The Compact Muon Solenoid Experiment

CMS Draft Note

Mailing address: CMS CERN, CH-1211 GENEVA 23, Switzerland

2018/07/04
Head Id: 467238
Archive Id: 467297
Archive Date: 2018/07/03
Archive Tag: trunk

Search for the associated production of a Higgs boson with top-quark pair in final states with electrons, muons, and τ_h with data recorded in 2017

S. Banerjee[5], S. Bhowmik[9], C. Botta[1], A. Carvalho[9], N. Chanon[13], R. K. Dewanjee[9], K. Ehatäht[9], C. Erice Cid[4], L. Gouskos[2], A. Gurrola[8], M. Kadastik[9], S. Karmakar[5], K. Lannon[7], B. Li[12], H. Liao[12], A. Nayak[10], C. M. Perez[6], M. Peruzzi[1], G. Petrucciani[1], H. Qu[2], A. Rinkevicius[11], F. Romeo[8], A. Ryd[11], S. Sanchez Cruz[4], S. Sawant[5], K. Skovpen[16], G. Smith[7], B. Stieger[15], T. Strebler[3], Z. Tao[11], J. Thomas-Wilsker[12], N. Tonon[14], C. Veelken[9], P. Vischia[4], A. Zabi[6], and H. Zhang[12]

[1] European Organization for Nuclear Research, Geneva, Switzerland
[2] University of California, Santa Barbara, USA
[3] Imperial College, London, U.K.
[4] Universidad de Oviedo, Principado de Asturias, Spain
[5] Tata Institute of Fundamental Research, Mumbai, India
[6] LLR Ecole Polytechnique, Palaiseau, France
[7] University of Notre Dame, Notre Dame, USA
[8] Vanderbilt University, Nashville, USA
[9] National Institute of Chemical Physics and Biophysics, Tallinn, Estonia
[10] Institute of Physics, Bhubaneswar, India
[11] Cornell University, Ithaca, USA
[12] Chinese Academy of Sciences, Beijing, China
[13] Institut de Physique Nucléaire de Lyon, Lyon, France
[14] Institut Pluridisciplinaire Hubert Curien, Strasbourg, France
[15] University of Nebraska-Lincoln, Lincoln, USA
[16] Vrije Universiteit Brussel, Brussels, Belgium

Abstract

Results of a search for the standard model Higgs boson produced in association with a top quark pair in final states with electrons, muons, and hadronic tau decays (τ_h) are presented. The analyzed dataset corresponds to an integrated luminosity of 41.5 fb^{-1}, recorded in proton-proton collisions at $\sqrt{s} = 13$ TeV center-of-mass energy in 2017. The sensitivity of the search is improved by using matrix element and machine learning methods to separate the signal from backgrounds. The results are combined with the ones obtained from 35.9 fb^{-1} of data collected in 2016. The measured observed (expected) signal rate amounts to $X.XX^{+Y.YY}_{-Z.ZZ}$ ($1.00^{+0.32}_{-0.28}$) times the production rate expected in the standard model, with an observed (expected) significance of $X.X\sigma$ (3.9σ).

This box is only visible in draft mode. Please make sure the values below make sense.

PDFAuthor:	TBD
PDFTitle:	Search for the associated production of Higgs boson with top-quark pair in final states with electrons, muons, and tauh with data recorded in 2017
PDFSubject:	CMS
PDFKeywords:	CMS, physics

Please also verify that the abstract does not use any user defined symbols

Contents

1 Introduction . . . 3
2 Data samples and Monte Carlo simulation . . . 4
3 Particle reconstruction and identification . . . 5
3.1 Electrons and muons . . . 7
3.2 Hadronic τ decays . . . 10
3.3 Jets . . . 12
4 Reconstruction of event level quantities . . . 13
5 Event selection . . . 13
5.1 Triggers . . . 13
5.2 E_T^{miss} filters . . . 14
5.3 $1\ell + 2\tau_h$ category . . . 14
5.4 2ℓss category . . . 15
5.5 $2\ell\text{ss} + 1\tau_h$ category . . . 15
5.6 3ℓ category . . . 16
5.7 $3\ell + 1\tau_h$ category . . . 16
5.8 4ℓ category . . . 16
5.9 $2\ell + 2\tau_h$ category . . . 17
6 Event categorization . . . 17
7 Data-to-MC corrections . . . 17
7.1 Pileup reweighting . . . 18
7.2 Trigger efficiency . . . 20
7.3 Identification and isolation efficiency for e and μ . . . 20
7.4 Identification efficiency for τ . . . 21
7.5 Energy scale of τ_h . . . 22
7.6 b-tag efficiency and mistag rate . . . 22
7.7 E_T^{miss} resolution and response . . . 23
8 Background estimation . . . 23
8.1 "Fake" background . . . 24
8.2 Charge "flip" background . . . 26
9 Systematic uncertainties . . . 26
10 Signal extraction . . . 29
11 Discriminating observables for signal extraction . . . 30
11.1 Multivariate analysis for categories with hadronic taus . . . 31
11.2 Multivariate analysis for categories without taus . . . 33
12 Results . . . 33
A Differences between 2016 and 2017 analyses . . . 45
B Combination with 2016 analysis . . . 47
C Reconstruction of hadronic top quark decays . . . 48
D Hj Tagger . . . 51
E Matrix Element Method . . . 53
F Measurement of lepton misidentification rates . . . 55
G Measurement of τ_h misidentification rate . . . 61

H Measurement of the lepton charge misidentification rate . . . 66
H.1 Electrons . . . 66
H.2 Muons . . . 68
I Background control regions . . . 70
I.1 $t\bar{t}W$ . . . 70
I.2 $t\bar{t}Z$ . . . 70
I.3 WZ+jets . . . 71
I.4 "Fake" background control plots . . . 71
I.5 $1\ell + 2ss\tau_h$ . . . 71
J Fit diagnosis on categories with taus . . . 74
K Alternative on categories with taus . . . 74
L Yields prefit . . . 77
M Pre-fit Distributions . . . 83
N Post-fit, fixed SM Higgs Distributions . . . 85
O Post-fit distributions for multilepton subcategories . . . 87

1 Introduction

The discovery of the Higgs (H) boson by the ATLAS and the CMS experiments [1, 2] represents a major step towards our understanding of the mechanism for electroweak symmetry breaking (EWSB). It also marks the beginning of a new era in particle physics that uses the H boson as a tool for discovering new physics through precision measurements of the H boson properties. The Standard Model (SM) makes precise predictions for all properties of the H boson, given its mass. The latter has been measured to be 125.09 ± 0.24 GeV by a combined analysis of the data recorded by ATLAS and CMS [3]. So far, all measured properties of the discovered resonance are consistent with expectations for the SM H boson within the uncertainties of these measurements, corroborating the the mechanism for EWSB of the SM.

The deviations of the H boson properties from the SM predictions, caused by the contributions of new physics, are expected to be small in many relevant models, however, in the order of a few percent [4]. This provides a strong motivation to improve upon the precision of the present measurements.

The measurement of the Yukawa coupling of the H boson to the top quark (y_t) is of high phenomenological interest for several reasons. In the SM, the coupling of the H boson to fermions is proportional to the fermion masses. The extraordinarily large value of the top quark mass, compared to the masses of all other known fermions, may indicate that the top quark has a still-unknown special role in the EWSB mechanism. The measurement of the cross section for the associated production of a H boson with a top quark pair ($t\bar{t}H$) provides the best sensitivity to determine y_t in a direct measurement.

A typical Feynman diagram for $t\bar{t}H$ production with subsequent decay of the H boson into a pair of τ leptons is shown in Fig. 1. The symbol τ_h denotes the system of charged and neutral particles that is produced in hadronic decays of these τ leptons.

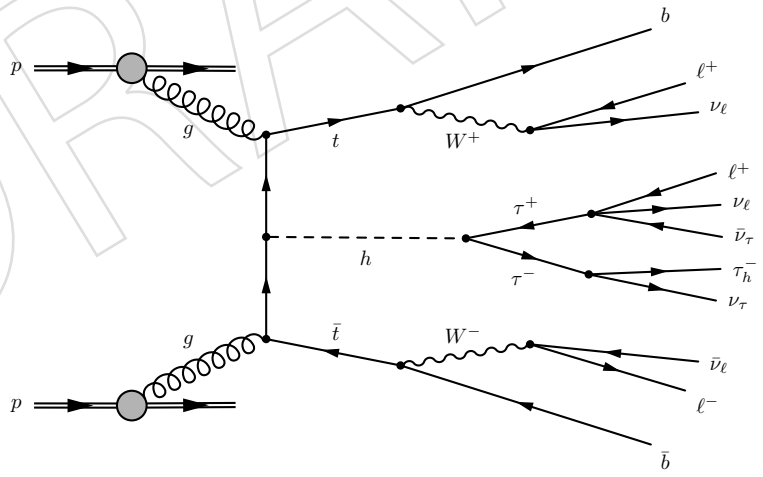


Figure 1: A typical Feynman diagram that depicts a process with a $t\bar{t}H$ intermediate state and a subsequent H decay to a pair of τ leptons.

This analysis note (AN) reports on the measurement of the $t\bar{t}H$ production cross section in events containing leptons (electrons and muons) and hadronic tau decays (τ_h) in the final state, using data recorded in proton–proton (pp) collisions at $\sqrt{s} = 13$ TeV center-of-mass energy in 2017, amounting to an integrated luminosity of 41.5 fb^{-1}. The analysis covers the decay channels H $\rightarrow$ WW, ZZ (excluding ZZ $\rightarrow 4\ell$), and $\tau\tau$. The sensitivity of the measurement

is enhanced by performing the analysis in event categories based on the multiplicity of leptons and τ_h and by using multivariate-analysis (MVA) techniques: the matrix element method (MEM) [5, 6] and machine learning techniques, in particular machine learning techniques based on boosted decision trees (BDT) [7, 8].

The note is organized as follows: in Section 2 we describe the datasets and Monte Carlo (MC) samples that were used. The particle identification and the reconstruction of global event quantities are described in Sections 3 and 4, respectively. Section 5 details the selection of events in the different final states considered in this study. Event categories, which are utilized in some of the final states to enhance the sensitivity of the analysis, are described in Section 6. Corrections that are applied to MC simulated events in order to improve the modeling of the data are detailed in Section 7. The estimation of background contributions is described in Section 8, followed by a discussion of systematic uncertainties in Section 9. The reconstruction of observables, based on MEM and BDT techniques, that are used for the signal extraction is described in Section 11. The statistical procedure used to extract the $\mathrm{t\bar{t}H}$ signal from the data is detailed in Section 10. The results are presented in Section 12. A summary, given in Section **??**, concludes the note.

The results of the analysis are combined with the results of a similar analysis based on data recorded in 2016 [9]. The combination is performed on datacard level. Details of the combination are presented in Section B of the Appendix. The differences between the 2016 and 2017 analyses are summarized in Section A of the Appendix.

2 Data samples and Monte Carlo simulation

The analyzed data has been collected in pp collisions at $\sqrt{s} = 13$ TeV center-of-mass energy and 25 ns bunch crossing period. The events have been recorded using a combination of single, double, and triple lepton triggers and by triggers based on the presence of a lepton and a τ_h. Only data-taking periods where all detector systems were fully operational are included in the analysis. The integrated luminosity of the analyzed dataset corresponds to 41.5 fb^{-1}. Approximately 30 inelastic pp interactions (pileup) occured per bunch crossing on average. The datasets analyzed by different channels are given in Table 1. The data has been reconstructed using `CMSSW_9_4_x` and is analyzed using the same release.

Samples of $\mathrm{t\bar{t}H}$ signal and of background events, produced by MC simulation, are used for the purpose of estimating signal and background yields in the analysis and to train machine learning algorithms. The following processes are simulated: Z+jets, W+jets, single top, diboson (WW, WZ, and ZZ) and triboson (WWW, WWZ, WZZ and ZZZ) production, the production of top quark pairs, and a few selected "exotic" processes. The exotic processes, such as $\mathrm{tt\bar{t}\bar{t}}$ and the production of same-sign WW boson pairs, typically have very small cross sections, but may nevertheless yield non-negligible background contributions in some event categories. Separate samples are produced to simulate the production of top quark pairs and of single top quarks in association with jets, photons, W and Z bosons. The $\mathrm{W}\gamma$+jets and $\mathrm{t\bar{t}}$+jets samples are produced using the leading-order (LO) matrix elements implemented in the MADGRAPH5_AMCATNLO 2.2.2 program [11]. Samples for other background processes and for the $\mathrm{t\bar{t}H}$ signal are generated using next-to-leading-order (NLO) matrix elements implemented in the programs MADGRAPH5_AMCATNLO and POWHEG $v2$ [12–14]. The signal events are generated for a H boson mass of $m_H = 125$ GeV. The samples produced by MADGRAPH and POWHEG, are generated with the NNPDF3.0 set of parton distribution functions, while the samples produced by PYTHIA use the NNPDF2.3LO set [15–17]. Parton shower and hadronization processes are modelled using the generator PYTHIA with the tune *CUETP8M1* [18],

Dataset name	Run-range	Luminosity (fb^{-1})
/SingleElectron/Run2017B-17Nov2017-v1/MINIAOD[1]	297047–299329	4.79
/SingleElectron/Run2017C-17Nov2017-v1/MINIAOD[1]	299368–302029	9.63
/SingleElectron/Run2017D-17Nov2017-v1/MINIAOD[1]	302030–302663	4.25
/SingleElectron/Run2017E-17Nov2017-v1/MINIAOD[1]	303818–304797	9.31
/SingleElectron/Run2017F-17Nov2017-v1/MINIAOD[1]	305040–306460	13.54
/SingleMuon/Run2017B-17Nov2017-v1/MINIAOD[1]	297047–299329	4.79
/SingleMuon/Run2017C-17Nov2017-v1/MINIAOD[1]	299368–302029	9.63
/SingleMuon/Run2017D-17Nov2017-v1/MINIAOD[1]	302031–302663	4.25
/SingleMuon/Run2017E-17Nov2017-v1/MINIAOD[1]	303824–304797	9.31
/SingleMuon/Run2017F-17Nov2017-v1/MINIAOD[1]	305040–306462	13.54
/DoubleEG/Run2017B-17Nov2017-v1/MINIAOD[1]	297047–299329	4.79
/DoubleEG/Run2017C-17Nov2017-v1/MINIAOD[1]	299368–302029	9.63
/DoubleEG/Run2017D-17Nov2017-v1/MINIAOD[1]	302030–302663	4.25
/DoubleEG/Run2017E-17Nov2017-v1/MINIAOD[1]	303818–304797	9.31
/DoubleEG/Run2017F-17Nov2017-v1/MINIAOD[1]	305040–306460	13.54
/DoubleMuon/Run2017B-17Nov2017-v1/MINIAOD[1]	297047–299329	4.79
/DoubleMuon/Run2017C-17Nov2017-v1/MINIAOD[1]	299368–302029	9.63
/DoubleMuon/Run2017D-17Nov2017-v1/MINIAOD[1]	302031–302663	4.25
/DoubleMuon/Run2017E-17Nov2017-v1/MINIAOD[1]	303824–304797	9.31
/DoubleMuon/Run2017F-17Nov2017-v1/MINIAOD[1]	305040–306462	13.54
/MuonEG/Run2017B-17Nov2017-v1/MINIAOD[1]	297047–299329	4.79
/MuonEG/Run2017C-17Nov2017-v1/MINIAOD[1]	299368–302029	9.63
/MuonEG/Run2017D-17Nov2017-v1/MINIAOD[1]	302031–302663	4.25
/MuonEG/Run2017E-17Nov2017-v1/MINIAOD[1]	303824–304797	9.31
/MuonEG/Run2017F-17Nov2017-v1/MINIAOD[1]	305040–306460	13.54

[1] Cert_294927-306462_13TeV_EOY2017ReReco_Collisions17_JSON_v1.txt

Table 1: List of datasets analyzed by different channels. The JSON file [10] used to apply a good-run selection is indicated by superscripts.

which is based on the *Monash* tune [19]. The decays of τ leptons, including polarization effects, are modelled by PYTHIA. The Z+jets, W+jets, and $t\bar{t}$+jets events are normalized to cross sections computed at next-to-next-to-leading-order (NNLO) accuracy [20, 21]. The cross sections for single top quark [22–24] and diboson [25] production are computed at NLO accuracy.

The complete list of signal and background samples is given in Table 2.

Minimum bias events generated with PYTHIA are overlaid on all simulated events, according to the luminosity profile of the analyzed data and for a pp inelastic cross section of 69.2 mb.

All generated events are passed through a detailed simulation of the CMS apparatus, based on GEANT4 [37], and are reconstructed using the same version of the CMS event reconstruction software as used for data.

3 Particle reconstruction and identification

The information provided by all CMS subdetectors is employed by a particle-flow (PF) algorithm [38–42] to identify and reconstruct individual particles in the event, namely muons, electrons, photons, charged and neutral hadrons. These particles are then used to reconstruct jets, τ_h candidates and the missing transverse momentum vector, as well as to quantify the isolation of leptons.

Some of the criteria used for particle identification, in particular the isolation of electrons and

Table 2: List of MC samples used to model the $t\bar{t}H$ signal and different background processes. In the first section of the table are listed the samples of the processes for which the simulation is used to extract the final yields and shapes, in the second section the samples of the processes estimated from data. The MC simulation is used to design the data driven methods and derive the associated systematics.

Process	Sample name	Cross section [pb]
$t\bar{t}H$	/ttHJetToNonbb_M125_TuneCP5_13TeV_amcatnloFXFX_madspin_pythia8/[1]	2.12×10^{-1} [26, 27]
$t\bar{t}W$	/TTWJetsToLNu_TuneCP5_13TeV-amcatnloFXFX-madspin-pythia8/[3]	2.04×10^{-1} [28]
	/TTWJetsToLNu_TuneCP5_PSweights_13TeV-amcatnloFXFX-madspin-pythia8/[1]	2.04×10^{-1} [28]
$t\bar{t}WW$	/TTWW_TuneCP5_13TeV-madgraph-pythia8/[4]	7.83×10^{-3} [29]
$t\bar{t}Z$	/TTZToLL_M-1to10_TuneCP5_13TeV-amcatnlo-pythia8/[1]	7.83×10^{-3} [30]
	/TTZToLLNuNu_M-10_TuneCP5_13TeV-amcatnlo-pythia8/[1]	2.53×10^{-1} [28]
$t\bar{t} + \gamma$+ jets	/TTGJets_TuneCP5_13TeV-amcatnloFXFX-madspin-pythia8/[1]	3.70 [28]
Single top quark $+\gamma$+ jets	/TGJets_TuneCUETP8M1_13TeV_amcatnlo_madspin_pythia8/[1]	2.97 [28]
WW	/WWTo2L2Nu_NNPDF31_TuneCP5_13TeV-powheg-pythia8/[1]	1.22×10^{1} [28]
	/WWTo2L2Nu_DoubleScattering_13TeV-herwigpp[5]	1.74×10^{-1} [28, 31]
WZ	/WZTo3LNu_TuneCP5_13TeV-amcatnloFXFX-pythia8/[3]	4.43 [28]
ZZ	/ZZTo4L_13TeV_powheg_pythia8/[3]	1.26 [28]
WWW	/WWW_4F_TuneCP5_13TeV-amcatnlo-pythia8/[5]	2.09×10^{-1} [28]
WWZ	/WWZ_4F_TuneCP5_13TeV-amcatnlo-pythia8/[5]	1.65×10^{-1} [28]
WZZ	/WZZ_TuneCP5_13TeV-amcatnlo-pythia8/[5]	5.57×10^{-2} [28]
ZZZ	/ZZZ_TuneCUETP8M1_13TeV-amcatnlo-pythia8/[5]	1.40×10^{-2} [28]
Single top quark +Z	/tZq_ll_4f_ckm_NLO_TuneCP5_PSweights_13TeV-amcatnlo-pythia8/[1]	7.58×10^{-2} [28]
$tt\bar{t}\bar{t}$	/TTTT_TuneCP5_13TeV-amcatnlo-pythia8/[1]	8.21×10^{-3} [32]
$t\bar{t}$+jets	/TTTo2L2Nu_TuneCP5_13TeV-powheg-pythia8/[1]	8.83×10^{1} [33]
	/TTTo2L2Nu_TuneCP5_PSweights_13TeV-powheg-pythia8/[1]	8.83×10^{1} [33]
	/TTToSemiLeptonic_TuneCP5_13TeV-powheg-pythia8/[1]	3.65×10^{2} [34]
	/TTToSemiLeptonic_TuneCP5_PSweights_13TeV-powheg-pythia8/[1]	3.65×10^{2} [34]
	/TTToHadronic_TuneCP5_13TeV-powheg-pythia8/[1]	3.78×10^{2} [35]
	/TTToHadronic_TuneCP5_PSweights_13TeV-powheg-pythia8/[1]	3.78×10^{2} [35]
$Z/\gamma^* \to \ell\ell$	/DYJetsToLL_M-10to50_TuneCP5_13TeV-madgraphMLM-pythia8/[1]	1.86×10^{4} [28]
	/DYJetsToLL_M-50_TuneCP5_13TeV-amcatnloFXFX-pythia8/[1,2]	5.77×10^{3} [28]
W+jets	/WJetsToLNu_TuneCP5_13TeV-madgraphMLM-pythia8/[1]	6.15×10^{4} [28]
Single top quark	/ST_s-channel_4f_leptonDecays_TuneCP5_13TeV-amcatnlo-pythia8/[1]	3.36 [28, 31, 36]
	/ST_t-channel_top_4f_inclusiveDecays_TuneCP5_13TeV-powhegV2-madspin-pythia8/[1]	1.36×10^{2} [36]
	/ST_t-channel_antitop_4f_inclusiveDecays_TuneCP5_13TeV-powhegV2-madspin-pythia8/[1]	8.10×10^{1} [36]
	/ST_tW_top_5f_inclusiveDecays_TuneCP5_13TeV-powheg-pythia8/[1]	3.56×10^{1} [28, 36]
	/ST_tW_antitop_5f_inclusiveDecays_TuneCP5_13TeV-powheg-pythia8/[1]	3.56×10^{1} [28, 36]

[1] RunIIFall17MiniAOD-94X_mc2017_realistic_v10-v1/MINIAODSIM

[2] RunIIFall17MiniAOD-94X_mc2017_realistic_v10_ext1-v1/MINIAODSIM

[3] RunIIFall17MiniAOD-94X_mc2017_realistic_v10-v2/MINIAODSIM

[4] RunIIFall17MiniAOD-94X_mc2017_realistic_v10_ext1-v2/MINIAODSIM

[5] RunIIFall17MiniAOD-94X_mc2017_realistic_v11-v1/MINIAODSIM

muons, depend on the choice of a primary collision vertex (PV). We choose this vertex to be the one that has the highest $\sum p_T^2$ of associated tracks.

3.1 Electrons and muons

The identification of electrons and muons is performed in two stages: Basic electron (muon) identification criteria, developed by the EGamma (Muon) POG, and loose isolation criteria are applied in the first stage to separate genuine leptons from jet backgrounds. In the second stage, the leptons originating from decays of W and Z bosons and from τ lepton decays are separated from leptons produced in the decays of charm (c) and bottom (b) quarks. We will refer to the former as "prompt" or signal leptons and to the latter as "non-prompt" or background leptons.

3.1.1 Basic electron identification

The identification of electrons is performed by a multivariate algorithm [43, 44], based on a BDT [7], which has been trained to separate electrons from jets. The training, performed by the EGamma POG, is done in two bins of p_T and three bins in η, using simulated samples of electrons and jets. Separate discriminants have been trained on electrons passing the requirements of the single electron trigger and for an inclusive sample of electron candidates. We use the discriminants trained on the inclusive electron selection. Loose and tight working-points (WP) are defined by corresponding cuts on the BDT output. The cuts, applied in different p_T and η bins, are given in Ref. [44].

In order to remove electron candidates that are due to photon conversions we require that the electron track is associated to a hit in each layer of the pixel detector that is crossed by the track, except for at most one layer. We further reject electron candidates in case there exist a track of opposite charge near the electron track that, if paired with the electron track, can be fitted to a common vertex within the volume of the tracking detector.

3.1.2 Basic muon identification

The identification of muons is based on linking track segments reconstructed in the silicon tracking detector and in the muon system [45]. The matching between track segments is done outside-in, starting from a track in the muon system, and inside-out, starting from a track reconstructed in the inner detector. In case a link can be established, the track parameters are refitted using the combination of hits in the inner and outer detectors and the track is referred to as global muon track. Quality cuts are applied on the multiplicity of hits, on the number of matched segments and on the quality of the global muon track fit, quantified by χ^2. The muon candidates used in the analysis are required to pass the medium PF muon identification criteria [46].

3.1.3 Electron and muon isolation

Electrons and muons in signal events are expected to be isolated, while leptons from c and b quark decays, as well as from in-flight decays of pions and kaons, are often reconstructed within jets. Isolated leptons are distinguished from leptons in jets by means of the sum of scalar p_T values of charged particles, neutral hadrons, and photons, that are reconstructed within a narrow cone centered on the lepton direction. The size R of the cone shrinks inversely proportional with the p_T of the lepton in order to increase the efficiency for leptons reconstructed in events with "boosted" topologies and/or high hadronic activity to pass the isolation criteria. The narrow cone size, referred to as "mini isolation", has the further advantage that

it reduces the effect of pileup (PU). Efficiency loss due to PU is further reduced by considering only charged particles originating from the lepton production vertex in the isolation sum. Residual contributions of PU to the neutral component of the isolation of the lepton is taken into account by means of so-called effective area corrections:

$$I_\ell = \sum_{\text{charged}} p_\mathrm{T} + \max\left(0, \sum_{\text{neutrals}} p_\mathrm{T} - \rho\,\mathcal{A}\,\left(\frac{R}{0.3}\right)^2\right), \tag{1}$$

where ρ represents the energy density of neutral particles reconstructed within the geometric acceptance of the tracking detectors (`fixedGridRhoFastjetAll`), computed as described in Refs. [47, 48]. The size of the cone is given by:

$$R = \begin{cases} 0.05 & \text{if } p_\mathrm{T} > 200 \text{ GeV} \\ 10 \text{ GeV}/p_\mathrm{T} & \text{if } 50 < p_\mathrm{T} < 200 \text{ GeV} \\ 0.20 & \text{if } p_\mathrm{T} < 50 \text{ GeV} \end{cases}. \tag{2}$$

The effective area $\mathcal{A}$ is obtained from the simulation, by studying the correlation between I_ℓ and ρ, and is determined in bins of η, separately for electrons and muons. Numerical values are given in Table 3.

