3.6, 7.1]	[2.9, 10.3]

Table 47: Expected and Observed 95% C.L. upper limit on Higgs cross section for C_t = -1 on qtH production. The $\pm1\sigma$ and $\pm2\sigma$ probability bands are also shown.

F.3 Results without Tau Veto Applied

Results on the upper limit when **not applying** the tau veto in the event selection are summarized in Table 48. Results for the three lepton final state remain the same, as shown in Table 21. Results shown here are the same presented in Section 8, but repeated here for completiness and to facilitate the comparison.

Channel	Observed	Expected	68% prob. band	95% prob. band
SS eμ	12.9	9.7	[7.5, 13.2]	[5.7, 17.7]
SS $\mu\mu$	8.5	7.6	[6.0, 10.8]	[5.3, 15.4]
3ℓ	11.5	8.6	[6.6, 12.4]	[5.7, 18.0]
combined	6.9	5.1	[3.8, 7.0]	[2.8, 9.7]

Table 48: Expected and Observed 95% C.L. upper limit on Higgs cross section for C_t = -1 on qtH production. The $\pm1\sigma$ and $\pm2\sigma$ probability bands are also shown.

G Overlap With $t\bar{t}H$ Multilepton Search

We assess the overlap of selected events between the pre-selection regions of this analysis and the related search for associated production of $t\bar{t}H$ in multileptonic channels [32, 64]. Five channels were investigated in that analysis, with similar selection criteria to the ones in the present analysis: three same-sign dilepton channels ($\mu\mu$, $e\mu$, ee), a three-lepton channel, and a four-lepton channel. The objects used for the event selection are identical between the tHq di-lepton channels and all the $t\bar{t}H$ channels, whereas the objects of the tHq three-lepton channels differ slightly. Therefore, by construction, there is no migration between the two tHq di-lepton channels and any of the $t\bar{t}H$ channels (however there is of course an overlap between the two selections in the same channels). Between the tHq three-lepton selection, and the $t\bar{t}H$ channels there is a very slight migration observed.

Below we summarize the main differences between the event selections and the resulting overlaps between all the relevant channels.

Same-Sign Channels The $t\bar{t}H$ selection is aimed at higher jet and b-tag multiplicities, and has an additional cut on missing transverse momentum, as is summarized in table 49.

tHq selection	$t\bar{t}H$ selection
≥ 1 un-tagged forward jet ($\lvert\eta\rvert > 1$)	≥ 4 jets ($\lvert\eta\rvert < 2.4$)
≥ 1 central jet ($\lvert\eta\rvert < 1$)	$p_{\mathrm{T}}^{\ell 1} + p_{\mathrm{T}}^{\ell 2} + E_{\mathrm{T}}^{\mathrm{miss}} > 100$ GeV
≥ 1 CSV loose-tagged jet	≥ 2 CSV loose OR ≥ 1 CSV medium tagged jet

Table 49: Main selection differences in the same-sign dilepton channels between the $t\bar{t}H$ and tHq analyses.

Three Lepton Channel The $t\bar{t}H$ selection is aimed at higher b-tag multiplicities, and has an additional cut on missing transverse momentum, as is summarized in table 50. Note that $E_{\mathrm{T}}^{\mathrm{miss}}$LD is defined as $E_{\mathrm{T}}^{\mathrm{miss}} \times 0.00397 + \mathrm{MH_T} \times 0.00265$, where $\mathrm{MH_T}$ is the absolute value of the vectorial sum of all jet momenta.

tHq selection	$t\bar{t}H$ selection
≥ 1 un-tagged forward jet ($\lvert\eta\rvert > 1.5$)	≥ 2 jets ($\lvert\eta\rvert < 2.4$)
$= 1$ CSV loose-tagged jet	≥ 2 CSV loose OR ≥ 1 CSV medium tagged jet
$E_{\mathrm{T}}^{\mathrm{miss}} > 30$ GeV	$E_{\mathrm{T}}^{\mathrm{miss}}\mathrm{LD} > 0.2$ OR ≥ 4 jets

Table 50: Main selection differences in the three lepton channels between the $t\bar{t}H$ and tHq analyses.

Observed Overlaps We report the observed yield of events for the data, the tHq signal, and the $t\bar{t}H$ signal separately in tables 51,52, and 53. As is expected, there is a significant overlap of selected events, and the selection efficiencies for the respective signals are higher for the respective dedicated analysis. i.e. the looser selection in the tHq channel allows for more tHq signal events without admitting more $t\bar{t}H$ events.

	Total	$\mu\mu$ ($t\bar{t}H$)	$e\mu$ ($t\bar{t}H$)	3ℓ($t\bar{t}H$)
Total		911	1549	2549
$\mu\mu$ (tHq)	5026	669	0	0
$e\mu$ (tHq)	8003	0	1089	0
3ℓ(tHq)	2841	2	0	902

Table 51: Total selected and shared MC events between the three tHq channels and the relevant $t\bar{t}H$ channels for the tHq **signal sample**.

	Total	$\mu\mu$ ($t\bar{t}H$)	$e\mu$ ($t\bar{t}H$)	3ℓ($t\bar{t}H$)
Total		1177	1794	1994
$\mu\mu$ (tHq)	951	686	0	0
$e\mu$ (tHq)	1473	0	1058	0
3ℓ(tHq)	654	7	6	353

Table 52: Total selected and shared MC events between the three tHq channels and the relevant $t\bar{t}H$ channels for the $t\bar{t}H$ **signal sample**.

	Total	$\mu\mu$ ($t\bar{t}H$)	$e\mu$ ($t\bar{t}H$)	3ℓ($t\bar{t}H$)
Total		39	51	62
$\mu\mu$ (tHq)	66	19	0	0
$e\mu$ (tHq)	117	0	39	0
3ℓ(tHq)	42	0	0	14

Table 53: Total selected and shared **data** events between the three tHq channels and the relevant $t\bar{t}H$ channels for the.