Electrons	
Pseudorapidity range	$\mathcal{A}$
$0.0 < \lvert\eta\rvert < 1.0$	0.1566
$1.0 < \lvert\eta\rvert < 1.479$	0.1626
$1.479 < \lvert\eta\rvert < 2.0$	0.1073
$2.0 < \lvert\eta\rvert < 2.2$	0.0854
$2.2 < \lvert\eta\rvert < 2.3$	0.1051
$2.3 < \lvert\eta\rvert < 2.4$	0.1204
$2.4 < \lvert\eta\rvert < 2.5$	0.1524

Muons	
Pseudorapidity range	$\mathcal{A}$
$0.0 < \lvert\eta\rvert < 0.8$	0.0566
$0.8 < \lvert\eta\rvert < 1.3$	0.0562
$1.3 < \lvert\eta\rvert < 2.0$	0.0363
$2.0 < \lvert\eta\rvert < 2.2$	0.0119
$2.2 < \lvert\eta\rvert < 2.5$	0.0064

Table 3: Effective areas $\mathcal{A}$ for electrons (top) and muons (bottom).

3.1.4 Separation of prompt from non-prompt electrons and muons

The final separation of prompt leptons from non-prompt and fake leptons is performed by a BDT-based algorithm [49, 50]. The algorithm has been retrained on MC samples with 2017 detector conditions, separately for electrons and muons. The following observables are used as inputs:

- `miniRelIsoCharged`, the isolation of the lepton with respect to charged particles, $I_\ell^{\text{charged}} = \sum_{\text{charged}} p_\mathrm{T}$.
- `miniRelIsoNeutral`, the isolation of the lepton with respect to neutral particles, corrected for PU effects, $I_\ell^{\text{neutrals}} = \max\left(0, \sum_{\text{neutrals}} p_\mathrm{T} - \rho\,\mathcal{A}\,\left(\frac{R}{0.3}\right)^2\right)$.

- `jetPtRatio`, the ratio of the transverse momentum of the lepton to the transverse momentum of the nearest jet, $p_{\mathrm{T}}^{\ell}/p_{\mathrm{T}}^{\mathrm{j}}$.
- `jetBTagCSV`, the discriminant value of the CSVv2 b-tagging algorithm of the jet (see Section 3.3).
- `jetNDauChargedMVASel`: the number N_{charged} of charged particles within the jet.
- `jetPtRelv2`, the component of the lepton momentum in direction transverse to the jet, $p_{\mathrm{T}}^{\mathrm{rel}} = p_{\ell} \sin\theta$. where θ denotes the angle between the lepton and jet momentum vectors.
- `dxy` and `dz`, the transverse and longitudinal impact parameters of the lepton track with respect to the PV.
- `sip3d`, the signed impact parameter, in three dimensions, of the lepton track with respect to the PV, divided by its uncertainty, which corresponds to its significance d/σ_d.
- `mvaIdFall17noIso`, the output of the BDT that separates electrons from jets, trained by the EGamma POG.
- `segmentCompatibility`, the compatibility of track segments in the muon system with the pattern expected for a minimum ionizing particle.

The observable `mvaIdFall17noIso` (`segmentCompatibility`) is used only for electrons (muons). Lower jet p_{T} threshold of 15 GeV is used for the matching of jets to leptons. In case no jet of $p_{\mathrm{T}} > 15$ GeV is within a distance $\Delta R < 0.4$ to the lepton, the value of the observable `jetPtRatio` is set to $p_{\mathrm{T}}^{\ell}/(p_{\mathrm{T}}^{\ell} + I_{\ell})$ and the value of the observables `jetBTagCSV`, `jetNDauChargedMVASel`, and `jetPtRelv2` is set to zero. The p_{T} of the jet is computed by applying the "lepton-aware" jet energy corrections (JEC) [51], which apply the JEC to the difference of uncalibrated jet minus lepton and subsequently add the lepton momentum to the calibrated jet. The inputs are complemented by the p_{T} and η of the lepton. We refer to the output of the BDT trained on electrons (muons) as prompt-e (prompt-μ) MVA.

The BDTs are trained on simulated samples of prompt leptons in $\mathrm{t\bar{t}H}$ events (signal) and non-prompt leptons in $\mathrm{t\bar{t}}$+jets events (background). The leptons used for the training are required to pass loose preselection criteria (referred to as "loose" lepton selection), given by Tables 4 and 5. Further "tight" and "fakeable" lepton selection criteria are used for the purpose of selecting events in the signal region and for estimating the fake lepton background from control regions in data, respectively. In order to reduce potential biases of the background estimation procedure, the p_{T} of leptons that pass the fakeable, but fail the tight lepton selection criteria is set to 0.90 times the p_{T} of the nearest jet in case the distance between lepton and nearest jet satisfies $\Delta R < 0.4$. In case $\Delta R > 0.4$ the p_{T} of fakeable leptons is set to $p_{\mathrm{T}}^{\ell}/(p_{\mathrm{T}}^{\ell} + I_{\ell})$. We refer to the p_{T} of fakeable leptons, computed in this way, as "cone-p_{T}". The cone-p_{T} in general exceeds the transverse momentum of the lepton that is determined by the electron and muon reconstruction algorithms. We refer to the latter as "reco-p_{T}" to distinguish it from the cone-p_{T} when there is danger of confusion. When there is no danger of confusion, we use the terms reco-p_{T} and p_{T} synonymously. Electrons passing the fakeable and tight lepton selection criteria are required to satisfy a set of conditions on the width of the electron cluster in η-direction ($\sigma_{i\eta i\eta}$), the ratio of energy in the HCAL to the energy in the ECAL that is associated to electron (H/E), and the difference between the reciprocal of the electron cluster energy and the reciprocal of its track momentum (1/E - 1/p). They mimic the electron identification criteria applied at trigger level.

The small lifetime of the τ lepton of $c\tau = 87$ μm [52] causes the distributions in d_{xy}, d_z and

d/σ_d to be more background-like for electrons and muons produced in τ decays compared to prompt leptons originating from decays of W and Z bosons, resulting in a few percent lower efficiency for the former.

Observable	Loose	Fakeable	Tight
p_T	> 7 GeV	> 10 GeV	> 10 GeV
$\|\eta\|$	< 2.5	< 2.5	< 2.5
$\|d_{xy}\|$	< 0.05 cm	< 0.05 cm	< 0.05 cm
$\|d_z\|$	< 0.1 cm	< 0.1 cm	< 0.1 cm
d/σ_d	< 8	< 8	< 8
I_e	$< 0.4 \times p_T$	$< 0.4 \times p_T$	$< 0.4 \times p_T$
EGamma POG MVA	$> \{-0.86 / -0.81 / -0.72\}$[1]	$> \{-0.86 / -0.81 / -0.72\}$[2]	$> \{-0.86 / -0.81 / -0.72\}$[2]
$\sigma_{i\eta i\eta}$	–	$< \{0.011 / 0.011 / 0.030\}$	$< \{0.011 / 0.011 / 0.030\}$
H/E	–	< 0.10	< 0.10
1/E - 1/p	–	> -0.04	> -0.04
Conversion rejection	–	✓	✓
Missing hits	≤ 1	$= 0$	$= 0$
p_T^e / p_T^j	–	> 0.6 † $(-)$	–
CSVv2 of nearby jet	–	< 0.07 † (< 0.4941)	< 0.4941
Prompt-e MVA	–	–	> 0.90

[1] $> \{-0.13 / -0.32 / -0.08\}$ if $p_T < 10$ GeV
[2] > 0.50 if prompt-e MVA < 0.90

Table 4: Loose, fakeable, and tight selection criteria for electrons. The requirement on the output of the BDT trained by the EGamma POG (see Section 3.1.1) and on the observables $\sigma_{i\eta i\eta}$, H/E, and 1/E - 1/p are varied as function of η of the electron candidate. The numbers are separated by slashes and refer to $|\eta| < 0.8$, $0.8 < |\eta| < 1.479$, and $|\eta| > 1.479$, respectively. The conditions on p_T^e / p_T^j and on the CSV discriminant of the jet nearest to the electron are tightened (relaxed) for fakeable electrons that fail (pass) the requirement prompt-e MVA > 0.90, in order to reduce the systematic uncertainty on the fake lepton background estimate on the jet flavour composition (see Section 7.4.1 in Ref. [53]). The tightened (relaxed) conditions are indicated by the symbol † (put in parentheses). A hyphen $(-)$ indicates selection criteria that are not applied.

3.2 Hadronic τ decays

Hadronic τ decays are reconstructed by the "hadrons plus strips" (HPS) algorithm [54, 55]. The algorithm allows to reconstruct individual hadronic decay modes of the τ: $h^\pm$, $h^\pm + 1\pi^0$, $h^\pm + 2\pi^0$ and $h^\pm h^\mp h^\pm$, where $h^\pm$ denotes either a charged pion or kaon. The decay modes of τ^+ are the charge conjugate of the τ^- decay modes. We refer to the set of decay modes $h^\pm$, $h^\pm + 1\pi^0$ and $h^\pm + 2\pi^0$ as hadronic "one-prong" τ decays, and to the decay mode $h^\pm h^\mp h^\pm$ as hadronic "three-prong" τ decays. Hadronic τ candidates are built by combining the charged hadrons reconstructed by the PF algorithm with neutral pions. The latter are reconstructed by clustering the photons reconstructed by the PF algorithm within rectangular strips, that are narrow in η-, but wide in ϕ-direction, which accounts for the broadening of energy deposits in the ECAL in case one of the photons produced in $\pi^0 \to \gamma\gamma$ decays converts within the tracking detector. The size of the strip is adjusted as function of p_T, taking into consideration that the bending of charged particles in the magnetic field increases inversely proportional to p_T [56]. Photon conversions within the silicon tracking detector are accounted for by considering not

Observable	Loose	Fakeable	Tight
p_T	> 5 GeV	> 10 GeV	> 10 GeV
$\|\eta\|$	< 2.4	< 2.4	< 2.4
$\|d_{xy}\|$	< 0.05 cm	< 0.05 cm	< 0.05 cm
$\|d_z\|$	< 0.1 cm	< 0.1 cm	< 0.1 cm
d/σ_d	< 8	< 8	< 8
I_μ	$< 0.4 \times p_T$	$< 0.4 \times p_T$	$< 0.4 \times p_T$
Loose PF muon	✓	✓	✓
Medium PF muon	–	–	✓
Segment compatibility	–	> 0.3 † $(-)$	–
p_T^μ / p_T^j	–	> 0.6 † $(-)$	–
CSVv2 of nearby jet	–	< 0.07 † (< 0.4941)	< 0.4941
Prompt-μ MVA	–	–	> 0.90

Table 5: Loose, fakeable, and tight selection criteria for muons. The conditions on p_T^μ / p_T^j and on the CSV discriminant of the jet nearest to the muon are tightened (relaxed) for fakeable muons that fail (pass) the requirement prompt-μ MVA > 0.90, in order to reduce the systematic uncertainty on the fake lepton background estimate on the jet flavour composition (see Section 7.4.1 in Ref. [53]). The tightened (relaxed) conditions are indicated by the symbol † (put in brackets). A hyphen $(-)$ indicates selection criteria that are not applied.

only the photons, but also the electrons reconstructed by the PF algorithm for the reconstruction of neutral pions.

A BDT is used to separate hadronic τ decays from the quark and gluon jet background [57]. The BDT uses as input the sums of scalar p_T values of charged particles and of photons, computed within a cone of size $R = 0.3$ that is centered on the τ_h direction, the reconstructed τ_h decay mode and observables that provide sensitivity to the lifetime of the τ. The size of the cone has been optimized to identify τ_h in events with high hadronic activity. The transverse impact parameter of the "leading" (highest p_T) track of the τ_h candidate with respect to the PV is used for τ_h candidates reconstructed in any decay mode. In case of τ_h candidates reconstructed in the decay mode $h^\pm h^\mp h^\pm$, a fit of the three tracks to a common secondary vertex is attempted and the distance to the PV is used as additional input variable to the BDT. The BDT has been trained on simulated samples of genuine hadronic τ decays and jets. We use the training `2017v2` [58]. Various WPs, corresponding to different τ_h identification efficiencies and misidentification rates for jets, are defined by varying the selections on the BDT output. The thresholds are adjusted as function of the p_T of the τ_h candidate, such that the τ_h identification efficiency for each WP is constant as function of p_T.

Two levels of τ_h identification criteria are utilized in the analysis. The selection events in the signal region, described in Section 5, is based on the collection of "tight" τ_h, while the collection of "fakeable" τ_h is used for the purpose of obtaining a data-driven estimate of the fake τ_h background. The selection criteria applied to tight and fakeable τ_h are given in Table 6. The τ_h are required to be reconstructed in one of the "old" decay modes $h^\pm$, $h^\pm + 1\pi^0$, $h^\pm + 2\pi^0$, or $h^\pm h^\mp h^\pm$ [54], to satisfy $p_T > 20$ GeV and $|\eta| < 2.3$, and not to overlap, within $\Delta R < 0.3$, with any electron or muon passing the loose lepton selection criteria. The τ_h identification criteria applied to separate hadronic τ decays from the jet background are channel specific. In the $2\ell ss + 1\tau_h$ and $3\ell + 1\tau_h$ channels the tight τ_h are required to pass the loose WP of the τ_h identification discriminant, whereas the hadronic τ decays selected in the signal region of the $1\ell + 2\tau_h$ and $2\ell + 2\tau_h$ channels are required to pass the medium WP (see Section 5). The dis-

Observable	Fakeable	Tight
p_{T}	> 20 GeV	> 20 GeV
$\lvert\eta\rvert$	< 2.3	< 2.3
$\lvert d_z\rvert$	< 0.2 cm	< 0.2 cm
Decay mode finding	old	old
Tau POG MVA ($R = 0.3$)	very-loose WP	loose / medium WP
Anti-e discriminant	–	–
Anti-μ discriminant	–	–

Table 6: Fakeable and tight selection criteria for hadronic τ decays. A hyphen ($-$) indicates selection criteria that are not applied. The loose WP is used in the $2\ell ss + 1\tau_{\mathrm{h}}$ and $3\ell + 1\tau_{\mathrm{h}}$ channels, while the medium WP is used in the $1\ell + 2\tau_{\mathrm{h}}$ and $2\ell + 2\tau_{\mathrm{h}}$ channels.

criminants provided by the Tau POG to separate hadronic τ decays from electrons and muons are not applied, as they are found not to improve the sensitivity of the analysis. Instead, all τ_{h} candidates considered in the analysis are required not to overlap, within $\Delta R = 0.3$, with electrons or muons passing the loose preselection criteria given in Tables 4 and 5.

3.3 Jets

Jets are reconstructed using the anti-k_{t} algorithm [59] with a distance parameter $R = 0.4$. Charged particles not originating from the PV are excluded from the jet clustering. Fake jets, mainly arising from calorimeter noise, are rejected by requiring reconstructed jets to pass a set of loose jet identification criteria [60]. The energy of reconstructed jets is calibrated as function jet p_{T} and η [61]. Corrections based on jet area and energy density [47, 48] are applied in order to compensate for PU effects. The global tags used to apply JEC to data and simulated events are given in Table 7.

Jets considered in the analysis are required to satisfy the conditions $p_{\mathrm{T}} > 25$ GeV and $\lvert\eta\rvert < 2.4$, and are required not to overlap, within $\Delta R < 0.4$, with electrons, muons, and τ_{h} passing the selection criteria for fakeable objects.

Table 7: Global tags used to apply jet energy corrections to data and simulated events.

Event type	global tag
Simulation	94X_mc2017_realistic_v13
Data	94X_dataRun2_v6

Jets originating from the hadronization of b quarks are identified by the "combined secondary vertex" (CSVv2 and DeepCSV) algorithm [62, 63], which exploits observables related to the long lifetime of b hadrons and to the higher particle multiplicity and mass of b-jets compared to light quark and gluon jets. While the discriminant computed by the CSVv2 algorithm is used for the purpose of separating prompt from non-prompt leptons (*cf.* Section 3.1.4), the b-jets selected on analysis level are identified using the DeepCSV algorithm. Two levels of cuts on the discriminant d computed by the DeepCSV algorithm are used in the analysis, corresponding to the loose ($d > 0.1522$) and medium ($d > 0.4941$) WP defined by the BTV POG [64].

4 Reconstruction of event level quantities

The missing transverse momentum vector is computed as the negative sum of the transverse momentum vectors of all the Particle-Flow candidates reconstructed in the event. The difference in momentum between calibrated and uncalibrated jets is applied as correction ("Type-1 corrected PFMET") [65]. The magnitude of the missing transverse momentum vector is referred to as E_T^{miss}. To mitigate the influence of pile-up on the missing transverse energy, the H_T^{miss} variable is considered as well and is defined in the same way as E_T^{miss}, but considering only jets of $p_T > 25$ GeV and $|\eta| < 2.4$ as well as electrons, muons, and τ_h passing the fakeable object selection criteria when evaluating the sum of transverse momenta. While the resolution in H_T^{miss} is worse compared to the resolution in E_T^{miss}, the observable H_T^{miss} has the advantage to be less sensitive to pileup, as soft (low p_T) hadrons, which predominantly originate from pileup, do not enter its computation. A linear combination of E_T^{miss} and H_T^{miss} is used in the event selection, utilizing the fact that the two observables are less correlated for events in which the reconstructed missing transverse momentum arises from instrumental effects compared to events with genuine missing transverse momentum. The linear combination, referred to as $E_T^{miss}LD$, is defined as:

$$E_T^{miss}LD = 0.6 \times E_T^{miss} + 0.4 \times H_T^{miss}\,. \tag{3}$$

As $t\bar{t}H$ signal events typically contain neutrinos originating both from the H boson and from top quark decays, the reconstruction of the H boson mass is very challenging. Algorithms for reconstructing the H boson mass, such as the SVfit [66, 67] algorithm used in $H \to \tau\tau$ analyses, cannot be used, as they crucially depend on the assumption that all E_T^{miss} in the event originates from the decay of the H boson. In the $1\ell + 2\tau_h$ and $2\ell + 2\tau_h$ channels we instead use the visible mass, denoted by m_{vis}, of the two τ_h, taking advantage of the fact that the two τ_h are likely to originate from the H boson decay.

5 Event selection

The event selection aims to select $t\bar{t}H$ signal events in which the Higgs boson decays into a pair of either W bosons, Z bosons, or τ leptons, with subsequent decay of the W and Z bosons and of the τ leptons to electrons, muons, and τ_h. The events are analyzed in seven mutually exclusive categories, also referred to as "channels", based on the multiplicity of reconstructed leptons and τ_h: $1\ell + 2\tau_h$, $2\ell ss$, $2\ell ss + 1\tau_h$, $2\ell + 2\tau_h$, 3ℓ, $3\ell + 1\tau_h$, and 4ℓ. Regardless of the multiplicity of leptons and τ_h, $t\bar{t}H$ signal events are expected to always contain two b-jets. Events selected in any one of the seven channels are therefore required to contain at least two jets of $p_T > 25$ GeV and $|\eta| < 2.4$. We further demand that at least two jets pass the loose WP of the DeepCSV discriminant or at least one jet passes the medium WP. Events containing two leptons of mass $m_{\ell\ell} < 12$ GeV, which pass the loose lepton selection criteria, are rejected, as these events are not well modeled by the simulation. The triggers used to select events in all channels are detailed in Section 5.1. Following the recommendation of the JetMET POG [68], events selected in any of the channels are required to pass filter algorithms, which remove events that are subject to different types of spurious detector signals. These algorithms are referred to as "E_T^{miss} filters" and are described in Section 5.2. Further channel specific event selection criteria are described in Sections 5.3 to 5.9, and are summarized in Table 10 and 11.

5.1 Triggers

The HLT paths used to record events in the $1\ell + 2\tau_h$, $2\ell ss$, $2\ell ss + 1\tau_h$, $2\ell + 2\tau_h$, 3ℓ, $3\ell + 1\tau_h$, and 4ℓ channels are given in Table 8. Events in the $1\ell + 2\tau_h$ channel are recorded using a

Single lepton triggers	HLT_Ele32_WPTight_Gsf HLT_Ele35_WPTight_Gsf HLT_IsoMu24 HLT_IsoMu27
Lepton+τ_h cross-triggers	HLT_Ele24_eta2p1_WPTight_Gsf_LooseChargedIsoPFTau30_eta2p1_CrossL1 HLT_IsoMu20_eta2p1_LooseChargedIsoPFTau27_eta2p1_CrossL1
Double lepton triggers	HLT_Ele23_Ele12_CaloIdL_TrackIdL_IsoVL HLT_Ele23_Ele12_CaloIdL_TrackIdL_IsoVL_DZ HLT_Mu23_TrkIsoVVL_Ele12_CaloIdL_TrackIdL_IsoVL HLT_Mu23_TrkIsoVVL_Ele12_CaloIdL_TrackIdL_IsoVL_DZ HLT_Mu12_TrkIsoVVL_Ele23_CaloIdL_TrackIdL_IsoVL_DZ HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_DZ HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_DZ_Mass3p8
Triple lepton triggers	HLT_Ele16_Ele12_Ele8_CaloIdL_TrackIdL HLT_Mu8_DiEle12_CaloIdL_TrackIdL HLT_DiMu9_Ele9_CaloIdL_TrackIdL_DZ HLT_TripleMu_12_10_5

Table 8: Triggers used to record events selected in the $1\ell + 2\tau_h$, $2\ell ss$, $2\ell ss + 1\tau_h$, $2\ell + 2\tau_h$, 3ℓ, $3\ell + 1\tau_h$, and 4ℓ channels.

combination of single lepton triggers and triggers based on the presence of a lepton and a τ_h. The latter are referred to as lepton+τ_h "cross-triggers". A combination of single and double lepton triggers are used to record events in the $2\ell ss$, $2\ell ss + 1\tau_h$, and $2\ell + 2\tau_h$ channels, while events in the 3ℓ, $3\ell + 1\tau_h$, and 4ℓ channels are recorded using a combination of single, double, and triple lepton triggers. For some triggers, we use a mix of HLT paths with different p_T thresholds or a mix of paths with and without a d_z requirement applied to the leptons. In case of the dimuon triggers we further use a mix of HLT paths with and without a requirement $m_{\mu\mu} > 3.8$ GeV on the mass of the muon pair. The motivation for choosing such a mix is to use the trigger of higher efficiency whenever it is available and to resort to triggers of lower efficiency whenever the trigger of higher efficiency is disabled or prescaled. Events recorded in any data-taking period as well as simulated events are selected in case they pass any of the HLT paths that are included in the mix. The effect of triggers that are disabled or prescaled in some part of the analyzed data is accounted for by applying suitably chosen weights to simulated events. These weights are detailed in Section 7.2.

5.2 E_T^{miss} filters

Events selected in any of the channels $1\ell + 2\tau_h$, $2\ell ss$, $2\ell ss + 1\tau_h$, $2\ell + 2\tau_h$, 3ℓ, $3\ell + 1\tau_h$, and 4ℓ are required to pass the filter algorithms given in Table 9. The Flag_eeBadScFilter filter is applied to events selected in data only, and not to simulated events, as recommended by the JetMET POG [68].

5.3 $1\ell + 2\tau_h$ category

The $1\ell + 2\tau_h$ category predominantly selects $t\bar{t}H$ signal events in which one top quarks decays leptonically, the other hadronically, and the H boson decays into two τ_h. Events selected in this category are required to contain one electron or muon passing the tight object selection criteria and two τ_h, which both pass the medium WP of the τ_h identification discriminant. The lepton is required to be within the geometric acceptance of the lepton+τ_h cross-trigger, $|\eta| < 2.1$, and to have $p_T > 30$ GeV (> 25 GeV) if it is an electron (muon). The two τ_h are required to have

Filter name	Applied to data	Applied to simulation
Flag_goodVertices	✓	✓
Flag_globalTightHalo2016Filter	✓	✓
Flag_HBHENoiseFilter	✓	✓
Flag_HBHENoiseIsoFilter	✓	✓
Flag_EcalDeadCellTriggerPrimitiveFilter	✓	✓
Flag_BadPFMuonFilter	✓	✓
Flag_BadChargedCandidateFilter	✓	✓
Flag_eeBadScFilter	✓	–
Flag_ecalBadCalibFilter	✓	✓

Table 9: E_T^{miss} filters applied to events selected in data and to simulated events. A hyphen (–) indicates that the filter is not applied.

opposite charge, as expected for a τ_h pair produced in a H boson decay. The contribution of the dominant background, arising from $t\bar{t}$+jets events in which one or two jets are misidentified as τ_h, is reduced by demanding that the τ_h of higher p_T satisfies the condition $p_T > 30$ GeV and the event contains at least three jets of $p_T > 25$ GeV and $|\eta| < 2.4$. Events with more than one electron or muon passing the tight object selection criteria are vetoed, to avoid overlap with the $2\ell + 2\tau_h$ category.

5.4 2ℓss category

The $2\ell ss$ category targets $t\bar{t}H$ signal events in which the H boson decays into a pair of W bosons, one of the W bosons produced in the H boson decay and one of the top quarks decays to leptons, while the other W boson and top quark decays hadronically. Selected events are required to contain two leptons of the same charge and passing the tight object selection criteria. The lepton of higher (lower) p_T is required to have $p_T > 25$ GeV (> 15 GeV). Requiring both leptons to be of the same charge cuts half of the $t\bar{t}H$ signal events, but removes almost all of the large $t\bar{t}$+jets background. Residual $t\bar{t}$+jets background contributions are further suppressed by requiring that the charge of all electrons and muons in the event, which pass the fakeable object selection criteria, is well measured. Electrons are required to pass the two conditions `isGsfCtfScPixChargeConsistent` and `isGsfScPixChargeConsistent`, which test the consistency between the independent measurements of the electron charge obtained from the position of the ECAL cluster and from its track, while muons must satisfy the condition that the estimated uncertainty on the p_T of the muon track is below 0.2 times its p_T. The latter is equivalent to requiring that the sign of the curvature of the muon track, which determines its charge, is measured with a significance of 5σ. After applying these conditions, the charge misidentification rate is on the per-mille level for electrons and negligible for muons [69]. Background from $t\bar{t}Z$ production is suppressed by requiring that the event contains no pair of loose electrons with mass close to the mass of the Z boson, $|m_{ee} - m_Z| < 10$ GeV, where $m_Z = 91.2$ GeV [52]. In case all leptons passing the fakeable object selection in the event are electrons, the event is further required to satisfy the condition $E_T^{miss} LD > 30$ GeV. Finally, the events are required to contain at least four jets of $p_T > 25$ GeV and $|\eta| < 2.4$. Events containing more than two tight leptons or a τ_h passing the loose τ_h identification discriminant are vetoed, to avoid overlap with the 3ℓ and $2\ell ss + 1\tau_h$ categories.

5.5 2ℓss $+ 1\tau_h$ category

The event selection criteria applied in the $2\ell ss + 1\tau_h$ category are identical to those applied in the $2\ell ss$ category, except that events selected in the former are required to contain one τ_h passing the loose τ_h identification discriminant and the requirement on the multiplicity of jets

is relaxed from four to three. Also, if the second highest p_T lepton is a muon, the p_T requirement is relaxed to $p_T > 10$ GeV. The charge of the τ_h is required to be opposite to the charge of the leptons. Overlap with the $2\ell + 2\tau_h$ category is avoided by applying a modified τ_h veto, which demands that events selected in the $2\ell ss + 1\tau_h$ category must not contain two τ_h passing the medium WP of the τ_h identification discriminant.

5.6 3ℓ category

The 3ℓ category predominantly selects events in which the H boson decays into a pair of W bosons, and in which either one of the W bosons from the H boson decay or one of the top quarks decay hadronically, while the remaining W bosons and top quarks decay to leptons. Selected events are required to contain exactly three leptons passing the tight object selection criteria. The lepton of highest, second, and third highest p_T is required to have $p_T > 25$ GeV, > 15 GeV, and > 10 GeV, respectively. The charge sum of the leptons is required to be either $+1$ or -1. Background from $t\bar{t}Z$ production is removed by vetoing events containing a pair of loose leptons of the same flavor, opposite charge, and mass $|m_{\ell\ell} - m_Z| < 10$ GeV. Events containing fewer than four jets are required to satisfy the condition $E_T^{miss}LD > 45$ GeV in case the event contains a pair of loose leptons of the same flavor and opposite charge and $E_T^{miss}LD > 30$ GeV if the event contains no such lepton pair. In case the event contains four or more jets, no requirement on $E_T^{miss}LD$ is applied, as the contributions of background processes are small in this case anyway. Events containing τ_h passing the loose τ_h identification discriminant are vetoed, as are events containing two pairs of loose leptons of the same flavor and opposite charge, which satisfy the condition $m_{\ell\ell\ell\ell} < 140$ GeV. While the first condition avoids overlap with the $3\ell + 1\tau_h$ category, the second condition avoids overlap with the $t\bar{t}H$-tagged category of the $H \to ZZ^* \to 4\ell$ analysis [70].

5.7 $3\ell + 1\tau_h$ category

The event selection criteria applied in the $3\ell + 1\tau_h$ category are identical to those applied in the 3ℓ category, except that no τ_h veto is applied in the $3\ell + 1\tau_h$ category and the events selected in this category are instead required to contain one τ_h passing the loose τ_h identification discriminant. The acceptance for the $t\bar{t}H$ signal is increased by lowering the p_T thresholds for the leptons to $p_T > 20$ GeV, > 10 GeV, and > 10 GeV for the lepton of highest, second, and third highest p_T, respectively. The charge sum of leptons and τ_h is required to be zero.

5.8 4ℓ category

The 4ℓ category targets events in which the H boson decays into a pair of W bosons and in which all W bosons and top quarks decay leptonically. Events selected in this category are required to contain four leptons passing the tight object selection criteria and passing p_T thresholds of $p_T > 25$ GeV, > 15 GeV, > 15 GeV, and > 10 GeV, for the lepton of highest, second, third, and fourth highest p_T, respectively. Remaining event selection criteria are similar to those of the 3ℓ category. The charge sum of the leptons is required to be zero. Events containing a pair of loose leptons of the same flavor, opposite charge, and mass $|m_{\ell\ell} - m_Z| < 10$ GeV are vetoed. Events containing fewer than four jets are required to satisfy the condition $E_T^{miss}LD > 45$ GeV in case the event contains a pair of fakeable leptons of the same flavor and opposite charge and $E_T^{miss}LD > 30$ GeV if the event contains no such lepton pair. Events containing two pairs of loose leptons of the same flavor and opposite charge, which satisfy the condition $m_{\ell\ell\ell\ell} < 140$ GeV, are again vetoed.

5.9 $2\ell + 2\tau_h$ category

The $2\ell + 2\tau_h$ category aims at selection signal events in which both of the top quarks decay leptonically, and the H boson decays into two τ_h. Events selected in this category are required to contain two leptons that pass the tight selection criteria. The lepton of higher p_T is required to have $p_T > 25$ GeV, and the lower p_T lepton is requried to have $p_T > 10$ GeV ($p_T > 15$ GeV) if it is a muon (an electron). To suppress background from $t\bar{t}Z$ process, events containing a pair of loose leptons of the same flavor and opposite charge, and mass $|m_{\ell\ell} - m_Z| < 10$ GeV are vetoed. Two τ_h passing the medium WP are required, and the charge sum of the two leptons and two τ_h is required to be zero. Events containing fewer than four jets are required to satisfy the condition $E_T^{miss}LD > 45$ GeV in case the event contains a pair of fakeable leptons of the same flavor and opposite charge and $E_T^{miss}LD > 30$ GeV if the event contains no such lepton pair. In case the event contains four or more jets, no requirement on $E_T^{miss}LD$ is applied.

6 Event categorization

The event statistics in the $2\ell ss$ and 3ℓ categories is high enough to allow that events selected in these categories are analyzed in subcategories. The sensitivity of the analysis is increased by splitting the $2\ell ss$ and 3ℓ categories into subcategories of different signal-to-background ratio. We exploit the fact that identification efficiencies are lower and misidentification rates are higher for electrons compared to muons, charge misidentification rates are sizeable for electrons, while negligible for muons, and that several background processes (ttW, WZ, W+jets, and single top production) more often contain leptons of positive compared to negative charge, which is not the case for signal events. Further, backgrounds that are due to the misidentification of non-prompt leptons as prompt ones (predominantly $t\bar{t}$+jets production) typically have fewer reconstructed b-jets compared to the $t\bar{t}H$ signal.

Events selected in the $2\ell ss$ channel are analyzed in subcategories based on the flavor of the leptons (ee, eμ, $\mu\mu$), the charge of the lepton pair ($++$ or $--$), and the number of jets passing the medium WP of the DeepCSV discriminant (≥ 2 or < 2). Events containing two electrons are not categorized by b-jet multiplicity, yielding ten subcategories. Using the label "2bM" to refer to events containing two or more jets passing the medium WP of the DeepCSV discriminant and the label "2bL" to refer to events not containing two such jets, we denote the ten subcategories of the $2\ell ss$ channel by: e^+e^+2bM, e^-e^-2bM, $e^+\mu^+$2bM, $e^-\mu^-$2bM, $\mu^+\mu^+$2bM, $\mu^-\mu^-$2bM, e^+e^+2bL, e^-e^-2bL, $e^+\mu^+$2bL, $e^-\mu^-$2bL, $\mu^+\mu^+$2bL, and $\mu^-\mu^-$2bL.

Events selected in the 3ℓ channel are analyzed in two subcategories, denoted by 2bM and 2bL, which are based on the multiplicity of jets passing the medium WP of the DeepCSV discriminant (≥ 2 or < 2).

7 Data-to-MC corrections

In order to improve the modeling of the data, we apply corrections to simulated events in terms of:

- Pileup reweighting
- Trigger efficiency
- e and μ identification and isolation efficiency
- τ_h identification efficiency
- τ_h energy scale

Table 10: Event selections applied in the $2\ell ss$, $2\ell ss + 1\tau_h$, $1\ell + 2\tau_h$ and $2\ell + 2\tau_h$ categories.

Selection	$2\ell ss$	$2\ell ss + 1\tau_h$
Targeted $t\bar{t}H$ decays	$t \to b\ell\nu$, $t \to bqq$, $H \to WW \to \ell\nu qq$	$t \to b\ell\nu$, $t \to bqq$, $H \to \tau\tau \to \ell\tau_h + \nu$'s
Trigger	Single- and double-lepton triggers	
Lepton p_T	$p_T > 25$ / 15 GeV	$p_T > 25$ / 15 (e) or 10 GeV (μ)
Lepton η	$\lvert\eta\rvert < 2.5$ (e) or 2.4 (μ)	
τ_h p_T	—	$p_T > 20$ GeV
τ_h η	—	$\lvert\eta\rvert < 2.3$
Charge requirements	2 same-sign leptons and charge quality requirements	$\sum_{\ell,\tau_h} q = \pm 1$
Jet multiplicity	$\geq$4 jets	$\geq$3 jets
b tagging requirements	$\geq$1 tight b-tagged jet or $\geq$2 loose b-tagged jets	
Missing transverse momentum	$E_T^{miss}LD > 30$ GeV **	
Dilepton mass	$m_{\ell\ell} > 12$ GeV * and $\lvert m_{ee} - m_Z\rvert > 10$ GeV	

Selection	$1\ell + 2\tau_h$	$2\ell + 2\tau_h$
Targeted $t\bar{t}H$ decays	$t \to b\ell\nu$, $t \to bqq$, $H \to \tau\tau \to \tau_h\tau_h + \nu$'s	$t \to b\ell\nu$, $t \to b\ell\nu$, $H \to \tau\tau \to \tau_h\tau_h + \nu$'s
Trigger	Single-lepton and lepton+τ_h triggers	Single-, double-lepton triggers
Lepton p_T	$p_T > 25$ (e) or 20 GeV (μ)	$p_T > 25$ / 15 (e) or 10 GeV (μ)
Lepton η	$\lvert\eta\rvert < 2.1$	$\lvert\eta\rvert < 2.5$ (e) or 2.4 (μ)
τ_h p_T	$p_T > 30$ / 20 GeV	$p_T > 20$ GeV
τ_h η	$\lvert\eta\rvert < 2.3$	$\lvert\eta\rvert < 2.3$
Charge requirements	$\sum_{\tau_h} q = 0$	$\sum_{\ell,\tau_h} q = 0$
Jet multiplicity	$\geq$3 jets	$\geq$2 jets
b tagging requirements	$\geq$1 tight b-tagged jet or $\geq$2 loose b-tagged jets	
Missing transverse momentum	—	No requirement if $N_j \geq 4$
	—	$E_T^{miss}LD > 45$ GeV †
	—	$E_T^{miss}LD > 30$ GeV otherwise
Dilepton mass	$m_{\ell\ell} > 12$ GeV *	

* Applied on all pairs of leptons that pass loose selection.
** If both leptons are electrons.
† If the event contains a SFOS lepton pair and $N_j \leq 3$.

- b-tag efficiency and mistag rate
- E_T^{miss} resolution and response

7.1 Pileup reweighting

The PU present in the MC samples does not exactly match the PU present in the data. The difference is corrected by reweighting simulated events to match the PU distribution in data [71]. We use the PU reweighting based on the mean of the Poisson distribution. The corresponding PU distribution in data is computed as described in Ref. [72], using a value of 69.2 mb for the inelastic pp scattering cross section. A problem has been reported with the simulation of PU in the `RunIIFall17` MC production [73]. We work around the problem using the distribution in `PileupSummaryInfo::getTrueNumInteractions`, obtained individually for each MC

Table 11: Event selections applied in the 3ℓ, $3\ell + 1\tau_h$ and 4ℓ categories.

Selection	3ℓ	$3\ell + 1\tau_h$
Targeted $t\bar{t}H$ decays	$t \to b\ell\nu$, $t \to b\ell\nu$, $H \to WW \to \ell\nu qq$ $t \to b\ell\nu$, $t \to bqq$, $H \to WW \to \ell\nu\ell\nu$	$t \to b\ell\nu$, $t \to b\ell\nu$, $H \to \tau\tau \to \ell\tau_h + \nu$'s
Trigger	Single-, double- and triple-lepton triggers	
Lepton p_T	$p_T > 25$ / 15 / 10 GeV	$p_T > 20$ / 10 / 10 GeV
Lepton η	$\lvert\eta\rvert < 2.5$ (e) or 2.4 (μ)	
τ_h p_T	—	$p_T > 20$ GeV
τ_h η	—	$\lvert\eta\rvert < 2.3$
Charge requirements	$\sum_\ell q = \pm 1$	$\sum_{\ell,\tau_h} q = 0$
Jet multiplicity	$\geq$2 jets	
b tagging requirements	$\geq$1 tight b-tagged jet or $\geq$2 loose b-tagged jets	
Missing transverse momentum	No requirement if $N_j \geq 4$ $E_T^{miss}LD > 45$ GeV † $E_T^{miss}LD > 30$ GeV otherwise	
Dilepton mass	$m_{\ell\ell} > 12$ GeV * and $\lvert m_{\ell\ell} - m_Z\rvert > 10$ GeV ‡	
Four-lepton mass	$m_{4\ell} > 140$ GeV §	—

Selection	4ℓ
Targeted $t\bar{t}H$ decays	$t \to b\ell\nu$, $t \to b\ell\nu$, $H \to WW \to \ell\nu\ell\nu$ $t \to b\ell\nu$, $t \to b\ell\nu$, $H \to ZZ \to \ell\ell qq$ or $\ell\ell\nu\nu$
Trigger	Single-, double- and triple-lepton triggers
Lepton p_T	$p_T > 25$ / 15 / 15 / 10 GeV
Lepton η	$\lvert\eta\rvert < 2.5$ (e) or 2.4 (μ)
τ_h p_T	—
τ_h η	—
Charge requirements	$\sum_\ell q = 0$
Jet multiplicity	$\geq$2 jets
b tagging requirements	$\geq$1 tight b-tagged jet or $\geq$2 loose b-tagged jets
Missing transverse momentum	No requirement if $N_j \geq 4$ $E_T^{miss}LD > 45$ GeV † $E_T^{miss}LD > 30$ GeV otherwise
Dilepton mass	$m_{\ell\ell} > 12$ GeV * and $\lvert m_{\ell\ell} - m_Z\rvert > 10$ GeV ‡
Four-lepton mass	$m_{4\ell} > 140$ GeV §

* Applied on all pairs of leptons that pass loose selection.
† If the event contains a SFOS lepton pair and $N_j \leq 3$.
‡ Applied to all SFOS lepton pairs.
§ Applied only if the event contains 2 SFOS lepton pairs.

493 sample before any cuts are applied, as input for the PU reweighting.

Lepton multiplicity and flavour	Leading lepton p_T	SF
$2\,\mu$	$p_T < 35$ GeV	0.972 ± 0.006
	$p_T \geq 35$ GeV	0.994 ± 0.001
$e + \mu$	$p_T < 35$ GeV	0.952 ± 0.008
	$35 \leq p_T < 50$ GeV	0.983 ± 0.003
	$p_T \geq 50$ GeV	1.000 ± 0.001
$2\,e$	$p_T < 30$ GeV	0.937 ± 0.027
	$p_T \geq 30$ GeV	0.991 ± 0.002
$\geq 3\,\ell$	–	1.000 ± 0.050

Table 12: Trigger efficiency SF applied to simulated events selected in different categories. The SF applied to events selected in the $1\ell + 2\tau_h$ are described in the text.

7.2 Trigger efficiency

Difference in efficiency for events in data and MC simulation to pass the triggers given in Table 8 are corrected by applying the ratio of the efficiency in data to the efficiency in the MC simulation as weight to simulated events. We refer to these weights as "scale factors" (SF).

In the $2\ell ss$, $2\ell ss + 1\tau_h$, $2\ell + 2\tau_h$, 3ℓ, $3\ell + 1\tau_h$, and 4ℓ channels, the SF are measured as function of lepton multiplicity and type of the leptons as well as of the leading lepton p_T. An additional uncertainty of 2% is attributed to 2ℓ-triggered events, associated with the parametrization of the trigger efficiency as a function of the leading lepton p_T alone. The SF are determined by comparing the combined efficiencies of single, double, and triple lepton triggers between data and MC simulation using events recorded by E_T^{miss} triggers. The sample predominantly consists of $t\bar{t}$+jets events. The SF are given in Table 12.

In the $1\ell + 2\tau_h$ channel, we parametrize the efficiency for events to pass the single lepton trigger and to pass the lepton and τ_h "legs" of the lepton+τ_h cross-triggers as function of p_T and η. The efficiency of these triggers has been measured via the Tag-and-Probe technique [74], using $Z/\gamma^* \to ee$, $Z/\gamma^* \to \mu\mu$, and $Z/\gamma^* \to \tau\tau$ events. The measurement has been performed by the Tau POG [75]. The efficiency for an event containing one lepton and two τ_h to pass the combination of single lepton trigger and lepton+τ_h cross-trigger is given by:

$$\epsilon = \begin{cases} \epsilon_L - \min(\epsilon_L, \epsilon_\ell) \times (1 - (1 - \epsilon_{\tau 1})(1 - \epsilon_{\tau 2})) & \text{if only the single lepton trigger fires} \\ (\epsilon_\ell - \epsilon_L) \times (1 - (1 - \epsilon_{\tau 1})(1 - \epsilon_{\tau 2})) & \text{if only the cross-trigger fires} \\ \min(\epsilon_L, \epsilon_\ell) \times (1 - (1 - \epsilon_{\tau 1})(1 - \epsilon_{\tau 2})) & \text{if both triggers fire} \end{cases} \quad (4)$$

where ϵ_L denotes the efficiency to pass the single lepton triggers, ϵ_ℓ the efficiency to pass the lepton leg of the cross-trigger, and $\epsilon_{\tau(1)}$ ($\epsilon_{\tau(2)}$) refers to the efficiency for the τ_h of higher (lower) p_T to pass the τ_h leg of the cross-trigger. The efficiencies ϵ_L and ϵ_ℓ ($\epsilon_{\tau(1)}$ and $\epsilon_{\tau(2)}$) depend on p_T and η of the lepton (of the two τ_h). Eq. (4) takes into account the fact that any of the two τ_h present in the event may fire the lepton+τ_h cross-trigger, and that although the single lepton trigger has a higher p_T threshold compared to the lepton leg of the cross-trigger, the single lepton trigger may nevertheless have a higher efficiency, due to differences in lepton identification and isolation criteria. The SF applied to events selected in the $1\ell + 2\tau_h$ channel is then given by the ratio of the efficiencies for data and MC simulation, computed separately according to Eq. (4).

7.3 Identification and isolation efficiency for e and μ

The efficiency for electrons and muons to pass the loose and the tight selection criteria defined in Tables 4 and 5 have been measured via the Tag-and-Probe technique using samples of

$Z/\gamma^* \to ee$ and $Z/\gamma^* \to \mu\mu$ events. The measurement described in Ref. [76] has been repeated for data recorded in 2017. The efficiencies are measured separately for electrons and muons and are parametrized as function of lepton p_T and η. The measurement is performed in two stages: the efficiency for leptons to pass the loose selection criteria is measured first, and then the conditional probability for leptons that pass the loose selection criteria to also pass the tight selection criteria is measured. Separate measurements are performed for the case that the tight lepton charge selection criteria are applied and for the case that they are not applied. The latter scale factors are determined separately for every run era. The efficiencies obtained in the complete dataset is shown in fig 2, while their evolution as a function of the run era is shown in figs 3 and 4.

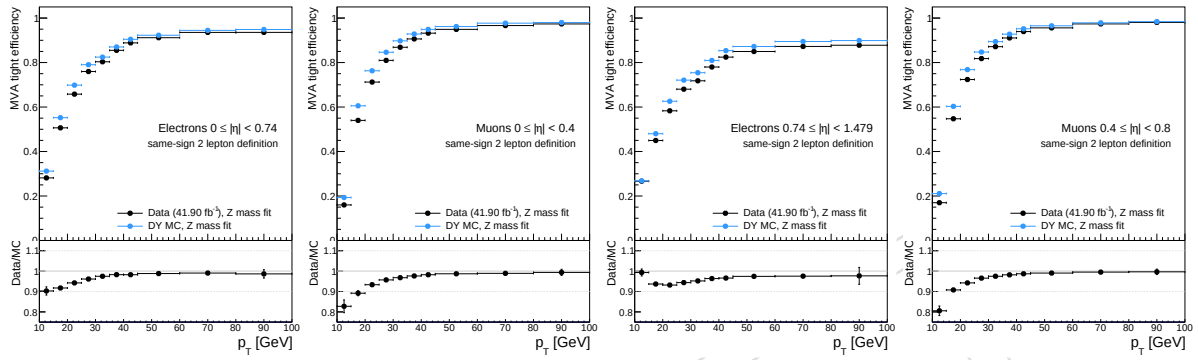


Figure 2: Efficiency of loose leptons to pass the tight selection criteria, as measured in $Z \to \ell\ell$ events using a tag-and-probe method as a function of p_T for different η regions.

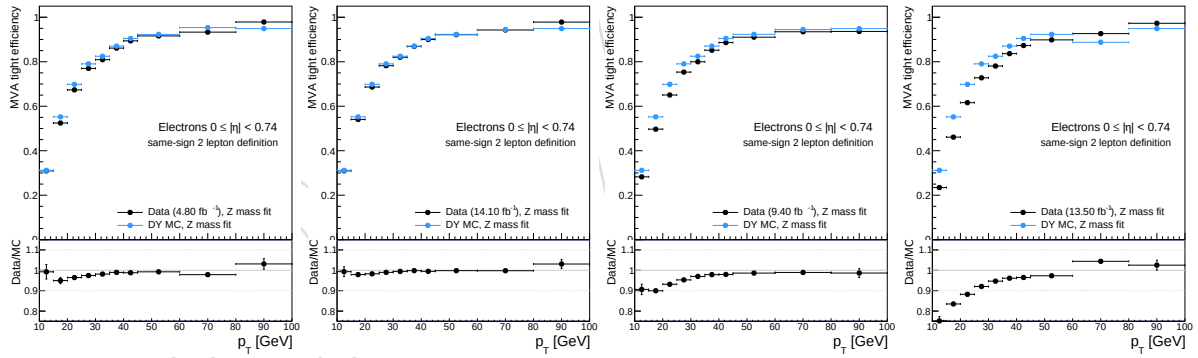


Figure 3: Efficiency of loose electrons to pass the tight selection criteria, as measured in $Z \to \ell\ell$ events using a tag-and-probe method as a function of p_T in the barrel for the B, C+D, E and F run eras, from left to right.

The efficiencies determined for the case that the tight lepton charge selection criteria are applied (not applied) are used for events selected in the $2\ell ss$ and $2\ell ss + 1\tau_h$ categories (in all other categories). The ratio of the efficiency measured in data to the efficiency in MC simulation yields a SF per lepton. The product of per-lepton SF is applied as weight to simulated events. The SF measured for leptons passing the loose selection criteria are also applied to leptons that pass the fakeable lepton selection criteria.

7.4 Identification efficiency for τ

The efficiency for hadronic τ decays to pass the τ_h identification criteria detailed in Section 3.2 has been measured by the Tau POG, using $Z/\gamma^* \to \tau\tau$, $t\bar{t} \to \bar{b}W\,bW \to \bar{b}\mu\nu\,b\tau\nu$, and $W \to \tau\nu$

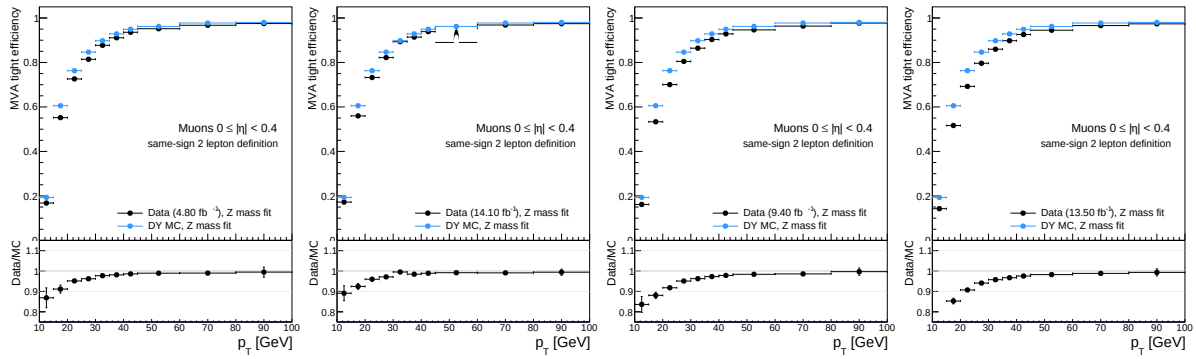


Figure 4: Efficiency of loose muons to pass the tight selection criteria, as measured in Z$\rightarrow \ell\ell$ events using a tag-and-probe method as a function of p_T in the barrel for the B, C+D, E and F run eras, from left to right.

WP	SF
Very-loose	0.85 ± 0.05
Loose	0.89 ± 0.05
Medium	0.89 ± 0.05

Table 13: SF for the τ_h identification efficiency. The SF are applied per reconstructed τ_h that is matched to a genuine τ_h on generator level.

events. The measurement described in Ref. [77] has been repeated for data recorded in 2017. The ratio of the efficiency measured in data to the efficiency in MC simulation depends on the τ_h identification discriminant. We use the SF measured for by the Tau POG for the τ_h identification discriminant with $R = 0.5$ [78], as the SF for the τ_h identification discriminant with $R = 0.3$ have not been measured yet. The SF are given in Table 13 and are applied per reconstructed τ_h that is matched, on generator level, to a genuine hadronic τ decay.

7.5 Energy scale of τ_h

A correction for the energy scale of τ_h (τ_h-ES) is determined by fitting the distribution in the mass of the τ_h candidate and in the mass of muon and τ_h reconstructed in $Z/\gamma^* \rightarrow \tau\tau$ events selected in the decay channel $\tau\tau \rightarrow \mu\nu\tau_h\nu\nu$ with shape templates for the $Z/\gamma^* \rightarrow \tau\tau$ signal and background processes [77]. Separate fits are performed for τ_h candidates reconstructed in the hadronic τ decay modes $h^\pm$, $h^\pm + 1\pi^0$ and $h^\pm + 2\pi^0$, and $h^\pm h^\mp h^\pm$.

The τ_h-ES is lower in data compared to the MC simulation. Depending on the reconstructed τ_h decay mode, the energy of τ_h that are matched, on generator level, to genuine hadronic τ decays in simulated signal and background events is scaled by the factors given in Table 14.

7.6 b-tag efficiency and mistag rate

Small differences between data and MC simulation in the efficiency for b-jets and c-jets to pass the loose and medium WPs of the DeepCSV algorithm, and in the mistag rate for light flavor (u,

Hadronic τ decay mode	SF
$h^\pm$	0.97 ± 0.03
$h^\pm + 1\pi^0$ and $h^\pm + 2\pi^0$	0.98 ± 0.03
$h^\pm h^\mp h^\pm$	0.99 ± 0.03

Table 14: SF for the τ_h-ES.

d, s) quark and gluon jets have been observed by the BTV POG. We correct for the differences by applying suitably chosen weights to simulated events, using SF and tools ("method 1d") provided by the BTV POG [79]. The per-jet SF are parametrized as function of jet p_{T} and η, of the b-tagging discriminant of the jet, and of the flavor of the quark or gluon that is matched to the jet on generator level. The per-event weight is then taken as the product of per-jet weight of all selected jets in the event.

7.7 E_T^{miss} resolution and response

Differences between data and MC simulation in E_T^{miss} response and resolution are corrected by "propagating" measured differences in jet response and resolution [80] to the E_T^{miss}, using the `runMETCorrectionsAndUncertainties` tool provided by the JetMET POG [81]. The validity of the corrections has been verified using $\mathrm{Z}/\gamma^* \to \mathrm{ee}$, $\mathrm{Z}/\gamma^* \to \mu\mu$, and γ+jets events [82].

8 Background estimation

Even small background contributions may be relevant in this analysis, due to the small rate of the $\mathrm{t\bar{t}H}$ signal. Thus, a multitude of different processes needs to be taken into account as backgrounds. We distinguish between "reducible" and "irreducible" background contributions.

A background is considered as "reducible" in case one or more of the reconstructed electrons, muons, or τ_{h} passing the tight object selection criteria detailed in Sections 3.1.4 and 3.2 are not due to genuine prompt leptons or τ_{h}. In the $2\ell ss$ and $2\ell ss + 1\tau_{\mathrm{h}}$ channels, a further source of reducible background arises from events containing lepton pairs of opposite charge, in which the charge of either lepton is mismeasured. Both of these reducible backgrounds are determined from data. We refer to the former as "fake" and to the latter as charge "flip" background. The procedures used to estimate the fake and charge flip backgrounds is described in Sections 8.1 and 8.2, respectively. The procedure for estimating the fake background is validated in control regions. The results of this validation are discussed in Section I.4 of the Appendix.

A small additional reducible background contribution arises from the production of top quark pairs in association with either real or virtual photons. This background has been studied in Refs. [53, 76, 83, 84]. The contribution of $\mathrm{t\bar{t}}$ events with real photons is typically due to asymmetric conversions of the type $\gamma \to \mathrm{e^+e^-}$, in which either the electron or the positron carries most of the energy of the photon, while the other electron or positron is of low energy and fails to get reconstructed. Such events are suppressed by the photon conversion rejection criteria and by requiring that the tracks of electron candidates have hits in each layer of the pixel detector that is crossed by the track (see Table 4). In case $\mathrm{t\bar{t}}$ events are produced in association with virtual photons, the virtual photons typically produce electron or muon pairs of low mass, which are suppressed very effectively by the $m_{\ell\ell} > 12$ GeV cut that is applied in all channels. As a consequence, the background contribution of $\mathrm{t\bar{t}}$ events with photons is small in the $2\ell ss$ and $2\ell ss + 1\tau_{\mathrm{h}}$ channels, and negligible in all other channels. The contribution of this background is modeled using the MC simulation.

We treat the uncertainties on background contributions that are due to conversions of real photons into $\mathrm{e^+e^-}$ pairs as fully correlated between the $\mathrm{t\bar{t}}$ background and other background processes, in which the reconstructed electrons are due to photon conversions. For the purpose of correlating uncertainties on the conversion background between different MC samples, we refer to the background contribution arising from events in which at least one reconstructed electron is matched, within $\Delta R < 0.3$, to a photon on generator level as "conversion" background. The p_{T} of the generator-level photon is required to exceed $0.5\,p_{\mathrm{T}}$ of the reconstructed

electron. All MC samples given in Table 2 are used to estimate the conversion background. All simulated events in which at least one reconstructed electron is due to photon conversion are included in the estimate of the conversion background, unless the event qualifies as fake or charge flip background. The latter events are excluded from the estimate of the conversion background, to avoid double-counting of background contributions.

Irreducible background contributions are modeled using the MC simulation. The dominant contributions are due to the production of top quark pairs in association with W or Z bosons and to the production of W or Z boson pairs in association with jets. Minor contributions arise from triboson production and from the production of single top quarks in association with W or Z bosons. The modeling of the data by the simulation is validated in specific control regions, each enriched in the contribution of one of the dominant irreducible background processes $t\bar{t}W$, $t\bar{t}WW$, $t\bar{t}Z$, and WZ+jets. Results of these studies are described in Sections I.1, I.2, and I.3 of the Appendix. The $t\bar{t}W$ and $t\bar{t}WW$ backgrounds are studied together.

In order to avoid double-counting background contributions that are included in the fake and charge flip background estimates obtained from data with background contributions modeled by the MC simulation, reconstructed electrons, muons and τ_h in simulated events are required to be matched, within $\Delta R < 0.3$, to genuine prompt leptons and hadronic τ decays. Electrons and muons are considered to be prompt if they originate from the decay of either a W boson, a Z boson, or a τ lepton. Reconstructed τ_h that are matched to an electron or muon on generator level are considered as irreducible background. The background contribution arising from the misidentification of electrons and muons as τ_h is small, as those electrons and muons that fail the loose object selection criteria get considered as τ_h. This background is modeled using the MC simulation.

8.1 "Fake" background

The estimation of the fake background is based on the fake-factor (FF) method. The method is applied to each of the channels $1\ell + 2\tau_h$, $2\ell ss$, $2\ell ss + 1\tau_h$, $2\ell + 2\tau_h$, 3ℓ, $3\ell + 1\tau_h$, and 4ℓ separately. In the $1\ell + 2\tau_h$, $2\ell ss$, $2\ell + 2\tau_h$, 3ℓ, and 4ℓ channels, the method selects events, which pass all selection criteria for the respective category, detailed in Section 5, except that the electrons, muons, and τ_h in these events are required to pass the fakeable instead of the tight object selection criteria. The selected event sample is referred to as the "application region" (AR) of the FF method. Events in which all e, μ, and τ_h pass the tight object selection criteria are vetoed in order to avoid overlap with the signal region (SR). An estimate of the fake background in the SR is obtained by applying appropriately chosen weights to the events selected in the AR.

The weights depend on the number of e, μ, and τ_h that pass the fakeable, but fail the tight object selection criteria. Expressions for the weights are derived in Refs. [83, 85]. For events containing a total of 2, 3, or 4 fakeable "objects", where "object" refers to either e, μ, or τ_h, the

expressions read:

$$
\begin{aligned}
N_{\text{pp}}^{\text{fake}} &= \sum_{\text{fp}} F_1 + \sum_{\text{pf}} F_2 - \sum_{\text{ff}} F_1 F_2 \\
N_{\text{ppp}}^{\text{fake}} &= \sum_{\text{fpp}} F_1 + \sum_{\text{pfp}} F_2 + \sum_{\text{ppf}} F_3 \\
&\quad - \sum_{\text{ffp}} F_1 F_2 - \sum_{\text{fpf}} F_1 F_3 - \sum_{\text{pff}} F_2 F_3 + \sum_{\text{fff}} F_1 F_2 F_3 \\
N_{\text{pppp}}^{\text{fake}} &= \sum_{\text{fppp}} F_1 + \sum_{\text{pfpp}} F_2 + \sum_{\text{ppfp}} F_3 + \sum_{\text{pppf}} F_4 \\
&\quad - \sum_{\text{ffpp}} F_1 F_2 - \sum_{\text{fpfp}} F_1 F_3 - \sum_{\text{fppf}} F_1 F_4 - \sum_{\text{pffp}} F_2 F_3 - \sum_{\text{pfpf}} F_2 F_4 - \sum_{\text{ppff}} F_3 F_4 \\
&\quad + \sum_{\text{fffp}} F_1 F_2 F_3 + \sum_{\text{ffpf}} F_1 F_2 F_4 + \sum_{\text{fpff}} F_1 F_3 F_4 + \sum_{\text{pfff}} F_2 F_3 F_4 - \sum_{\text{ffff}} F_1 F_2 F_3 F_4 \,. \qquad (5)
\end{aligned}
$$

The symbol F_i is a short-hand notation for $F_i \equiv \frac{f_i}{1-f_i}$. The label "pfpp" refers to events in which the object of highest, third-highest, and fourth-highest p_{T} pass the tight object selection criteria, while the object of second highest p_{T} fails the tight object selection criteria, and similar for the other labels. The FF f_i represents the probability for a e, μ, or τ_{h} that passes the fakeable selection criteria to pass the tight selection criteria, where i refers to the i-th e, μ, or τ_{h}, sorted in the order of decreasing p_{T}. The different sums correspond to specific combinations of leptons and hadronic τ decays passing and failing the tight selection criteria. Events enter the sums with a weight the magnitude of which is equal to a product of factors F_i, each factor F_i representing a single lepton or hadronic τ decay that passes the fakeable, but fails the tight selection criteria. For example, the sum $\sum_{\text{fpf}}$ extends over all events that contain three e, μ, or τ_{h} passing the fakeable selection criteria, of which the "leading" and "third" ones fail and the "subleading" one passes the tight selection criteria, and the events in the sum are weighted by the product $F_1 F_3$. The sign of the weights alternates for events with different numbers of e, μ, or τ_{h} failing the tight selection criteria. The symbols $N_{\text{pp}}^{\text{fake}}$ and $N_{\text{ppp}}^{\text{fake}}$ of the left-hand-side of the equations represent the estimated contributions of the "fake" background in the signal region, determined by the method, in case the FF method is applied to events containing 2 and 3 e, μ, and τ_{h}, respectively. The contamination from irreducible backgrounds with prompt leptons and genuine τ_{h} in the application region is subtracted from this estimate to avoid double-counting in the signal region.

The FF f_i are measured separately for e, μ, and τ_{h} and are parametrized as function of η and cone-p_{T} of the lepton. The measurement of the FF is described in Sections F and G of the Appendix. The control regions used to measure the FF are referred to as "measurement regions" (MR). The event selection criteria that are applied in the MR used to measure the FF for leptons (τ_{h}) are chosen such that the relative fractions of non-prompt leptons and hadrons (of quark jets of different flavor and gluon jets) are similar between AR and MR. Likewise, the fakeable lepton selection criteria are chosen such as to reduce the difference between the FF for non-prompt leptons and hadrons. These choices guarantee that the FF yield an unbiased estimate of the fake background in the SR when the weights given by Eq. (5) are applied to the events in the AR. Remaining differences of the fake factors for e and μ between the measurement and application regions have been studied in simulated events and found to be small (see Fig. 48 in Ref. [83] for illustration). The FF for τ_{h} are measured separately for the loose and medium τ_{h} ID discriminants.

In the $2\ell ss + 1\tau_{\text{h}}$ and $3\ell + 1\tau_{\text{h}}$ channels, a modified version of the FF method is used. In the modified version, only the lepton selection criteria are relaxed in the AR, while the τ_{h} selection

criteria are kept tight, and the FF are applied to the leptons only. The estimate of the "fake" background obtained in the $2\ell ss + 1\tau_h$ ($3\ell + 1\tau_h$) channel is given by the expression N_{pp}^{fake} (N_{ppp}^{fake}) in Eq. (5), i.e. the τ_h is not considered in the count of fakeable objects, the expression suitable for events containing 2 (3) fakeable objects is used, and the factors F_i refer to the leptons only. The contribution of background events, in which the reconstructed leptons match prompt leptons at generator level and in which the reconstructed τ_h is due to the misidentification of a quark and gluon jet, is taken from the MC simulation. The latter background contribution predominantly arises from $t\bar{t}W$, $t\bar{t}WW$, $t\bar{t}Z$, and diboson production. The motivation for this modification is that, in the $2\ell ss + 1\tau_h$ and $3\ell + 1\tau_h$ categories, in about one third of selected $t\bar{t}H$ signal events the reconstructed τ_h is a misidentified quark or gluon jet (mainly for $H \to WW$ decays). These events would be included in the estimate of the "fake" background, in case the FF method was applied to the $2\ell ss + 1\tau_h$ and $3\ell + 1\tau_h$ channels without this modification, and could not be used for the purpose of inferring the production rate of the $t\bar{t}H$ signal, which would reduce the signal-rate parameter (μ) sensitivity by $\sim 30\%$ (in the channels). The loss in sensitivity can be reduced by applying tighter τ_h identification discriminants, but the sensitivity would still be suboptimal compared to using the modified FF method for these channels.

8.2 Charge "flip" background

The charge "flip" background in the $2\ell ss$ and $2\ell ss + 1\tau_h$ categories is dominated by $t\bar{t}$+jets events with two prompt leptons, produced in the decay $t\bar{t} \to bW^+\bar{b}W^- \to b\ell^+\nu\bar{b}\ell^-\bar{\nu}$, in which the charge of either prompt lepton is mismeasured. The background is estimated from data, following a strategy similar to the one used for the estimation of the fake background. The control region used to estimate the charge flip background contains events that pass all selection criteria of the SR, except that the two leptons are required to be of opposite charge. The sum of the probabilities to mismeasure the charge of either one of the two leptons is then applied as event weight. The charge misidentification rates for electrons are measured using $Z/\gamma^* \to ee$ events and are parametrized as function of p_T and η of the electron. The measurement is described in Section H of the Appendix. The charge misidentification rates for muons are negligible [69]. As a consequence, the charge flip background in the $2\ell ss$ and $2\ell ss + 1\tau_h$ categories is significantly higher for events containing either two electrons or one electron and one muon compared to events that contain two muons.

9 Systematic uncertainties

Various imprecisely known or simulated effects may alter the event yield of the $t\bar{t}H$ signal and of background processes, as well as the shape of the distributions in the "discriminating" observables that are used for the signal extraction. The following sources are considered as systematic uncertainties in the analysis:

- **Trigger efficiency**
  The uncertainty for events selected in the $2\ell ss$, $2\ell ss + 1\tau_h$, $2\ell + 2\tau_h$, 3ℓ, $3\ell + 1\tau_h$, and 4ℓ channels to pass the trigger requirements are given in Table 12. The uncertainty on the efficiency to pass the trigger requirements is fully correlated among the 3ℓ, $3\ell + 1\tau_h$, and 4ℓ channels. In the $2\ell ss$, $2\ell ss + 1\tau_h$, and $2\ell + 2\tau_h$ channels, the uncertainties are treated as fully correlated between events containing leptons of the same flavour, while the uncertainties between events containing leptons of different flavour ($2\,\mu$, $e + \mu$, or $2\,e$) are treated as uncorrelated. In the $1\ell + 2\tau_h$ channel, the uncertainties on the trigger efficiency SF of the single lepton triggers, as well as of the lepton and τ_h legs of the lepton+τ_h cross-triggers, depend on the p_T and η of the lepton and

τ_h. The resulting effect on the normalization amounts to 3%, while the effect on the shape is found to be small and neglected in the analysis.

- **Identification and isolation efficiency for** e **and** μ
  The uncertainty for electrons (muons) to pass the loose, fakeable, or tight selection criteria defined in Table 4 (5) amounts to 3% for electrons of $p_T > 25$ GeV and 5% for electrons of $p_T < 25$ GeV) (2% for muons of all p_T). See Section 7.3 and Ref. [86] for details.
- **Identification efficiency for** τ_h
  The uncertainty to reconstruct and identify hadronic τ decays amounts to 5% [87]. The uncertainty is fully correlated among the $1\ell + 2\tau_h$, $2\ell ss + 1\tau_h$, $2\ell + 2\tau_h$, and $3\ell + 1\tau_h$ channels. The resulting effect on the normalization amounts to 5% (10%) in the $2\ell ss + 1\tau_h$ and $3\ell + 1\tau_h$ ($1\ell + 2\tau_h$ and $2\ell + 2\tau_h$) channels.
- **Energy scale of** e, μ, **and** τ_h
  The energy scales of electrons and muons are known with an uncertainty of less than 1% and are neglected in the analysis. The energy scale of τ_h is varied by 1.2%, following the recommendation of the Tau POG [87].
- **Jet energy scale**
  Jet energy scale uncertainties are provided as function of jet p_T and η by the JetMET POG [88].
- b**-tag efficiency and mistag rate**
  Uncertainties on b-tagging efficiencies and mistag rates as function of jet p_T and η are provided by the BTV POG [89]. The effect of these uncertainties on the analysis is evaluated by varying the data-to-MC correction factors described in Section 7.6 within their uncertainties and reanalyzing the events. We build separate shape templates for each individual source of systematic uncertainty on the b-tag efficiency and on the mistag rate. While the effect of the uncertainties on b-tag efficiency and mistag rate is small compared to the statistical uncertainties, and the splitting of the uncertainties into individual sources is not absolutely necessary for our analysis, it simplifies future combinations with $t\bar{t}H$ analyses performed in other decay channels, notably $H \to b\bar{b}$, $H \to \gamma\gamma$, and $H \to ZZ \to 4\ell$.
- E_T^{miss} **resolution and response**
  Uncertainties on E_T^{miss} resolution and response are accounted for by varying the jet energy scale and resolution within their respective uncertainties and recomputing E_T^{miss} and all E_T^{miss} related observables after each such variation. The variations are performed using the `runMETCorrectionsAndUncertainties` tool developed by the JetMET POG [81].
- **Signal rate**
  As the signal rate is measured in units of the SM $t\bar{t}H$ production rate, the measurement is affected by uncertainties on the $t\bar{t}H$ cross section. The uncertainty on the SM $t\bar{t}H$ cross section, computed at NLO accuracy, amounts to $^{+6.8\%}_{-10.0\%}$, of which $^{+5.8\%}_{-9.3\%}$ are due to missing higher orders and 3.6% arises from uncertainties on the PDF and α_s [90]. The uncertainty on the branching fraction for the H boson to decay into WW and ZZ ($\tau\tau$) amounts to 1.5% (1.7%). Uncertainties on the acceptance that are due to missing higher orders are treated as shape systematics, as they affect the shape of the distribution in the observable used for signal extraction, and are estimated by varying the renormalization (μ_R) and factorization (μ_F) scales scales between 0.5 and 2 times their default values, with the constraint that $0.5 \leq \mu_F/\mu_R < 2$.

- **Background rates**
  The contribution of the WZ+jets background is known with an uncertainty of 50%. The uncertainty represents the uncertainness on the extrapolation from the WZ+jets dominated control region described in Section I.3 of the Appendix to the SR. It includes the statistical uncertainty in the control region, the uncertainness on the b-tag efficiency and mistag rate, and the uncertainty on the relative composition of light quark and gluon, c quark, and b quark jets in WZ+jets events contributing to the control region and to the signal region by the MC simulation [76]. The uncertainty on the background arising from production of top quark pairs in association with real or virtual photons amounts to 50%. The yield of the fake background amounts to 30% in the $2\ell ss$, $2\ell ss + 1\tau_h$, 3ℓ, $3\ell + 1\tau_h$, and 4ℓ categories and to 50% in the $1\ell + 2\tau_h$ and $2\ell + 2\tau_h$ categories. Uncertainties on the shape of the fake background arise from statistical uncertainties in the MR and AR, from the subtraction of the prompt lepton contamination in the MR, and from differences in the background composition between MR (dominated by multijet background) and AR (dominated by $t\bar{t}$+jets background). The latter is quantified by comparing the shape templates obtained by selecting simulated $t\bar{t}$+jets in the AR and applying the weights given by Eq. (5) to these events to the shape templates obtained by requiring simulated $t\bar{t}$+jets to pass the event selection criteria of the SR. The FF f_i for leptons are obtained from simulated multijet events that pass the event selection criteria detailed in Section F, while the f_i for τ_h are obtained from simulated simulated $t\bar{t}$+jets that pass the selection criteria described in Section G of the Appendix. We refer to the shape templates obtained from simulated $t\bar{t}$+jets events passing the selection criteria of the SR as "nominal MC" shape templates, and to those obtained from simulated $t\bar{t}$+jets events selected in the AR and weighted by the expressions given by Eq. (5) as the "MC closure" shape templates. The ratio of the nominal MC to the MC closure shape templates in each channel is fitted by a linear function and the deviation of the constant term (slope) from 1 (0) is taken as an additional uncertainty. The fit is illustrated in Fig. **??** for the $1\ell + 2\tau_h$ channel. The yield of the charge flip background in the $2\ell ss$ and $2\ell ss + 1\tau_h$ categories is known with an uncertainty of 30%. No assumption is made on the rates of the irreducible $t\bar{t}W$, $t\bar{t}WW$, and $t\bar{t}Z$ backgrounds. The yield of these backgrounds is instead determined simultaneously with the rate of the $t\bar{t}H$ signal by the maximum-likelihood fit that is used for the signal extraction, which is described in Section 10. An additional uncertainty of 30% is attributed to the contribution of $t\bar{t}W$, $t\bar{t}WW$, and $t\bar{t}Z$ backgrounds, which arises in the $2\ell ss + 1\tau_h$ and $3\ell + 1\tau_h$ categories from events in which the reconstructed leptons are genuine prompt leptons and the reconstructed τ_h is due to the misidentification of a quark or gluon jet, and that are modeled using the MC simulation. An uncertainty of 50% is assigned to other, rare, backgrounds the contribution of which is modeled using the MC simulation.
- **Luminosity**
  The uncertainty on the integrated luminosity amounts to 2.3% [91].
- **Pileup**
  Uncertainties on the instantaneous luminosity and on the pp inelastic cross section may affect the event yield of the $t\bar{t}H$ signal and of backgrounds obtained from the MC simulation, as the efficiency to pass the event selection criteria detailed in Section 5 may vary with PU conditions. We vary the product of instantaneous luminosity and pp inelastic cross section, i.e. the number of PU interactions, by 5%, reweight simulated events to the different PU conditions, and rerun the analysis. We find that

the effect on the signal and of background yields amounts to less than 1%.

10 Signal extraction

The rate of the $t\bar{t}H$ signal is determined by means of a maximum-likelihood (ML) fit to the distributions in "discriminating" observables. The observables are chosen such that the maximal separation, in a shape of the distribution, is attained between the $t\bar{t}H$ signal and backgrounds. The differences in a shape (differential analysis) allow the fit to determine the $t\bar{t}H$ signal rate more precisely compared to simply counting events (scalar analysis), thereby improving the sensitivity of the analysis. One such discriminating observable is chosen per channel. The observables used in different channels are explained in detail in Section 11.

The likelihood function $\mathcal{L}\left(\text{data} \mid \mu, \Theta\right)$ depends on the $t\bar{t}H$ signal rate, which is taken as the parameter of interest (POI) in the fit and is denoted by μ, and on the values of nuisance parameters θ_k:

$$\mathcal{L}\left(\text{data} \mid \mu, \Theta\right) = \prod_i \mathcal{P}\left(n_i|\mu, \Theta\right) \prod_k \rho\left(\tilde{\theta}_k|\theta_k\right) . \tag{6}$$

The index i refers to individual bins of the distributions in the discriminating observables. All channels are fitted simultaneously. The nuisance parameters θ_k represent the systematic uncertainties discussed in Section 9. The set of all nuisance parameters θ_k is denoted by the symbol Θ. The probability to observe n_i events in a given bin i, when $\nu_i(\mu, \Theta)$ events are expected in that bin is given by the Poisson distribution:

$$\mathcal{P}\left(n_i|\mu, \Theta\right) = \frac{\left(\nu_i(\mu, \Theta)\right)^{n_i}}{n_i!} \exp\left(-\nu_i(\mu, \Theta)\right) . \tag{7}$$

The number of events expected in each bin corresponds to the sum of the number of signal (ν_i^{S}) and background (ν_i^{B}) events: $\nu_i(\mu, \Theta) = \nu_i^{\mathrm{S}}(\mu, \Theta) + \nu_i^{\mathrm{B}}(\Theta)$. The estimate in the number of background events is obtained as described in Section 8. The number of signal events is proportional to μ, with the coefficient of proportionality depending on the signal acceptance and on the signal selection efficiency, both are estimated from the MC simulation.

The function $\rho\left(\tilde{\theta}_k|\theta_k\right)$ represents the probability to observe a value $\tilde{\theta}_k$ in an auxiliary measurement of the nuisance parameter, given that the true value is θ_k. The nuisance parameters are treated via the frequentist paradigm, as described in Refs. [92, 93]. Systematic uncertainties that affect only the normalization, but not the shape of the distribution in the discriminating observable, are represented by a Gamma distribution if they are of statistical origin, e.g. corresponding to the number of events observed in a control region, and, otherwise, by a log-normal probability density function. Systematic uncertainties that affect the shape of the distribution are incorporated into the likelihood fit via the technique detailed in Ref. [94] and represented by Gaussian probability density functions.

Additional uncertainties arise from the limited statistics available to model the shape of the distribution in the discriminating observable for the $t\bar{t}H$ signal and background processes. Limitations in event statistics on the shape templates are accounted for by the approach described in Refs. [95, 96]. Further nuisance parameters θ_k are added to the likelihood function $\mathcal{L}\left(\text{data} \mid \mu, \Theta\right)$, which allow the fit to vary the number of events expected in a given bin within its statistical uncertainties. In order to make the ML fit less demanding in terms of computing time, we reduce the number of bins for which we consider statistical fluctuations by neglecting statistical uncertainties which are smaller than 10% times the background contribution expected in a given bin.

Results are presented both in terms of the measured value of the $t\bar{t}H$ signal rate and its associated uncertainty, and in terms of an upper limit on the signal rate. The measured value of the signal rate, μ_{obs}, is given by the value of μ which maximizes the likelihood function $\mathcal{L}(\text{data} \mid \mu, \Theta)$. Its associated uncertainty is given by the value of μ for which the likelihood function, which follows the χ^2_1, decreases by half a unit with respect to the maximum corresponding to a coverage probability of 68%. Upper limits on the signal rate, at 95% confidence level (CL), are computed following a modified frequentist approach, known as the CL_s method [97]. Further details about the CL_s method are given in [98]. We use the implementation provided in the `combine` [99] tool.

11 Discriminating observables for signal extraction

The observables used for the signal extraction are based on the output of BDTs which are trained to separate the $t\bar{t}H$ signal from the irreducible $t\bar{t}W$ and $t\bar{t}Z$ backgrounds and from the $t\bar{t}$+jets background. The $t\bar{t}$+jets background constitutes the dominant contribution to the reducible background in the analysis. In the $1\ell + 2\tau_h$, $2\ell ss + 1\tau_h$, $2\ell + 2\tau_h$, and $3\ell + 1\tau_h$ categories, to which we refer to as "channels with taus", a single BDT is trained in each category to separate the $t\bar{t}H$ signal from the irreducible as well as from the reducible backgrounds. The distribution in BDT output is binned such that the relative statistical uncertainty on the sum of backgrounds amounts to less than 30% in each bin. In the $2\ell ss$ and 3ℓ categories, referred to as "channels without taus", two separate BDTs per category are used. The first (second) BDT is trained to separate the $t\bar{t}H$ signal from the irreducible (reducible) background. The output of both BDTs is combined into a single discriminant by the procedure described in Section D in the Appendix of Ref. [76], based on a likelihood-ratio ordering performed on the phase space described by the outputs of the two BDTs, combined with a clustering algorithm that identifies the optimal number of bins for the 1-dimensional distribution. We refer to the BDTs used for the signal extraction in the $1\ell + 2\tau_h$, $2\ell ss$, $2\ell ss + 1\tau_h$, $2\ell + 2\tau_h$, 3ℓ, and $3\ell + 1\tau_h$ channels as "event-level" BDTs. In the 4ℓ category the event statistics is too low to allow the use of a variable shape. As a result, there the signal extraction is based on a simple event counting instead.

The observables used as BDT inputs for the channels with and without taus are described in Sections 11.1 and 11.2.

The `TMVA` package [7] is used to train the event-level BDTs for the channels without taus. In the channels with taus, the `xgboost` training algorithm [100] is used instead. The performance of the xgboost training algorithm has been compared to the performance of the TMVA package, in the channels with taus, and it has been found to improve the sensitivity by 10–15%, see Ref. [101]. We use the `scikit-learn` package [8] to perform the BDT training with the xgboost algorithm and convert the trained BDT to the TMVA format using a customized version of the mlglue, see Ref. [102]. The xgboost training algorithm uses a different normalization convention for the BDT output compared to the TMVA package, thus it is necessary to transform all the trained BDTs. The transformation is defined as the following:

$$\mathcal{D}_{\text{BDT}} = \frac{\sqrt{1 + \mathcal{D}_{\text{BDT}}}}{\sqrt{1 + \mathcal{D}_{\text{BDT}}} + \sqrt{1 - \mathcal{D}_{\text{BDT}}}}\,. \tag{8}$$

The events selected in the $1\ell + 2\tau_h$, $2\ell ss$, and $2\ell ss + 1\tau_h$ categories are expected to contain a hadronically decaying top quark. A BDT-based algorithm has been developed to reconstruct hadronic top quark decays in these events. The algorithm, referred to as "hadronic top tagger", is described in details in Section C of the Appendix. The algorithm computes the likelihood for

a given combination of three jets, referred to as "jet triplet", to be compatible with originating from the hadronic decay of a top quark. All jet triplets are given to the algorithm. The p_T and η of the hadronic top candidate, corresponding to the jet triplet with the highest BDT output, as well as the value of the BDT output for this triplet are given as inputs to the event-level BDTs used for the signal extraction.

Events selected in the $2\ell ss$ category are expected to contain a decay of the type $H \to WW^* \to \ell\nu qq'$. One of the jets produced in the hadronic decay of the W boson may be outside of the $|\eta| < 2.4$ region, where we select the jets, or may fail the condition $p_T > 25$ GeV. A dedicated algorithm, to which we refer to as "Hj tagger", has been developed to identify the jets produced in decays of the type $H \to WW^* \to \ell\nu qq'$ in events, where one of the jets produced in the W boson decay fails to get reconstructed. The algorithm is based on a BDT approach. It has been used in the previous $t\bar{t}H$ analysis with the data recorded in 2016 and is described in Section 5.1.2 of Ref. [76]. The BDT has been retrained for the 2017 data-taking conditions. The updated algorithm and its performance is described in Section D of the Appendix. **N**ote: The updated Hj tagger is not yet used by the event-level BDT that is used for the signal extraction in the $2\ell ss$ category in the present analysis.

Discriminants based on the matrix element method (MEM) [5, 6] have been developed to improve the separation of the $t\bar{t}H$ signal from backgrounds in the $2\ell ss + 1\tau_h$ and 3ℓ categories [76, 103]. The MEM discriminant developed for the 3ℓ channel is used as an input to the event-level BDT for this channel, while the MEM discriminant developed for the $2\ell ss + 1\tau_h$ channel is currently being optimized for the 2017 data-taking conditions and is not yet used in the present analysis.

11.1 Multivariate analysis for categories with hadronic taus

The BDTs have been trained on simulated samples of $t\bar{t}H$ signal and $t\bar{t}W$, $t\bar{t}Z$, and $t\bar{t}$+jets background events. The samples are given in Table 15. They have been produced using the fast simulation (`FastSim`) [104] for 2016 data-taking conditions. We plan to perform a new BDT training once the full MC statistics from the `RunIIFall17` campaign is available.

The BDTs are trained on events passing the selection criteria described in Section 5. We increase the MC statistics available for BDT training by relaxing the lepton and τ_h selection criteria. In the $2\ell ss + 1\tau_h$ and $3\ell + 1\tau_h$ categories, the electrons and muons in events used for BDT training are merely required to pass the loose lepton selection. The τ_h selection criteria are relaxed to the loose WP of the τ_h identification discriminant in the $2\ell ss + 1\tau_h$ category, to the very-loose WP in the $3\ell + 1\tau_h$ category, and to the very-very-loose WP in the $2\ell + 2\tau_h$ category. Note that the latter WPs refer to the `2015` training of the τ_h identification discriminant and do not directly correspond to the WPs of the `2017v2` training, which is used for the analysis of data recorded in 2017 (*cf.* Section 3.2) [105]. The relaxed lepton and τ_h selection criteria affect the relative fractions of genuine versus fake leptons and τ_h in the sample of simulated signal and background events used to train the event-level BDTs for the signal extraction. This bias, which would result in a suboptimal performance of the trained BDTs in separating the $t\bar{t}H$ signal from backgrounds, is avoided by reweighting the events such that the relative fractions of genuine leptons and τ_h versus fakes in the event samples used for BDT training match the fractions in the SR of the respective category.

Receiver operating characteristic (ROC) curves that quantify the separation of the $t\bar{t}H$ signal from backgrounds are shown in Fig. 5. The ROC curves demonstrate that the bias, which results from the relaxed lepton and τ_h selection that is applied to increase the MC statistics available for BDT training, is small, and also show that the BDT performance is very similar

Table 15: List of MC samples used for BDT training.

Dataset name	Number of events
/TTTo2L2Nu_TuneCUETP8M2_ttHtranche3_13TeV-powheg-pythia8/matze-faster_v8_ttjets_dl_maod_p1_3a2fa29ab1d54ae0995b28f27b405be9-v1/USER /TTTo2L2Nu_TuneCUETP8M2_ttHtranche3_13TeV-powheg-pythia8/matze-faster_v8_ttjets_dl_maod_p2_3a2fa29ab1d54ae0995b28f27b405be9-v1/USER /TTTo2L2Nu_TuneCUETP8M2_ttHtranche3_13TeV-powheg-pythia8/matze-faster_v8_ttjets_dl_maod_p3_3a2fa29ab1d54ae0995b28f27b405be9-v1/USER	90528029
/TTToSemilepton_TuneCUETP8M2_ttHtranche3_13TeV-powheg-pythia8/matze-faster_v8_ttjets_sl_maod_p1_3a2fa29ab1d54ae0995b28f27b405be9-v1/USER /TTToSemilepton_TuneCUETP8M2_ttHtranche3_13TeV-powheg-pythia8/matze-faster_v8_ttjets_sl_maod_p2_3a2fa29ab1d54ae0995b28f27b405be9-v1/USER /TTToSemilepton_TuneCUETP8M2_ttHtranche3_13TeV-powheg-pythia8/matze-faster_v8_ttjets_sl_maod_p3_3a2fa29ab1d54ae0995b28f27b405be9-v1/USER	85311662
/ttHToNonbb_M125_TuneCUETP8M2_ttHtranche3_13TeV-powheg-pythia8/matze-faster_v8_ttH_maod_p1_3a2fa29ab1d54ae0995b28f27b405be9-v1/USER /ttHToNonbb_M125_TuneCUETP8M2_ttHtranche3_13TeV-powheg-pythia8/matze-faster_v8_ttH_maod_p2_3a2fa29ab1d54ae0995b28f27b405be9-v1/USER /ttHToNonbb_M125_TuneCUETP8M2_ttHtranche3_13TeV-powheg-pythia8/matze-faster_v8_ttH_maod_p3_3a2fa29ab1d54ae0995b28f27b405be9-v1/USER	16459381
/TTWJetsToLNu_TuneCUETP8M1_13TeV-amcatnloFXFX-madspin-pythia8/matze-faster_v9_ttW_maod_54aa74f75231422e9f4d3766cb92a64a-v1/USER	7158309
/TTZToLLNuNu_M-10_TuneCUETP8M1_13TeV-amcatnlo-pythia8/matze-faster_v9_ttZ_maod_54aa74f75231422e9f4d3766cb92a64a-v1/USER	8826709

when evaluated on MC samples produced with the full CMS detector simulation (FullSim), based on GEANT4, and on MC samples produced with the fast simulation. The input variables used to each of the the individual BDT are listed on Tab. 16 and their ranking in performance to separate the $t\bar{t}H$ signal from backgrounds is shown in Fig. 6.

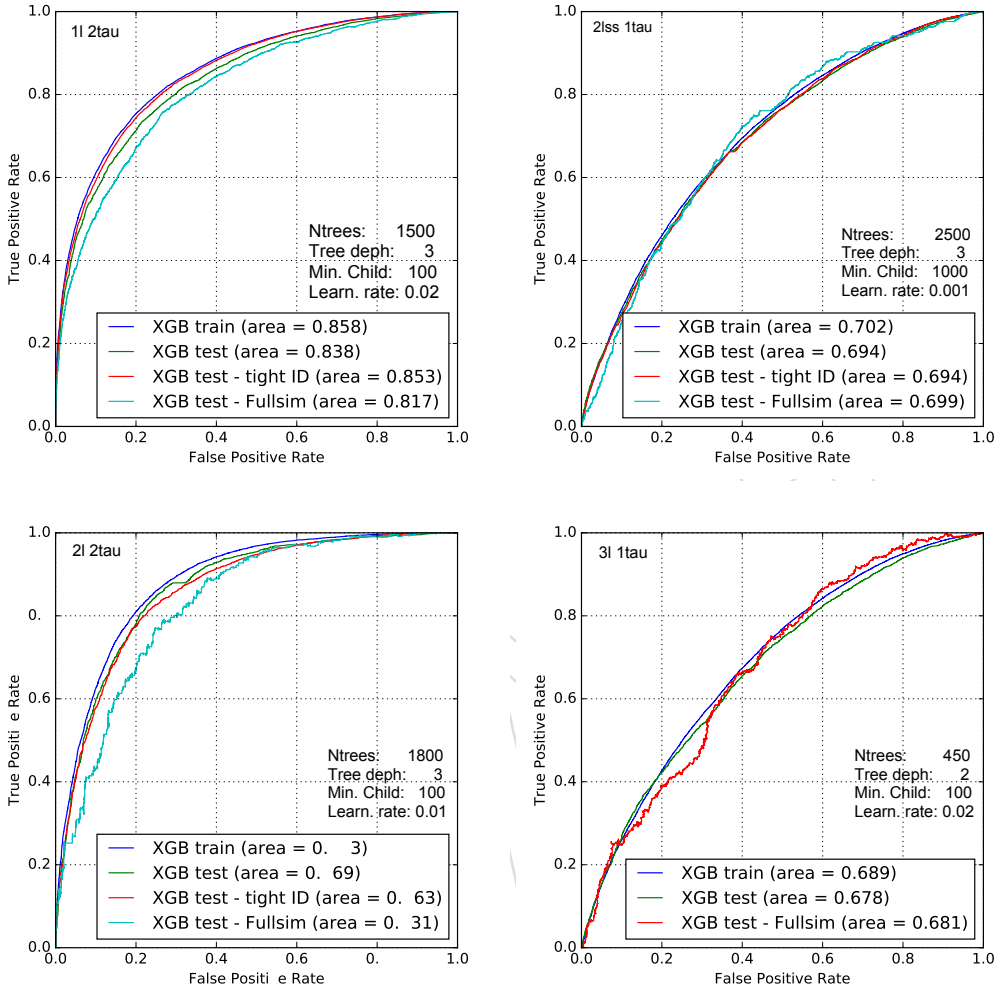


Figure 5: ROC curves, which quantify the separation of the $t\bar{t}H$ signal from the sum of $t\bar{t}W$, $t\bar{t}Z$ and $t\bar{t}$+jets backgrounds, in the $1\ell + 2\tau_h$ (top left), $2\ell ss + 1\tau_h$ (top right), $2\ell + 2\tau_h$ (bottom left), and $3\ell + 1\tau_h$ (bottom right) categories.

11.2 Multivariate analysis for categories without taus

The signal extraction is performed with the same methods as used for Ref. [9]. The variables used as inputs to the event-level BDTs for the $2\ell ss$ and 3ℓ categories are given in Table 17. Due to the low event statistics, the signal extraction in the 4ℓ category is based on simple event counting. No shape analysis is performed in the 4ℓ category.

12 Results

The number of events selected in different categories and in the control regions included on the fit is compared to the contributions of the $t\bar{t}H$ signal and of different background processes in Tables 18 to 22. The sum of background contributions from Z+jets, W+jets, and diboson

Table 16: Variables used in the multivariate discriminators for the channels with taus. Only one multivariate discriminant is used in the categories with τ_h. The $cos(\theta)^*$ between leading and trailing τ_hs stands for the cosinus between taus in the rest frame of the τ_h pair. Masses of ℓs + leading τ_h, in the $3\ell + 1\tau_h$ category, are constructed based on the leptons that have opposite signs to τ_h.

Category	$1\ell+$ $2\tau_h$	2ℓss $1\tau_h$	2ℓ $+2\tau_h$	$3\ell+$ $1\tau_h$	name on Fig. 6
Leading ℓ cone p_T	X		X	X	lep1_conePt
Trailing ℓ cone p_T		X		X	lep2_conePt
Min. ΔR between leading ℓ and a jet	X	X	X	X	mindr_tau1_jet
Min. ΔR between trailing ℓ and a jet		X			mindr_tau2_jet
ΔR between leading and trailing ℓ		X		X	dr_leps
Transverse Mass of leading ℓ	X	X			mT_lep1 or mT_lep
Transverse Mass of trailing ℓ		X			mT_lep2
Maximum $\|\eta\|$ of ℓ collection		X		X	max_lep_eta
Signal leading $\ell\times$ signal trailing ℓ			X		is_OS
Average of ΔR between the jets	X	X	X		avg_dr_jet
Number of jets		X		X	nJet
Number of loose b-jets	X		X		nBJetLoose
Mass of leading medium b-jet pair		X			mbb
Mass of leading loose b-jet pair				X	mbb_loose
E_T^{miss}	X	X		X	ptmiss
res-hTT	X	X			mvaOutput_hadTopTaggerWithKinFit
Hadronic t p_T	X	X			HadTop_pt or unfittedHadTop_pt
Leading τ_h p_T	X	X	X	X	tau1_pt or tau_pt
Trailing τ_h p_T	X		X		tau2_pt
Mass of leading τ_h + trailing τ_h	X		X		mTauTauVis
ΔR between leading and trailing τ_h	X		X		dr_taus
$cos(\theta)^*$ between leading and trailing τ_h	X		X		costS_tau
Min. ΔR between leading τ_h and a jet	X	X	X	X	mindr_lep1_jet or mindr_tau_jet
Min. ΔR between training τ_h and a jet	X				mindr_lep2_jet
Mass of leading ℓ + leading τ_h				X	mTauTauVis1
Mass of trailing ℓ + leading τ_h		X		X	mTauTauVis2
ΔR between leading ℓ and leading τ_h	X	X			dr_lep1_tau
ΔR between trailing ℓ and leading τ_h		X			dr_lep2_tau
$\Delta R(\ell, \tau_h)$ for same-sign pair of (ℓ, τ_h)	X				dr_lep_tau_ss
Average of ΔR between ℓ, and τ_h			X		avr_dr_lep_tau
Number of variables	17	18	13	12	—

production is referred to as "electroweak" background, while the contribution of triboson production and of the few other "exotic" processes described in Section 2 is referred to as "other" background. Background contributions arising from single top and top quark pair production are fully covered by the estimates for the fake and charge flip backgrounds, obtained from data, and are not given separately in the table. The event yields are computed for the values of the $t\bar{t}H$ signal rate and of the nuisance parameters obtained from the ML fit described in Section 10. The event yields observed in data are in agreement with the sum of signal plus background contributions.

The distributions in the discriminating observables that are used for the signal extraction are

Table 17: Variables used in the multivariate discriminators for the channels without taus. As explained in the text, for the categories without taus two different multivariate discriminants are constructed, each one discriminating against a different type of the background (this is added as a comment).

Category	2ℓss		3ℓ	
Comment	$t\bar{t}$	$t\bar{t}V$	$t\bar{t}$	$t\bar{t}V$
Leading ℓ cone p_T		X		X
Trailing ℓ cone p_T		X		X
Minimum of ΔR(leading ℓ, j)	X	X	X	X
Minimum of ΔR(trailing ℓ, j)	X	X		
Transverse Mass of leading ℓ	X	X	X	X
Transverse Mass of trailing ℓ			X	X
Maximum $\lvert\eta\rvert$ of ℓ collection	X	X	X	X
Number of jets	X	X		
res-hTT				
Hj-tagger				
Number of variables				

shown in Figs. 26 to 8. The distributions observed in data are in agreement with the sum of signal plus background contributions.

The $t\bar{t}H$ signal rate μ is extracted from a fit to all seven categories individually and from all categories combined. The rates of the $t\bar{t}W$, $t\bar{t}WW$, and $t\bar{t}Z$ backgrounds is left freely floating in each of these fits, and the control regions described in Sections I.1 and I.2 of the Appendix are included in each fit to constrain these backgrounds. The rates of the $t\bar{t}W$ and $t\bar{t}WW$ backgrounds are assumed to scale by the same factor, to which we refer to as $\mu_{t\bar{t}W}$. The scale factor for the $t\bar{t}Z$ background is denoted by $\mu_{t\bar{t}Z}$. The results are shown in Fig. **??**. For the combined fit, the observed (expected) signal rate is $\mu = X.XX^{+Y.YY}_{-Z.ZZ}$ ($1.00^{+0.38}_{-0.34}$) times the SM $t\bar{t}H$ production rate, which corresponds to an observed (expected) local significance of $QQQ\sigma$ (3.1σ). Such a deviation from the background-only hypothesis presents the evidence for the $t\bar{t}H$ production in the considered final states. While the 2ℓss, 3ℓ and 4ℓ categories are mostly sensitive to the $t\bar{t}H$ signal in the $H \to WW$ and $H \to ZZ$ decay modes, the $1\ell + 2\tau_h$, $2\ell\text{ss} + 1\tau_h$, $2\ell + 2\tau_h$, and $3\ell + 1\tau_h$ categories enhance the sensitvity to the $H \to \tau\tau$ decay mode.

The rates of the $t\bar{t}W + t\bar{t}WW$ and of the $t\bar{t}Z$ backgrounds are given in Table 24. The measured rates are in agreement with the SM expectation within their uncertainties.

A combination of the results obtained using the 2017 data with those using the 2016 (analysis) data, published in Ref. [9], yields an observed (expected) signal strength of $\mu = X.XX^{+Y.YY}_{-Z.ZZ}$ ($1.00^{+0.31}_{-0.27}$), which corresponds to an observed (expected) local significance of $QQQ\sigma$ (3.95σ). Such a deviation from the background-only hypothesis hints about the presence of a $t\bar{t}H$ signal. The 2016 and 2017 analyses are combined at datacard level.

Table 18: The selected number of events in the $2\ell ss$ (sub)categories. The rates are adjusted using the maximum-likelihood fit (post-fit). Smaller uncertainties than 0.01 are denoted as 0.00.

Category	$2\ell ss$									
b-tag	no req.		Loose				Tight			
Leptons	*ee*		*em*		*mm*		*em*		*mm*	
Charge	−	+	−	+	−	+	−	+	−	+
$t\bar{t}H$, $H \to ZZ$	0.15 ± 0.09	0.12 ± 0.07	0.22 ± 0.13	0.16 ± 0.09	0.11 ± 0.06	0.13 ± 0.08	0.22 ± 0.13	0.14 ± 0.09	0.07 ± 0.04	0.07 ± 0.04
$t\bar{t}H$, $H \to WW$	2.52 ± 1.45	2.68 ± 1.75	5.27 ± 3.11	4.90 ± 2.95	3.48 ± 2.15	3.62 ± 2.12	3.54 ± 2.17	3.89 ± 2.50	2.59 ± 1.62	2.29 ± 1.33
$t\bar{t}H$, $H \to \tau\tau$	0.60 ± 0.35	0.58 ± 0.34	0.93 ± 0.54	1.01 ± 0.59	0.57 ± 0.33	0.65 ± 0.39	0.76 ± 0.43	0.65 ± 0.40	0.46 ± 0.29	0.39 ± 0.24
$t\bar{t}H$, $H \to \mu\mu$	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05
$t\bar{t}H$, $H \to Z\gamma$	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05
ttH (sum)	3.27 ± 1.49	3.41 ± 1.79	6.46 ± 3.16	6.09 ± 3.01	4.17 ± 2.18	4.43 ± 2.15	4.56 ± 2.22	4.70 ± 2.54	3.14 ± 1.65	2.76 ± 1.35
$t\bar{t}W$	12.64 ± 2.36	20.06 ± 3.86	22.19 ± 4.49	39.43 ± 7.92	14.00 ± 2.61	24.95 ± 5.04	15.07 ± 2.76	30.39 ± 5.31	11.00 ± 1.99	17.97 ± 3.41
$t\bar{t}WW$	0.93 ± 0.23	0.70 ± 0.17	1.86 ± 0.45	1.36 ± 0.33	0.92 ± 0.23	0.95 ± 0.25	1.40 ± 0.35	1.14 ± 0.27	0.71 ± 0.17	0.77 ± 0.19
ttW + ttWW	13.58 ± 2.37	20.75 ± 3.87	24.06 ± 4.52	40.80 ± 7.93	14.92 ± 2.62	25.90 ± 5.04	16.47 ± 2.78	31.53 ± 5.32	11.71 ± 2.00	18.74 ± 3.42
tHq $H \to \tau\tau$	< 0.05	0.08 ± 0.02	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05
tHW $H \to \tau\tau$	< 0.05	< 0.05	< 0.05	0.12 ± 0.02	< 0.05	0.14 ± 0.03	< 0.05	< 0.05	< 0.05	< 0.05
tHq $H \to WW$	0.06 ± 0.01	< 0.05	0.10 ± 0.02	0.14 ± 0.01	0.07 ± 0.01	0.14 ± 0.01	0.10 ± 0.01	< 0.05	< 0.05	< 0.05
tHW $H \to WW$	< 0.05	< 0.05	0.19 ± 0.04	0.12 ± 0.01	0.24 ± 0.07	0.14 ± 0.02	< 0.05	0.22 ± 0.04	< 0.05	0.11 ± 0.02
tHq $H \to ZZ$	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05
tHW $H \to ZZ$	< 0.05	0.10 ± 0.02	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05
tH (sum)	0.11 ± 0.02	0.24 ± 0.02	0.30 ± 0.04	0.41 ± 0.03	0.31 ± 0.07	0.46 ± 0.04	0.16 ± 0.01	0.23 ± 0.04	0.05 ± 0.00	0.11 ± 0.02
ZZ	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05
WZ	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05
WZ + ZZ	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05
$t\bar{t}Z$	17.86 ± 2.86	17.62 ± 2.71	20.03 ± 3.10	23.08 ± 3.77	7.55 ± 1.22	8.68 ± 1.42	14.79 ± 2.43	16.97 ± 2.67	5.91 ± 0.99	5.74 ± 0.93
Misidentified	7.89 ± 2.05	10.44 ± 2.49	26.61 ± 4.30	28.27 ± 4.80	20.60 ± 4.40	18.23 ± 3.92	5.64 ± 1.87	6.60 ± 1.87	3.67 ± 1.59	4.19 ± 1.75
Conversions	1.52 ± 0.46	2.18 ± 1.83	2.52 ± 0.87	3.48 ± 2.13	< 0.05	< 0.05	0.88 ± 0.29	1.05 ± 0.32	< 0.05	< 0.05
signal flip	4.53 ± 1.10	4.59 ± 1.14	3.71 ± 0.90	3.79 ± 0.99	< 0.05	< 0.05	3.10 ± 0.75	3.07 ± 0.76	< 0.05	< 0.05
Other	1.44 ± 0.61	2.45 ± 1.06	4.62 ± 1.99	5.69 ± 2.22	1.65 ± 0.64	3.33 ± 1.29	2.18 ± 0.91	2.39 ± 1.01	1.11 ± 0.45	1.80 ± 0.81
SM expectation	50.20 ± 4.69	61.68 ± 6.12	88.32 ± 8.00	111.60 ± 10.94	49.21 ± 5.73	61.03 ± 7.01	47.78 ± 4.85	66.54 ± 6.86	25.60 ± 3.22	33.35 ± 4.26
Observed data	54	67	86	108	41	79	50	70	32	27

Table 19: The selected number of events in the 3ℓ and 4ℓ categories. The rates are adjusted using the maximum-likelihood fit (post-fit). Smaller uncertainties than 0.01 are denoted as 0.00.

Category	3ℓ				4ℓ
b-tag	Loose		Tight		no req.
Total charge	$-$	$+$	$-$	$+$	no req.
$t\bar{t}H$, $H \to ZZ$	0.34 ± 0.21	0.12 ± 0.07	0.14 ± 0.09	0.13 ± 0.08	0.06 ± 0.04
$t\bar{t}H$, $H \to WW$	5.60 ± 3.22	5.14 ± 2.94	2.28 ± 1.35	2.04 ± 1.15	0.46 ± 0.29
$t\bar{t}H$, $H \to \tau\tau$	0.93 ± 0.55	0.40 ± 0.24	0.50 ± 0.30	0.71 ± 0.43	0.13 ± 0.08
$t\bar{t}H$, $H \to \mu\mu$	0.07 ± 0.04	0.18 ± 0.11	0.05 ± 0.03	0.12 ± 0.07	0.05 ± 0.03
$t\bar{t}H$, $H \to Z\gamma$	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05
ttH (sum)	6.94 ± 3.27	5.84 ± 2.95	2.98 ± 1.39	3.03 ± 1.23	0.70 ± 0.30
$t\bar{t}W$	12.12 ± 2.52	20.72 ± 4.15	5.64 ± 1.16	8.55 ± 1.64	< 0.05
$t\bar{t}WW$	1.22 ± 0.30	1.34 ± 0.33	0.76 ± 0.18	0.69 ± 0.16	0.13 ± 0.03
ttW + ttWW	13.34 ± 2.54	22.05 ± 4.17	6.40 ± 1.17	9.24 ± 1.64	0.13 ± 0.03
tHq $H \to \tau\tau$	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05
tHW $H \to \tau\tau$	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05
tHq $H \to WW$	0.08 ± 0.01	0.28 ± 0.03	0.27 ± 0.09	< 0.05	< 0.05
tHW $H \to WW$	0.05 ± 0.01	< 0.05	< 0.05	< 0.05	< 0.05
tHq $H \to ZZ$	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05
tHW $H \to ZZ$	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05
tH (sum)	0.15 ± 0.01	0.36 ± 0.03	0.29 ± 0.09	0.08 ± 0.01	< 0.05
ZZ	0.37 ± 0.07	0.29 ± 0.08	< 0.05	< 0.05	0.15 ± 0.10
WZ	3.67 ± 0.76	7.05 ± 1.46	0.22 ± 0.05	0.44 ± 0.09	< 0.05
WZ + ZZ	4.04 ± 0.76	7.34 ± 1.47	0.22 ± 0.05	0.44 ± 0.09	0.15 ± 0.10
$t\bar{t}Z$	24.86 ± 3.98	23.18 ± 3.82	13.15 ± 2.23	12.93 ± 2.13	3.91 ± 0.62
Misidentified	12.30 ± 2.96	10.27 ± 2.44	1.67 ± 0.59	2.59 ± 1.10	< 0.05
Conversions	1.48 ± 0.52	1.97 ± 0.63	2.35 ± 0.93	0.82 ± 0.29	< 0.05
signal flip	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05
Other	2.40 ± 1.09	3.86 ± 1.62	1.57 ± 0.63	1.83 ± 0.69	0.09 ± 0.05
SM expectation	65.52 ± 6.62	74.87 ± 7.19	28.61 ± 3.15	30.96 ± 3.24	5.01 ± 0.70
Observed data	55	84	23	33	6

Table 20: The selected number of events in the $1\ell + 2\tau_h$, $2\ell ss + 1\tau_h$, $2\ell + 2\tau_h$, and $3\ell + 1\tau_h$ categories. The rates are adjusted using the maximum-likelihood fit (post-fit). Smaller uncertainties than 0.01 are denoted as 0.00.

Category	$1\ell + 2\tau_h$	$2\ell + 2\tau_h$	$3\ell + 1\tau_h$	$2\ell ss + 1\tau_h$
$t\bar{t}H$, $H \to ZZ$	0.08 ± 0.05	< 0.05	0.03 ± 0.02	0.06 ± 0.04
$t\bar{t}H$, $H \to WW$	0.34 ± 0.20	0.07 ± 0.04	0.39 ± 0.24	1.73 ± 0.98
$t\bar{t}H$, $H \to \tau\tau$	6.15 ± 3.57	0.80 ± 0.47	0.54 ± 0.30	3.33 ± 1.87
$t\bar{t}H$, $H \to \mu\mu$	< 0.05	< 0.05	< 0.05	< 0.05
$t\bar{t}H$, $H \to Z\gamma$	< 0.05	< 0.05	< 0.05	< 0.05
ttH (sum)	6.57 ± 3.58	0.88 ± 0.47	0.99 ± 0.38	5.13 ± 2.11
$t\bar{t}W$	1.03 ± 0.25	< 0.05	0.37 ± 0.12	12.39 ± 2.43
$t\bar{t}WW$	0.10 ± 0.03	< 0.05	0.14 ± 0.03	0.71 ± 0.16
ttW + ttWW	1.13 ± 0.25	< 0.05	0.51 ± 0.12	13.11 ± 2.43
tHq	< 0.05	< 0.05	< 0.05	0.32 ± 0.03
tHW	0.22 ± 0.07	0.02 ± 0.09	< 0.05	0.13 ± 0.01
tH (sum)	0.25 ± 0.07	0.02 ± 0.09	< 0.05	0.45 ± 0.03
$WZ + ZZ$	1.50 ± 0.77	0.04 ± 0.03	0.13 ± 0.03	2.76 ± 1.95
$t\bar{t}Z$	11.55 ± 2.64	1.55 ± 0.46	4.48 ± 0.68	15.35 ± 2.43
Misidentified	299.60 ± 19.10	5.31 ± 2.19	0.25 ± 0.26	5.28 ± 2.18
Conversions	0.25 ± 0.11	< 0.05	< 0.05	< 0.05
signal flip	< 0.05	< 0.05	< 0.05	< 0.05
Other	1.18 ± 0.46	0.06 ± 0.04	0.33 ± 0.15	3.23 ± 1.14
SM expectation	322.02 ± 19.64	7.86 ± 2.29	6.70 ± 0.85	45.31 ± 5.11
Observed data	324	7	4	53

Table 21: The selected number of events in the $2\ell ss$ (3j) control regions ($t\bar{t}W$ enrich.). The rates are adjusted using the maximum-likelihood fit (post-fit). Smaller uncertainties than 0.01 are denoted as 0.00.

Category	$2\ell ss$									
b-tag	no req.		Loose				Tight			
Leptons	ee		$e\mu$		$\mu\mu$		$e\mu$		$\mu\mu$	
Charge	$-$	$+$	$-$	$+$	$-$	$+$	$-$	$+$	$-$	$+$
$t\bar{t}H$, $H \to ZZ$	< 0.05	< 0.05	0.03 ± 0.02	< 0.05	0.04 ± 0.02	< 0.05	< 0.05	0.04 ± 0.02	< 0.05	< 0.05
$t\bar{t}H$, $H \to WW$	0.55 ± 0.34	0.59 ± 0.38	1.46 ± 0.93	1.48 ± 0.85	1.03 ± 0.61	0.99 ± 0.56	0.62 ± 0.39	0.70 ± 0.45	0.43 ± 0.26	0.37 ± 0.21
$t\bar{t}H$, $H \to \tau\tau$	0.17 ± 0.11	0.17 ± 0.10	0.35 ± 0.21	0.33 ± 0.18	0.20 ± 0.12	0.22 ± 0.12	0.19 ± 0.11	0.11 ± 0.07	0.14 ± 0.09	0.15 ± 0.09
$t\bar{t}H$, $H \to \mu\mu$	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05
$t\bar{t}H$, $H \to Z\gamma$	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05
ttH (sum)	0.75 ± 0.36	0.78 ± 0.39	1.87 ± 0.96	1.85 ± 0.87	1.27 ± 0.62	1.25 ± 0.58	0.84 ± 0.40	0.85 ± 0.46	0.59 ± 0.28	0.53 ± 0.23
$t\bar{t}W$	5.99 ± 1.17	10.70 ± 2.06	12.33 ± 2.39	23.16 ± 4.52	8.15 ± 1.70	16.56 ± 3.16	6.33 ± 1.15	11.73 ± 2.04	4.09 ± 0.85	7.74 ± 1.47
$t\bar{t}WW$	0.05 ± 0.01	0.23 ± 0.06	0.39 ± 0.11	0.24 ± 0.06	0.27 ± 0.08	0.28 ± 0.07	0.10 ± 0.02	0.14 ± 0.03	0.08 ± 0.03	0.06 ± 0.01
ttW + ttWW	6.04 ± 1.17	10.93 ± 2.06	12.72 ± 2.39	23.40 ± 4.52	8.42 ± 1.71	16.84 ± 3.16	6.43 ± 1.15	11.87 ± 2.05	4.17 ± 0.85	7.80 ± 1.47
tHq $H \to \tau\tau$	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05
tHW $H \to \tau\tau$	< 0.05	< 0.05	< 0.05	0.18 ± 0.03	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05
tHq $H \to WW$	< 0.05	< 0.05	0.14 ± 0.08	0.16 ± 0.03	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05
tHW $H \to WW$	< 0.05	< 0.05	0.37 ± 0.07	0.57 ± 0.07	< 0.05	0.16 ± 0.04	< 0.05	0.11 ± 0.01	< 0.05	< 0.05
tHq $H \to ZZ$	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05
tHW $H \to ZZ$	< 0.05	< 0.05	0.06 ± 0.01	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05
tH (sum)	0.05 ± 0.00	0.06 ± 0.01	0.60 ± 0.11	0.93 ± 0.08	0.06 ± 0.01	0.19 ± 0.04	< 0.05	0.11 ± 0.01	< 0.05	< 0.05
ZZ	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05
WZ	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05
WZ + ZZ	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05
$t\bar{t}Z$	8.26 ± 1.55	8.22 ± 1.38	10.13 ± 1.87	11.58 ± 1.95	3.99 ± 0.69	3.84 ± 0.67	6.10 ± 1.34	5.30 ± 0.97	2.39 ± 0.82	2.17 ± 0.38
Misidentified	7.75 ± 2.00	9.32 ± 2.13	26.30 ± 4.30	23.94 ± 4.45	20.50 ± 4.65	17.35 ± 4.25	2.88 ± 1.02	4.86 ± 1.51	2.28 ± 1.40	1.92 ± 0.81
Conversions	1.51 ± 0.58	2.50 ± 1.63	1.42 ± 0.41	1.96 ± 0.84	< 0.05	< 0.05	0.53 ± 0.18	0.57 ± 0.18	< 0.05	< 0.05
signal flip	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05
Other	1.72 ± 0.77	2.23 ± 0.85	3.67 ± 1.65	5.91 ± 2.38	2.52 ± 1.01	3.88 ± 1.53	0.90 ± 0.37	1.19 ± 0.60	0.12 ± 0.08	0.53 ± 0.22
SM expectation	26.09 ± 2.97	34.04 ± 3.77	56.71 ± 5.61	69.55 ± 7.16	36.76 ± 5.14	43.35 ± 5.58	17.71 ± 2.12	24.74 ± 2.83	9.57 ± 1.86	12.99 ± 1.75
Observed data	33	33	64	105	49	41	27	27	15	8

Table 22: The selected number of events in the 3ℓ and 4ℓ control regions. The rates are adjusted using the maximum-likelihood fit (post-fit). Smaller uncertainties than 0.01 are denoted as 0.00.

Category	3ℓ					4ℓ
CR	Z-peak ($t\bar{t}Z$ enrich.)				WZ enrich.	ZZ enrich.
b-tag	Loose		Tight		no req.	
Total charge	$-$	$+$	$-$	$+$	no req.	
Process						
$t\bar{t}H$, $H \to ZZ$	0.61 ± 0.37	0.52 ± 0.32	0.25 ± 0.17	0.23 ± 0.15	0.19 ± 0.19	0.05 ± 0.08
$t\bar{t}H$, $H \to WW$	0.46 ± 0.29	0.75 ± 0.44	0.54 ± 0.32	0.21 ± 0.15	0.38 ± 0.35	< 0.05
$t\bar{t}H$, $H \to \tau\tau$	0.19 ± 0.12	0.13 ± 0.08	0.09 ± 0.06	< 0.05	0.05 ± 0.06	< 0.05
$t\bar{t}H$, $H \to \mu\mu$	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05
$t\bar{t}H$, $H \to Z\gamma$	0.04 ± 0.03	0.08 ± 0.05	< 0.05	< 0.05	< 0.05	< 0.05
ttH (sum)	1.29 ± 0.49	1.49 ± 0.55	0.91 ± 0.37	0.49 ± 0.21	0.64 ± 0.41	0.06 ± 0.08
$t\bar{t}W$	1.66 ± 0.36	2.63 ± 0.53	1.01 ± 0.24	1.78 ± 0.38	0.98 ± 0.92	< 0.05
$t\bar{t}WW$	0.23 ± 0.07	0.16 ± 0.08	0.18 ± 0.08	0.03 ± 0.02	0.07 ± 0.06	< 0.05
ttW + ttWW	1.89 ± 0.36	2.79 ± 0.54	1.20 ± 0.25	1.82 ± 0.38	1.05 ± 0.92	< 0.05
tHq $H \to \tau\tau$	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05
tHW $H \to \tau\tau$	< 0.05	0.17 ± 0.02	< 0.05	< 0.05	< 0.05	< 0.05
tHq $H \to WW$	< 0.05	< 0.05	0.07 ± 0.01	0.34 ± 0.30	< 0.05	< 0.05
tHW $H \to WW$	< 0.05	< 0.05	< 0.05	0.04 ± 0.04	< 0.05	< 0.05
tHq $H \to ZZ$	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05
tHW $H \to ZZ$	< 0.05	< 0.05	< 0.05	< 0.05	0.07 ± 0.06	< 0.05
tH (sum)	0.05 ± 0.02	0.18 ± 0.02	0.10 ± 0.02	0.40 ± 0.31	0.09 ± 0.06	< 0.05
ZZ	4.47 ± 0.90	4.19 ± 0.87	0.66 ± 0.56	0.29 ± 0.09	97.77 ± 116.12	55.58 ± 9.04
WZ	20.69 ± 4.51	30.16 ± 6.95	1.49 ± 0.49	1.12 ± 0.29	694.46 ± 117.54	< 0.05
$WZ + ZZ$	25.16 ± 4.59	34.35 ± 7.01	2.15 ± 0.75	1.40 ± 0.30	792.23 ± 165.23	55.58 ± 9.04
$t\bar{t}Z$	79.37 ± 13.72	73.64 ± 12.16	52.09 ± 8.16	40.28 ± 7.61	45.23 ± 44.09	3.25 ± 4.12
Misidentified	5.60 ± 1.40	6.43 ± 1.45	1.35 ± 0.52	1.12 ± 0.80	18.50 ± 18.00	< 0.05
Conversions	0.08 ± 0.03	0.71 ± 0.23	0.95 ± 0.96	0.06 ± 0.07	4.25 ± 4.25	< 0.05
signal flip	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05
Other	15.98 ± 6.43	24.61 ± 9.63	5.35 ± 2.12	7.12 ± 2.83	56.79 ± 67.99	2.42 ± 2.97
SM expectation	129.43 ± 15.91	144.20 ± 17.10	64.09 ± 8.54	52.68 ± 8.18	918.77 ± 184.96	61.30 ± 10.37
Observed data	123	154	54	49	919	61

Table 23: Signal rates μ, in units of the SM $t\bar{t}H$ production rate, measured in individual categories and for the combination of all seven categories, and for the combination of all categories of the 2017 analysis with the analysis of the 2016 data, published in Ref. [9] ("combined with HIG-17-018").

Category	signal strength $\pm 1\sigma$ Measured	Expected	significance
$1\ell + 2\tau_h$	$X.XX^{+X.XX}_{-X.XX}$	$1.00^{+1.16}_{-0.95}$	$--$
$2\ell ss$	$X.XX^{+X.XX}_{-X.XX}$	$1.00^{+0.52}_{-0.48}$	$--$
$2\ell ss + 1\tau_h$	$X.XX^{+X.XX}_{-X.XX}$	$1.00^{+0.70}_{-0.59}$	$--$
$2\ell + 2\tau_h$	$X.XX^{+X.XX}_{-X.XX}$	$1.00^{+2.52}_{-1.73}$	$--$
3ℓ	$X.XX^{+X.XX}_{-X.XX}$	$1.00^{+0.68}_{-0.60}$	$--$
$3\ell + 1\tau_h$	$X.XX^{+X.XX}_{-X.XX}$	$1.00^{+1.57}_{-1.13}$	$--$
4ℓ	$X.XX^{+X.XX}_{-X.XX}$	$1.00^{+2.40}_{-1.67}$	$--$
Combined	$X.XX^{+X.XX}_{-X.XX}$	$1.00^{+0.38}_{-0.34}$	3.10σ
Combined with HIG-17-018	$X.XX^{+X.XX}_{-X.XX}$	$1.00^{+0.31}_{-0.27}$	3.95σ

Table 24: Signal rates μ, in units of the SM $t\bar{t}H$ and $t\bar{t}Z$ and $t\bar{t}W + t\bar{t}WW$ production rates, measured including the $t\bar{t}Z$ and $t\bar{t}W(W)$ control regions on the fit. If the result is redefined to $\mu_{t\bar{t}H}$ only the significance is 3.10σ.

	signal strength $\pm 1\sigma$	
	Measured	Expected
$\mu_{t\bar{t}H}$	$X.XX^{+X.XX}_{-X.XX}$	$1.00^{+0.38}_{-0.34}$
$\mu_{t\bar{t}W}$	$X.XX^{+X.XX}_{-X.XX}$	$1.00^{+0.26}_{-0.22}$
$\mu_{t\bar{t}Z}$	$X.XX^{+X.XX}_{-X.XX}$	$1.00^{+0.22}_{-0.19}$

Table 25: 95% CL upper limits on the $t\bar{t}H$ signal rate μ, obtained in individual categories and for the combination of all seven categories, and for the combination of all categories of the 2017 analysis with the analysis of the 2016 data, published in Ref. [9] ("combined with HIG-17-018"). The observed limit is compared to the limits expected for the background-only hypothesis ($\mu = 0$) and to the SM expectation ($\mu = 1$).

Category	Observed limit	Expected limit ($\mu = 0$)	Expected limit ($\mu = 1$)
$1\ell + 2\tau_h$	$X.XX$	$X.XX$	$X.XX$
$2\ell ss$	$X.XX$	$X.XX$	$X.XX$
$2\ell ss + 1\tau_h$	$X.XX$	$X.XX$	$X.XX$
$2\ell + 2\tau_h$	$X.XX$	$X.XX$	$X.XX$
3ℓ	$X.XX$	$X.XX$	$X.XX$
$3\ell + 1\tau_h$	$X.XX$	$X.XX$	$X.XX$
4ℓ	$X.XX$	$X.XX$	$X.XX$
Combined	$X.XX$	$X.XX$	$X.XX$
Combined with HIG-17-018	$X.XX$	$X.XX$	$X.XX$

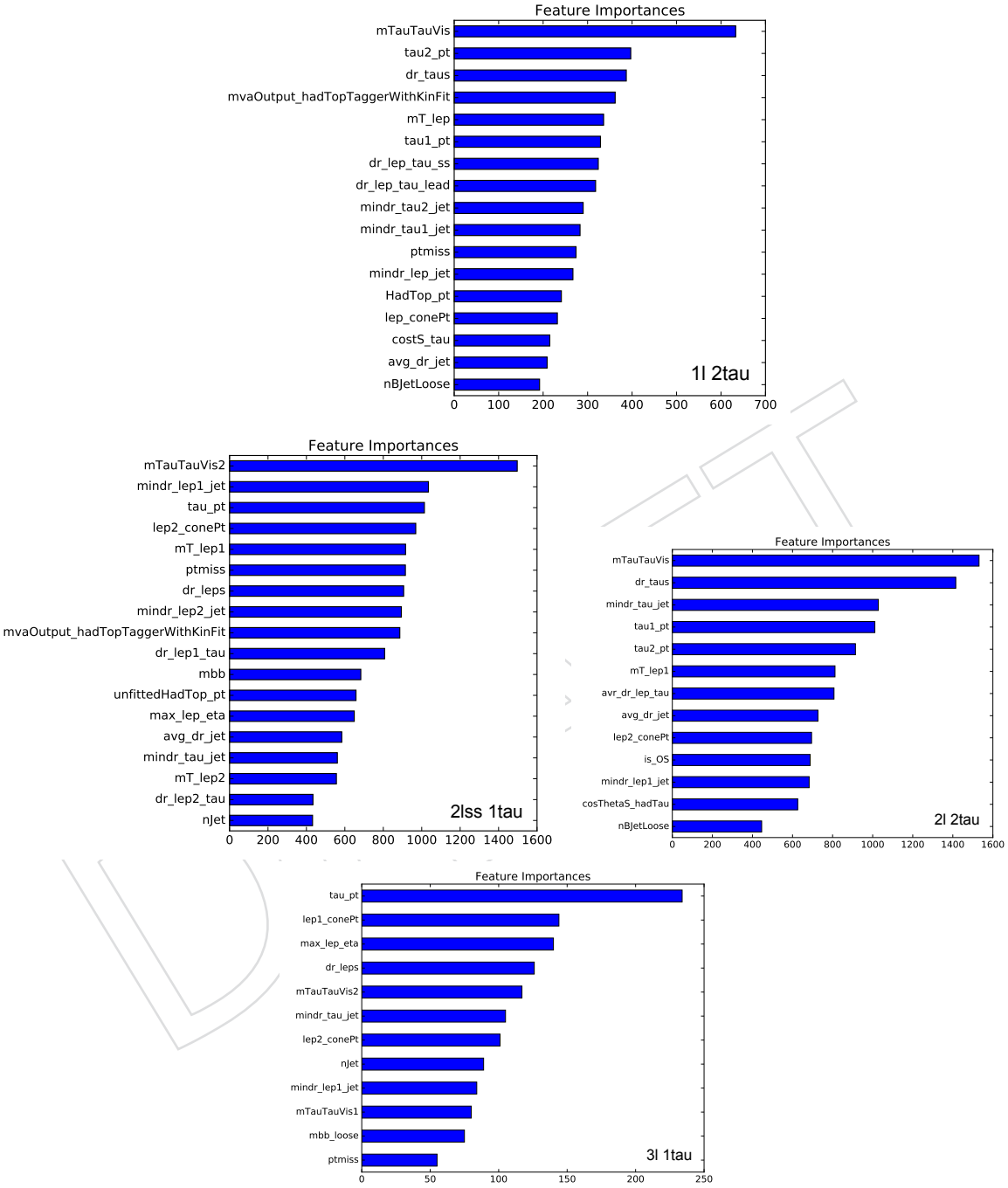
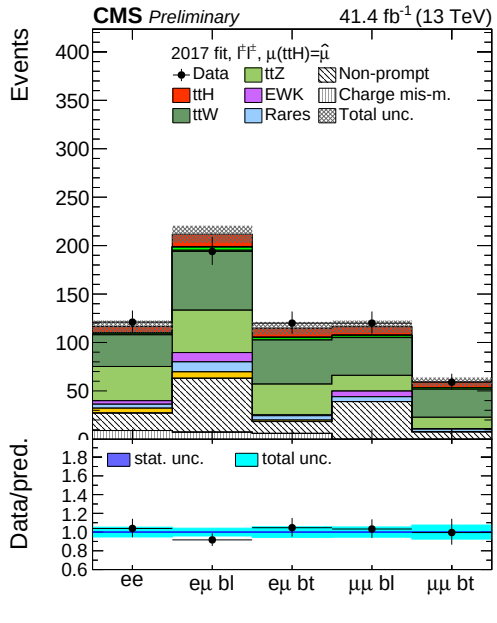
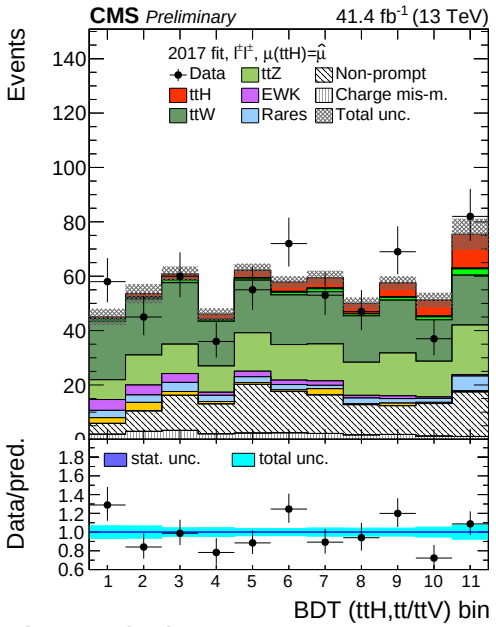


Figure 6: Ranking in performance of individual BDT input variables to separate the $t\bar{t}H$ signal from backgrounds in the $1\ell + 2\tau_h$ (top left), $2\ell ss + 1\tau_h$ (top right), $2\ell + 2\tau_h$ (bottom left), and $3\ell + 1\tau_h$ (bottom right) categories.

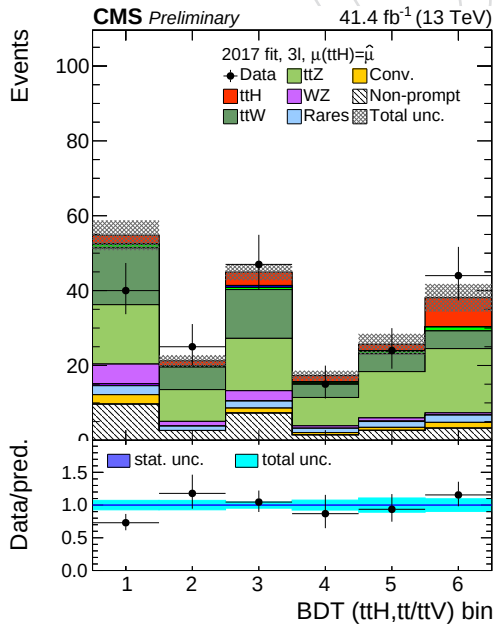
Figure 7: Distributions in the discriminating observables used for the signal extraction in the categories without taus. The SM expectations for the $t\bar{t}H$ signal and for the background processes are presented in colors. The post-fit rates and uncertainties are used. Figures for subcategories are found in appendix O.



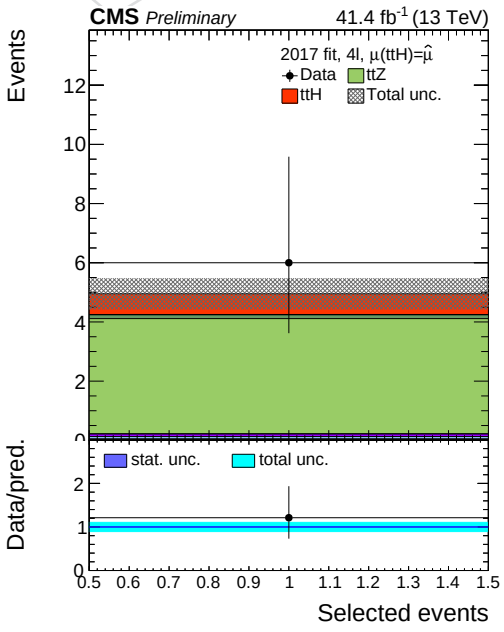
(a)



(b)

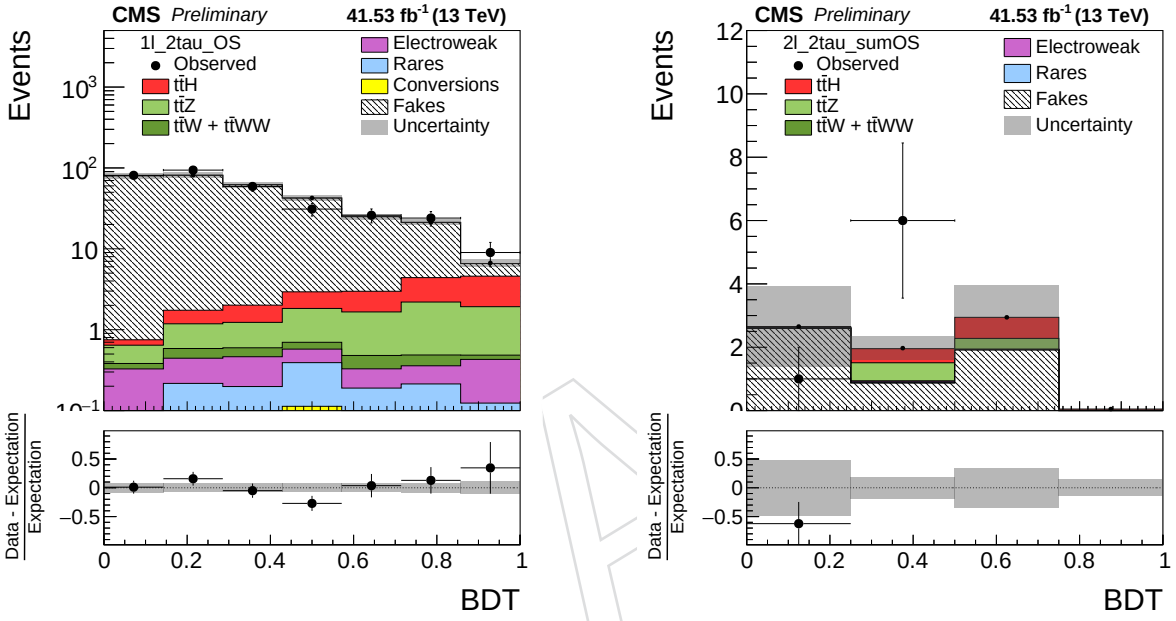


(c)



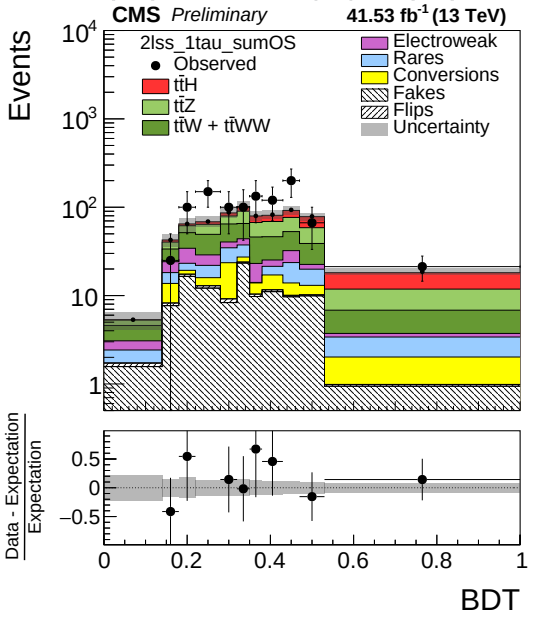
(d)

Figure 8: Distributions in the discriminating observables used for the signal extraction in (a) the $1\ell + 2\tau_h$, (b) the $2\ell + 2\tau_h$, (c) the $2\ell ss + 1\tau_h$, and (d) $3\ell + 1\tau_h$ categories. The SM expectations for the $t\bar{t}H$ signal and for the background processes are presented in colors. The post-fit rates and uncertainties are used.

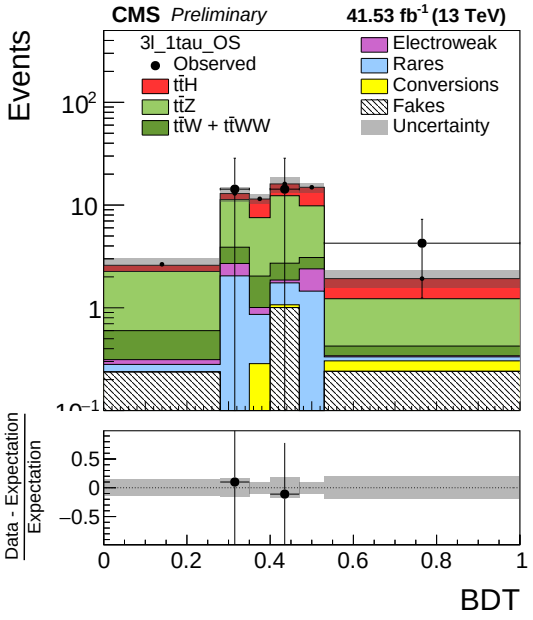


(a)

(b)



(c)



(d)

A Differences between 2016 and 2017 analyses

The search for $t\bar{t}H$ production in final states with leptons and τ_h performed with the data recorded in 2017 has been updated compared to the corresponding analysis of the 2016 data in four areas: the sensitivity of the analysis has been increased by new and/or improved analysis techniques, the treatment of the irreducible $t\bar{t}W$ and $t\bar{t}Z$ backgrounds has been revised, improved algorithms for particle reconstruction and identification are used and the event-level BDTs have been retrained, and the data-to-Monte Carlo corrections have been remeasured and updated in the analysis.

The sensitivity of the analysis has been improved by the following developments:

- The addition of the $2\ell + 2\tau_h$ category.
- The improvement of the signal extraction in the categories with τ_h. In the 2017 analysis, a single BDT has been trained with the `xgboost` algorithm [100] to separate the $t\bar{t}H$ from the sum of the reducible $t\bar{t}$+jets background as well as from the irreducible $t\bar{t}W$ and $t\bar{t}Z$ backgrounds. Compared to training two separate BDTs for the separation of the $t\bar{t}H$ signal from reducible and irreducible backgrounds with the TMVA package [7] and mapping the output of the two BDTs into a single discriminant, as described in Appendix D of Ref. [76], the new signal extraction improves the sensitivity of the categories with τ_h by 10-15% in the 2017 analysis.
- The application of the resHTT algorithm, described in Section C of the Appendix, to reconstruct had-top decays in the $1\ell + 2\tau_h$ and $2\ell ss + 1\tau_h$ categories. In the 2016 analysis, the reconstruction of had-top decays was used in the $2\ell ss$ category only.

In the 2016 analysis, the event yield of the irreducible $t\bar{t}W$ and $t\bar{t}Z$ backgrounds in the SR have been obtained from the MC simulation, using the SM expectation for the $t\bar{t}W$ and $t\bar{t}Z$ cross sections, whereas the yields of these backgrounds are fully determined from data in the 2017 analysis. The latter is accomplished by introducing two unconstrained nuisance parameters for the $t\bar{t}W$ and $t\bar{t}Z$ cross sections into the ML fit described in Section 10 and adding the CR described in Sections I.1 and I.2 to the fit, in order to constrain these backgrounds. The yield of the (small) $t\bar{t}WW$ background is assumed to scale by the same factor with respect to its SM expectation as the $t\bar{t}W$ background. The update in the treatment of the irreducible $t\bar{t}W$ and $t\bar{t}Z$ backgrounds does not improve the sensitivity, but addresses concerns that the $t\bar{t}W$ and $t\bar{t}Z$ cross sections measured by CMS both exceed the SM expectation by about one standard deviation [106], and the normalization of the $t\bar{t}W$ and $t\bar{t}Z$ backgrounds to the SM cross section may cause a bias of the $t\bar{t}H$ signal rate, as the event-level BDTs described in Section 11, which are used for the signal extraction, provide only limited discrimination between the $t\bar{t}H$ signal and the $t\bar{t}W$ and $t\bar{t}Z$ backgrounds.

The algorithms used for particle identification have been updated:

- A new training of the BDT that separates prompt from non-prompt leptons, described in Section 3.1.4, has been performed. The definition of the observable `jetPtRatio` has been improved. In case no jet of $p_T > 15$ GeV is within a distance $\Delta R < 0.4$ to the lepton, the observable is set to $p_T^\ell/(p_T^\ell + I_\ell)$ in the 2017 analysis, while the observable was set to a constant value in the 2016 analysis in this case.
- The new training `2017v2`, provided by the Tau POG [58], is used to identify τ_h in the 2017 analysis. The choice of the WP has been reoptimized. The reoptimization of the choice of WP turned-out to be important, as the WP have shifted between the `2017v2` training and the `2015` training [107], which was used to identify τ_h in the

2016 analysis.

- The switch from the CSVv2 to the DeepCSV algorithm for the identification of jets originating from the hadronization of b quarks.
- A new training of the Hj tagger, described in Section D of the Appendix, that is used in the $2\ell ss$ category.

The updates are related to changes in the detector configuration and in the CMS reconstruction software (CMSSW). The event-level BDTs used for the signal extraction have also been retrained.

The following data-to-MC corrections have been remeasured for the 2017 analysis:

- The corrections for trigger efficiencies, described in Section 7.2.
- The corrections for electron and muon reconstruction and isolation efficiencies, described in Section 7.3.
- The corrections for data-to-MC differences in b-tagging efficiencies and mistag rates, which are obtained from the BTV POG and described in Section [79].

The MEM-based discriminants used in the 3ℓ and $2\ell ss + 1\tau_{\mathrm{h}}$ categories in the 2016 analysis are currently being updated to the 2017 data-taking conditions and are not yet used in the 2017 analysis.

B Combination with 2016 analysis

In this section, we highlight the components of systematic uncertainties that are considered correlated in the combination of this analysis (2017 data) with the one described in PAS-HIG-17-18 (2016 data). We follow the nomenclature introduced in Sec. 9.

- b**-tag efficiency and mistag rate**
  In the 2016 version of the analysis, the CSVv2 was used, while in the 2017 version DeepCSV is the new default b-tagger (described in this analysis note). Even though we consider the b-tag uncertainties that are not related with statistics ("HF", "LF", "cErr1", "cErr2") to be correlated between the analysis eras.
- **Identification and isolation efficiency for** e **and** μ
- **Jet energy scale**
- **Fake background estimation**
- **Normalization of charge flip background**
  This component is considered as correlated as it is related to the physics of the background events. The stat. uncertainty on the flip rate would be uncorrelated, but it's expected to be small.
- **Rare background**
- **Diboson background**
- **Theoretical sources of systematics**
  That comprises sources coming from PDF and α_s (both normalization and shape), and as well the ones related with H branching fractions.
- **Fake background estimation (categories with at least one** τ_h**)**
- **Identification efficiency for** τ_h
- **Energy scale of** τ_h

C Reconstruction of hadronic top quark decays

An algorithm has been developed to identify hadronic top quark ("had-top") decays. The algorithm, referred to as "resHTT", targets the case that the b quark and the two quarks produced in the W boson decay are reconstructed as separate jets. It computes the likelihood of three jets passing the selection criteria described in Section 3.3 and not overlapping, within $\Delta R < 0.4$, with an electron, muon, or τ_h passing the fakeable object selection criteria to be compatible with originating from a had-top decay. We refer to the combination of three such jets as "jet triplet". The algorithm uses a BDT to discriminate triplets of jets in which all jets are generator-level matched to quarks from a had-top decay (the "signal") from jet triplets in which one or more jets fail to satisfy this matching criterion (the "background").

Observables reconstructed by a kinematic fit of the three jets are used as input to the BDT. The kinematic fit maximizes the likelihood function:

$$L = W(p_T^b|\hat{p}_T^b)\, W(p_T^{W1}|\hat{p}_T^{W1})\, W(p_T^{W2}|\hat{p}_T^{W2})\,, \tag{9}$$

where the symbol b refers to the jet corresponding to the b quark, and the symbols W1 and W2 refer to the quarks of higher and lower p_T that are produced in the decay of the W boson, respectively. The symbols with (without) a hat indicate the true (reconstructed) p_T. The maximization of the likelihood function L is performed subject to the constraint that the mass of W1+W2 equals $m_W = 80.4$ GeV and the mass of b + W1 + W2 equals $m_t = 173.1$ GeV [52]. The functions $W(p_T|\hat{p}_T)$, referred to as "transfer functions" in the literature [108], are taken from Ref. [109], with updates to the LHC Run 2 data-taking conditions taken from Ref. [110].

The following variables are used as input to the BDT:

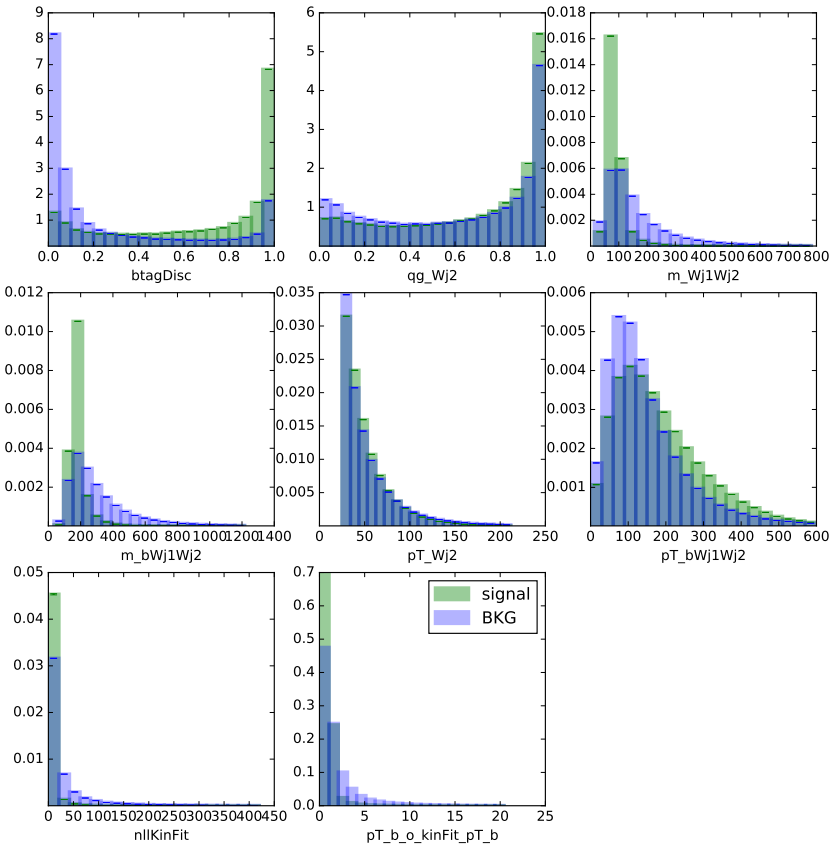
- `btagDisc`, the b-tag discriminant of b.
- `qj_Wj2`, quark-gluon discriminant [111] of jet with the lower p_T between W1 of W2.
- `m_Wj1Wj2`, mass of W1 + W2.
- `m_bWj1Wj2`, mass of b + W1 + W2.
- `pT_Wj2`, p_T of the jet with the lower p_T between W1 of W2.
- `pT_bWj1Wj2`, p_T of b + W1 + W2.
- `nllkinfit`, the likelihood function described on Eq. 9
- `pT_b_o_kinFit_pT_b`, ratio between the p_T of b before and after the kinematic fit.

The distributions of these variables for jet triplets considered as signal and background are shown in Fig. 9.

The BDT is trained using a mixture of $t\bar{t}H$, $t\bar{t}W$, $t\bar{t}Z$, and $t\bar{t}$+jets samples, produced by the MC simulation. No further event selection criteria are applied to the events used for the BDT training, the idea being that the resHTT algorithm is trained once for all categories in which $t\bar{t}H$ signal events are expected to contain top quarks that decay hadronically ($1\ell + 2\tau_h$, $2\ell ss$, $2\ell ss + 1\tau_h$, and 3ℓ).

The performance of the algorithm is quantified by the fraction of $t\bar{t}H$ signal events in which the jet triplet with the highest BDT output corresponds to a genuine hadronic top decay, i.e. in which all three jets of this jet triplet match either the b quark or one of the two quarks produced in the W boson decay on generator level on a cone of 0.2. The fraction of $t\bar{t}H$ signal events in which all three jets produced in hadronic top quark decays satisfy the conditions $p_T > 25$ GeV, $|\eta| < 2.4$ (and therefore could be tagged) is $\approx 25\%$ of the events that pass the selections of the selections of the $1\ell + 2\tau_h$ and $2\ell ss + 1\tau_h$ categories, which limits the usefulness of the resHTT

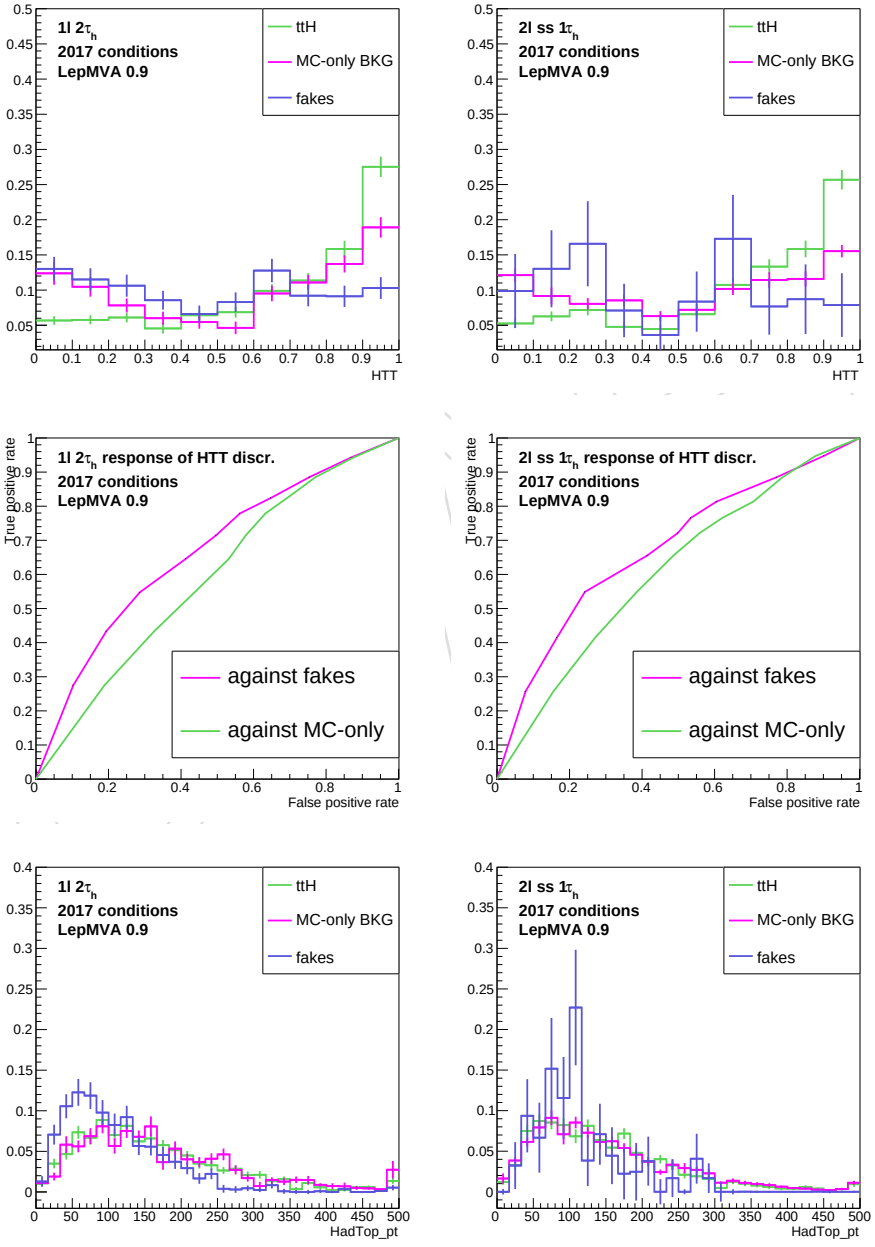
Figure 9: Distribution of BDT input variables used for reconstructing had-top decays for signal (green) and background (blue) jet triplets in simulated $t\bar{t}$+jets events produced with the full CMS detector simulation for 2017 detector conditions.



algorithm as inputs to the event-level BDTs described in Section 11. The fraction of such $\approx 25\%$ portion of events which contain at least one generator-level top quark (three sub-jets matched with 0.2 matching cone) amounts to $\approx 70\%$ in $t\bar{t}H$ signal events and $\approx 60\%$ in $t\bar{t}W$, $t\bar{t}W$ and top quark pair production .

Distributions in BDT output and in the p_T of the had-top candidate reconstructed by the resHTT algorithm after applying the selections of the $1\ell + 2\tau_h$ and $2\ell ss + 1\tau_h$ categories are shown in Fig. 10, where the had-top candidate reconstructed in a given event is taken to be the jet triplet with the highest BDT output. In this figure we also show the discrimination power between signal and the reducible (fakes, mainly composed by top pair production) and irreducible (MC-only, composed by events that are matched to generation level ℓ's and τ_h's) parts of BKG. As expected the discriminator is more potent against reducible BKG, where the had-top kinematics is expected to be more discriminant. As the $1\ell + 2\tau_h$ category is dominated by fake BKG the impact of the tagger on the final analysis is more prominent.

Figure 10: Distributions in BDT output (top), response (middle) and in top quark p_{T} (bottom), for had-top candidates reconstructed by the resHTT algorithm in events selected in the $1\ell + 2\tau_{\mathrm{h}}$ (left) and $2\ell ss + 1\tau_{\mathrm{h}}$ (right) category, where the had-top candidate reconstructed in a given event is taken to be the jet triplet with the highest BDT output. The label "MC-only BKG" refers to the had-top candidates reconstructed from the parts of the processes $\mathrm{t\bar{t}W}$, $\mathrm{t\bar{t}Z}$, WW, WZ and rares that and are considered irreducible background (that are matched to generation level ℓ's and τ_{h}'s) and "fakes" refer to the had-top candidates reconstructed with the part of the BKG tagged as "fake" (reducible) background.



D Hj Tagger

The "Higgs-jet" (Hj) tagger is a BDT-based discriminant that targets the identification of jets originating from a H boson decaying into two W bosons. Jets that are identified as originating from had-top decays by the algorithm described in Section C are not considered by the Hj tagger. The algorithm presented in this paragraph has been developed and used in the analysis with 2016 data. In the following, we present updated results for the Hj tagger for the current analysis and show that it has a similar performance compared to the 2016 version.

The Hj tagger focuses on the $2\ell ss$ category, where the H boson decays to two W bosons via:

$$\mathrm{t\bar{t}H} \rightarrow \mathrm{bW\,bW\,WW^{*}} \rightarrow \mathrm{bjj\,b}\ell^{\circ}{}_{\ell}\,\ell^{\circ}{}_{\ell}\mathrm{jj}\,.$$

The selection criteria described in Section 5.4 require the presence of two same-sign leptons and at least four jets, which means that one of the two W bosons originating from the H boson decay decays hadronically. The two jets do not necessarily both pass the $p_{\mathrm{T}} > 25$ GeV and $|\eta| < 2.4$ cuts. Allowing for one of the two jets to fail the p_{T} and η cuts, we have developed the Hj tagger to assess the likelihood for a jet to originate from the decay chain $\mathrm{H} \rightarrow \mathrm{WW^{*}} \rightarrow \ell\nu_{\ell}\mathrm{jj}$.

The algorithm is trained on simulated samples of $\mathrm{t\bar{t}H}$, $\mathrm{t\bar{t}W}$, and $\mathrm{t\bar{t}Z}$ events. Jets originating from $\mathrm{H} \rightarrow \mathrm{WW^{*}} \rightarrow \ell\nu_{\ell}\mathrm{jj}$ decays are considered as "signal" and jets that do not originate from such decays are considered as "background".

The following observables are used as input to the BDT:

- minimum ΔR between the jet and any electron or muon passing the fakeable lepton selection criteria
- maximum ΔR between the jet and any electron or muon passing the fakeable lepton selection criteria
- jet p_{T}
- DeepCSV b-tagging discriminant
- quark-gluon discriminant [111] of the jet

The distribution in BDT output of the Hj tagger for signal and background jets and the corresponding ROC curves are shown in Fig. 11. The ROC curve demonstrates that the performance of the Hj tagger to separate signal from background jets is similar for the 2016 and 2017 versions of the analysis.

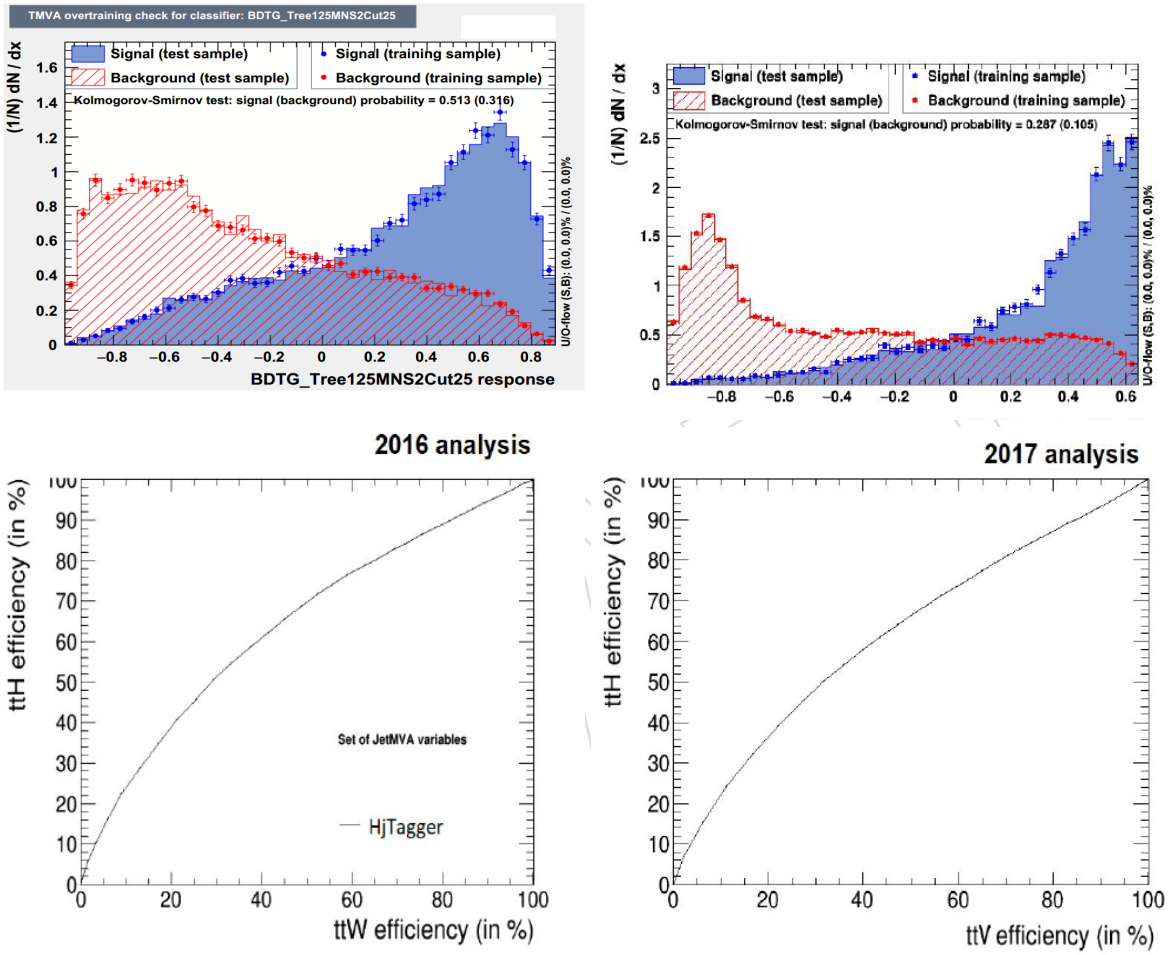


Figure 11: Distributions in the Hj discriminant (top) and ROC curve for the separation between signal and background jets (bottom) for the Hj taggers used in the 2016 (left) and 2017 (right) versions of the analysis (see text).

E Matrix Element Method

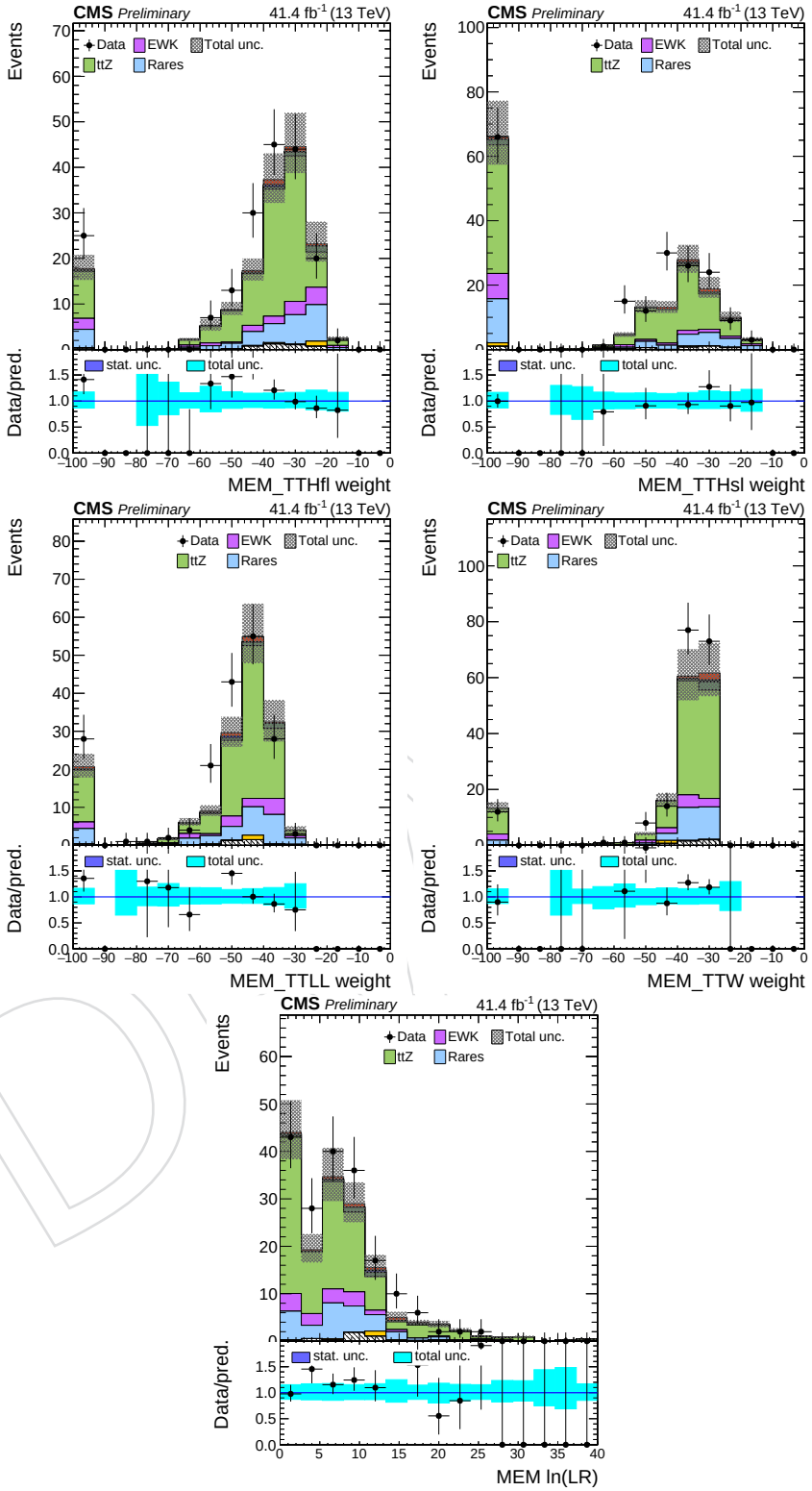


Figure 12: Distributions of signal, TT di-lepton and TTW background MEM weights, as well as discribution of MEM likelihood ratio, in $t\bar{t}Z$ control region as defined in I.2.

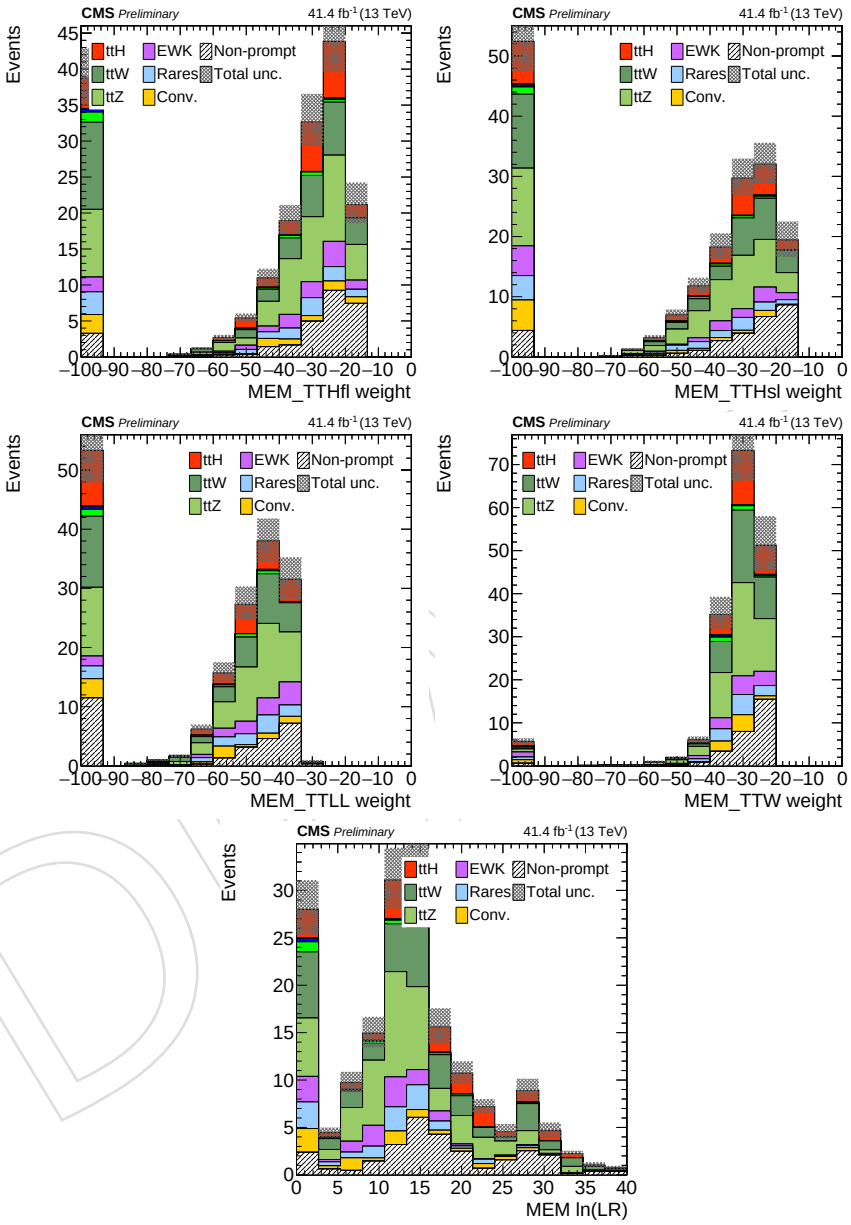


Figure 13: Distributions of signal, TT di-lepton and TTW background MEM weights, as well as discribution of MEM likelihood ratio in 3ℓ signal region. NOTE: data is still blinded.

HLT path	Prescale	Lepton cone-p_T	Lepton reco-p_T	Jet p_T
HLT_Ele8_CaloIdM_TrackIdM_PFJet30	1.14×10^4	15-45 GeV	$>$ 27 GeV	$>$ 30 GeV
HLT_Ele17_CaloIdM_TrackIdM_PFJet30	1.17×10^3	25-100 GeV	$>$ 17 GeV	$>$ 30 GeV
HLT_Ele23_CaloIdM_TrackIdM_PFJet30	1.07×10^3	32-100 GeV	$>$ 23 GeV	$>$ 30 GeV
HLT_Mu3_PFJet40	8.99×10^3	10-32 GeV	$>$ 3 GeV	$>$ 45 GeV
HLT_Mu8	1.59×10^4	15-45 GeV	$>$ 8 GeV	$>$ 30 GeV
HLT_Mu17	5.94×10^2	32-100 GeV	$>$ 17 GeV	$>$ 30 GeV
HLT_Mu20	2.25×10^2	32-100 GeV	$>$ 20 GeV	$>$ 30 GeV
HLT_Mu27	2.25×10^2	45-100 GeV	$>$ 27 GeV	$>$ 30 GeV

Table 26: Triggers used to record events for the measurement of the lepton misidentification rate. A hyphen ($-$) indicates requirements that are not applied.

F Measurement of lepton misidentification rates

The lepton misidentification rates are measured using an event sample enriched in the contribution from multijet production. The sample is selected by requiring the presence of exactly one electron or muon that passes the fakeable lepton selection criteria given by Tables 4 and 5 and at least one jet, which is separated from the lepton by $\Delta R > 0.7$. The events are recorded using single electron and muon triggers of various p_T thresholds. The corresponding HLT paths are given in Table 26. Some of the triggers require the presence of a jet on trigger level, where the jet p_T threshold varies by HLT path. Events selected in the multijet sample, referred to as MR, are required to contain at least one "offline" reconstructed jet with a p_T above the p_T threshold of the trigger. None of the triggers demands that the lepton satisfies an isolation condition on trigger level, which results in high trigger rates and necessitate to apply prescales to the triggers, in order to satisfy the bandwidth limitations of the data-acquisition system. The values of the prescales are given in the Table.

The offline reconstructed leptons are required to satisfy conditions on the reco-p_T and on the cone-p_T, depending on the HLT path that triggers the event. The requirement on the reco-p_T is applied to ensure that the trigger is fully efficient, while the condition on the cone-p_T is applied to reduce biases in the measurement of the lepton misidentification rates. The issue is that the p_T thresholds that are applied on trigger level impose an implicit condition on the isolation of the lepton, because the lepton misidentification rates are parametrized by the cone-p_T, and not the reco-p_T, of the lepton. Such bias on the isolation of the lepton would cause a bias on the output of the BDT that is used to separate prompt leptons from non-prompt and fake leptons, increasing the probability for fakeable leptons to pass the tight lepton selection criteria. This in turn would cause the estimate of the fake background in the SR to be overestimated, as all leptons selected in the MR are required to pass the trigger requirements, whereas events selected in the AR contain multiple leptons and only one of these leptons must pass the trigger requirements. A bias of this kind is avoided by requiring that the cone-p_T of the offline reconstructed lepton exceeds the p_T threshold applied on trigger level by about a factor two. The condition $I_\ell < 0.4 \times$ reco-p_T that is applied in the selection of fakeable electrons and muons then ensures that leptons passing the threshold on the cone-p_T pass the p_T threshold of the trigger regardless of their isolation, avoiding the aforementioned bias. A small residual bias, which is caused by the electron identification criteria applied on trigger level, is reduced by requiring fakeable electrons to satisfy conditions on $\sigma_{i\eta i\eta}$, H/E, and 1/E - 1/p (*cf.* Table 4).

The p_T cuts that are applied in the offline event selection on the p_T of the jet and on the reco-p_T and cone-p_T of the lepton are also given in Table 26. For events selected in the MR, we demand that at least of the HLT paths given in the Table triggered the event and the event contains at least one pair of fakeable lepton plus jet, which satisfies the conditions on cone-p_T as well as reco-p_T of the lepton and on the p_T of the jet p_T for any of the HLT paths that triggered the

event.

The event yield of processes that are modeled by the MC simulation is corrected for the effect of the trigger prescales by applying suitable chosen weights to the simulated events. As the prescales are applied independently for each HLT path, the weights are given by the expression:

$$w = 1 - \prod_{\mathrm{p}} \left(1 - \frac{1}{\mathrm{prescale}_{\mathrm{p}}} \right) , \tag{10}$$

where the product extends over all HLT paths p that "fired" for a given event.

The fake-factors (FF) are measured in bins of η and cone-p_{T} of the lepton. In each bin, the the FF is measured by determining the number of multijet events with fakeable leptons that pass the tight lepton selection criteria and those that fail. The number of multijet events in the pass and fail regions, denoted by N_{pass} and N_{fail}, is determined by a maximum-likelihood (ML) fit of the distribution in the observable:

$$m_{\mathrm{T}}^{\mathrm{fix}} = \sqrt{2\, p_{\mathrm{T}}^{\mathrm{fix}}\, E_T^{miss}\ (1 - \cos\Delta\phi)} \tag{11}$$

with shape templates for W+jets, Z+jets, diboson, top quark pair, and multijet production. Given the event yields N_{pass} and N_{fail} of multijet production in the pass and fail regions, the FF f_i for a given bin i in η and cone-p_{T} of the lepton are computed by:

$$f_i = \frac{N_{\mathrm{pass}}}{N_{\mathrm{pass}} + N_{\mathrm{fail}}} . \tag{12}$$

The symbol $\Delta\phi$ in Eq. (11) refers to the angle in the transverse plane between the lepton momentum and the E_T^{miss} vector, while the symbol $p_{\mathrm{T}}^{\mathrm{fix}}$ denotes a constant of value 35 GeV. The observable $m_{\mathrm{T}}^{\mathrm{fix}}$ exploits the fact that the E_T^{miss} reconstructed in multijet events is mainly due to resolution effects and is typically small, resulting in a falling distribution in $m_{\mathrm{T}}^{\mathrm{fix}}$, while the distribution in W+jets events, which constitute the main background for the FF measurement, exhibits a broad maximum around $m_{\mathrm{W}} \approx 80$ GeV. Compared to the usual transverse mass, the observable $m_{\mathrm{T}}^{\mathrm{fix}}$ has the advantage of not depending on the p_{T} of the lepton, which is better suited for the purpose of measuring the FF in bins of lepton p_{T} [76]. The contributions from W+jets, Z+jets, and diboson production are assumed to scale by a common factor with respect to their MC expectation in the fit, and their sum is referred to as "electroweak" background. The shape templates for multijet production are obtained from the MC simulation, using the samples given in Table 27. The contributions of other processes to the pass and fail regions is small. These contributions are estimated using the MC simulation and subtracted from the data before the fit. The pass and fail regions for each p_{T} and η bin are fitted simultaneously, taking correlations between the systematic uncertainties between both regions into account, while individual p_{T} and η bins are fitted independently. The fit is illustrated for two example p_{T} and η bins, one for electrons and one for muons, in Fig. 14.

A correction needs to be applied to the measured FF for electrons to account for the contribution of photon conversions to the pass and fail regions. The event yields N_{pass} and N_{fail} obtained by the fit are scaled by the relative fraction of multijet events in which the reconstructed fakeable electron is matched, within $\Delta R < 0.3$, to a generator level prompt photon (i.e. not a photon arising from the decays of hadrons) The fraction is determined separately for the pass and fail region and for each bin in cone-p_{T} and η of the lepton. The effect of this correction is to reduce the FF for electrons by 10-20%.

Table 27: List of MC samples used to obtain the shape templates for multijet production for the measurement of lepton misidentification rates.

Sample name	Cross section [pb]
/QCD_Pt-20toInf_MuEnrichedPt15_TuneCP5_13TeV_pythia8/[3]	2.38×10^5
/QCD_Pt-20to30_MuEnrichedPt5_TuneCP5_13TeV_pythia8/[1]	2.49×10^6
/QCD_Pt-30to50_MuEnrichedPt5_TuneCP5_13TeV_pythia8/[1]	1.36×10^6
/QCD_Pt-50to80_MuEnrichedPt5_TuneCP5_13TeV_pythia8/[1]	3.78×10^5
/QCD_Pt-80to120_MuEnrichedPt5_TuneCP5_13TeV_pythia8/[1]	8.84×10^4
/QCD_Pt-120to170_MuEnrichedPt5_TuneCP5_13TeV_pythia8/[1]	2.13×10^4
/QCD_Pt-20to30_EMEnriched_TuneCP5_13TeV_pythia8/[1]	4.93×10^6
/QCD_Pt-30to50_EMEnriched_TuneCP5_13TeV_pythia8/[1]	6.41×10^6
/QCD_Pt-50to80_EMEnriched_TuneCP5_13TeV_pythia8/[1]	1.99×10^6
/QCD_Pt-80to120_EMEnriched_TuneCP5_13TeV_pythia8/[1]	3.71×10^5
/QCD_Pt-120to170_EMEnriched_TuneCP5_13TeV_pythia8/[1]	6.68×10^4
/QCD_Pt_20to30_bcToE_TuneCP5_13TeV_pythia8/[2]	3.14×10^5
/QCD_Pt_30to80_bcToE_TuneCP5_13TeV_pythia8/[2]	3.62×10^5
/QCD_Pt_80to170_bcToE_TuneCP5_13TeV_pythia8/[2]	3.38×10^4

[1] RunIIFall17MiniAOD-94X_mc2017_realistic_v10-v1/MINIAODSIM

[2] RunIIFall17MiniAOD-94X_mc2017_realistic_v11-v1/MINIAODSIM

[3] RunIIFall17MiniAOD-RECOSIMstep_94X_mc2017_realistic_v10-v1/MINIAODSIM

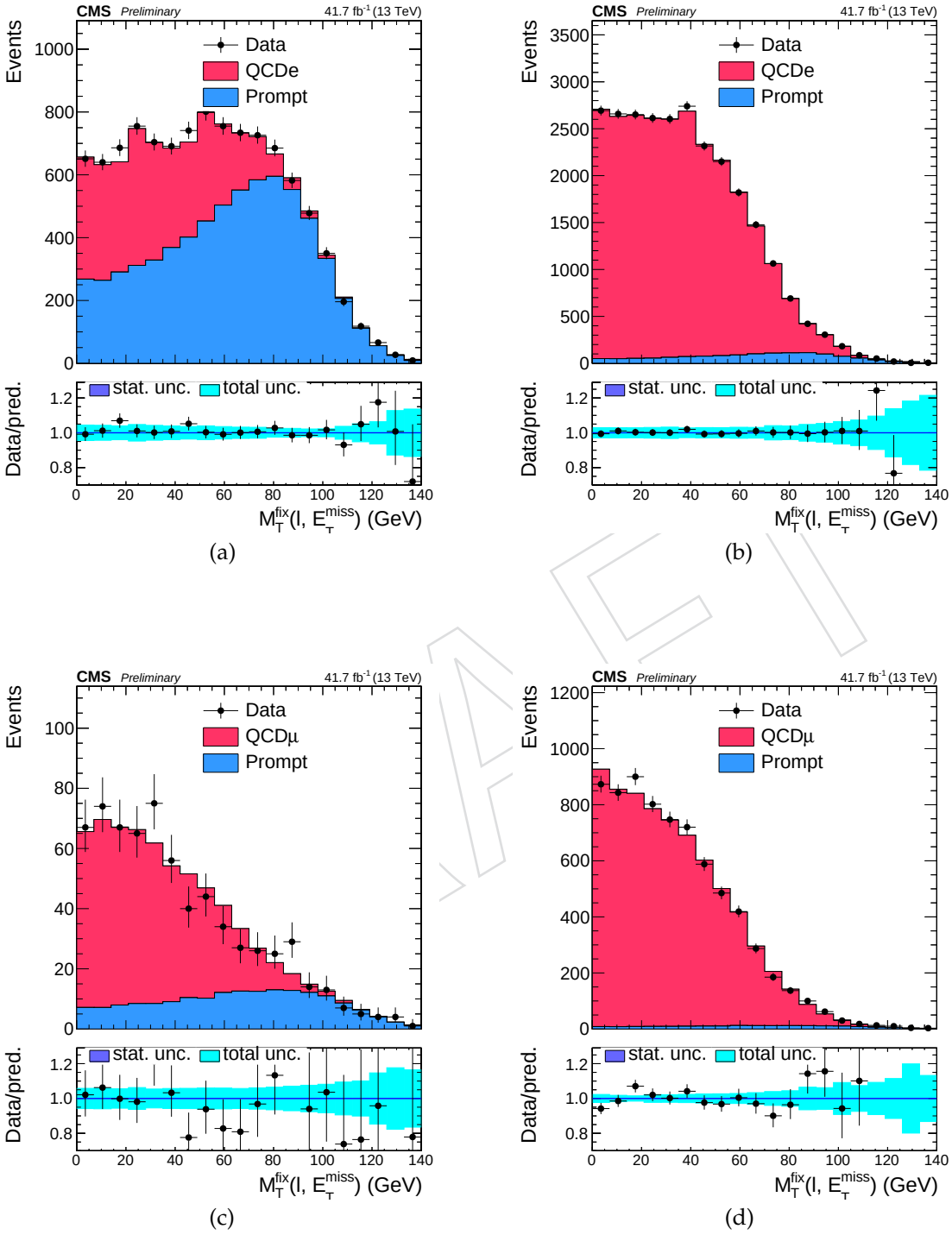


Figure 14: Distributions in the observable $m_{\mathrm{T}}^{\mathrm{fix}}$ in the pass (a, c) and fail (b, d) regions. for two example bins in η and cone-p_{T} of the lepton: for electrons of $25 <$ cone-$p_{\mathrm{T}} < 35$ GeV and $|\eta| < 1.479$ (a, b) and for muons of $15 <$ cone-$p_{\mathrm{T}} < 20$ GeV and $|\eta| > 1.2$ (b, d).

The FF for electrons and muons, obtained from the fit, and the effect of the correction that is applied to account for photon conversions are shown in Fig. 15. Numerical values are given in Table 28.

Electrons		
	$\lvert\eta\rvert < 1.479$	$\lvert\eta\rvert > 1.479$
$15 <$ cone-$p_{\mathrm{T}} < 25$ GeV	$0.132^{+0.012}_{-0.012}$	$0.095^{+0.015}_{-0.015}$
$25 <$ cone-$p_{\mathrm{T}} < 35$ GeV	$0.121^{+0.008}_{-0.008}$	$0.084^{+0.009}_{-0.009}$
$35 <$ cone-$p_{\mathrm{T}} < 45$ GeV	$0.118^{+0.019}_{-0.019}$	$0.087^{+0.012}_{-0.012}$
$45 <$ cone-$p_{\mathrm{T}} < 65$ GeV	$0.109^{+0.025}_{-0.025}$	$0.077^{+0.026}_{-0.026}$
cone-$p_{\mathrm{T}} > 65$ GeV	$0.118^{+0.055}_{-0.055}$	$0.062^{+0.041}_{-0.041}$

Muons		
	$\lvert\eta\rvert < 1.2$	$\lvert\eta\rvert > 1.2$
cone-$p_{\mathrm{T}} < 15$ GeV	$0.024^{+0.005}_{-0.005}$	$0.007^{+0.004}_{-0.004}$
$15 <$ cone-$p_{\mathrm{T}} < 20$ GeV	$0.086^{+0.006}_{-0.006}$	$0.066^{+0.005}_{-0.005}$
$20 <$ cone-$p_{\mathrm{T}} < 32$ GeV	$0.095^{+0.008}_{-0.008}$	$0.109^{+0.009}_{-0.009}$
$32 <$ cone-$p_{\mathrm{T}} < 45$ GeV	$0.080^{+0.019}_{-0.019}$	$0.090^{+0.012}_{-0.012}$
$45 <$ cone-$p_{\mathrm{T}} < 65$ GeV	$0.093^{+0.039}_{-0.039}$	$0.089^{+0.050}_{-0.050}$
cone-$p_{\mathrm{T}} > 65$ GeV	$0.093^{+0.078}_{-0.078}$	$0.084^{+0.093}_{-0.084}$

Table 28: Probabilities f for non-prompt leptons and hadrons, which pass the fakeable lepton selection criteria, to pass the tight lepton selection criteria, measured as function of η and cone-p_{T} of the lepton in multijet events.

Small differences between the FF in simulated multijet and top quark pair production events can be seen in Fig. 15. The differences in the FF reflect differences in the fractions of non-prompt leptons and hadrons between multijet and top quark pair production events, with multijet events containing a higher fraction of fake leptons from light quark and gluon jets, and top quark pair production events containing a higher fraction of non-prompt leptons from decays of heavy (c and b) quarks. The differences between the FF have been reduced by tuning the fakeable lepton selection criteria, in particular by adjusting the cut on the CSVv2 b-tagging discriminant that is applied in the selection of fakeable electrons and muons. The small residual differences are not corrected for, but considered as systematic uncertainties (*cf.* Section 9).

An alternative procedure to measure the FF, which does not rely on the MC simulation to obtain the shape templates for multijet production, is used as a cross-check. In the alternative procedure, the event yield N_{fail} of multijet production is determined by subtracting the MC expectation for W+jets, Z+jets, diboson, and top quark pair production from the data in the fail region. The distribution in $m_{\mathrm{T}}^{\mathrm{fix}}$ observed in the fail region, corrected by subtracting the expected contributions of W+jets, Z+jets, diboson, and top quark pair production, is then used as shape template for multijet production when performing the ML fit in the pass region. The event yield N_{pass} of multijet production in the pass region is determined by an ML fit as before. The results of this cross-check agree with the nominal values of the FF within 10-20% [112]. The difference is within the systematic uncertainties associated to the FF method.

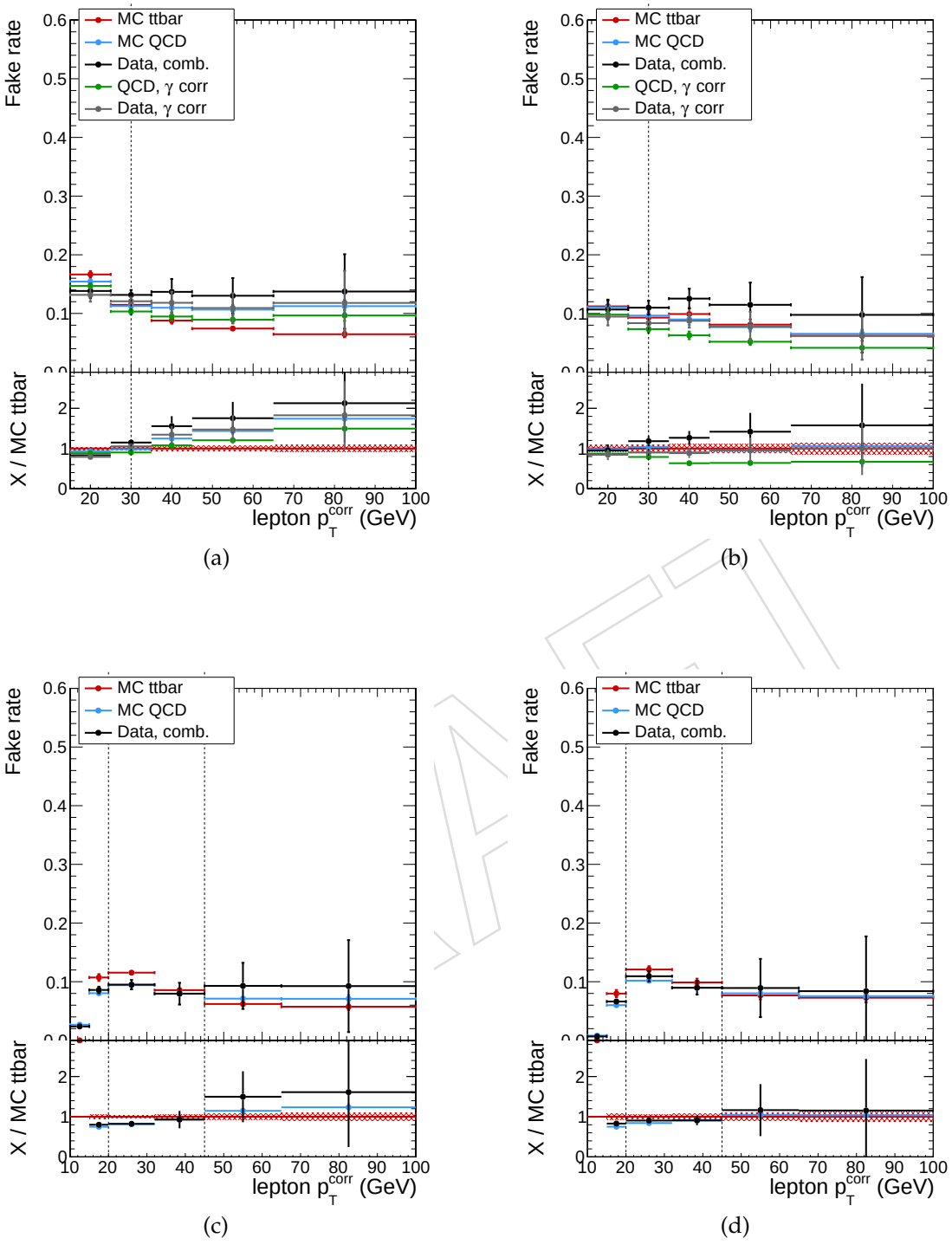


Figure 15: Probabilities f for non-prompt leptons and hadrons, which pass the fakeable lepton selection criteria, to pass the tight lepton selection criteria, measured as function of cone-p_{T} of the lepton in multijet events, for electrons of $|\eta| < 1.479$ (a) and $|\eta| > 1.479$ (b), and for muons of $|\eta| < 1.2$ (c) and $|\eta| > 1.2$ (d). The FF in simulated multijet and top quark pair production events are shown for comparison. In case of simulated top quark pair production events, only those reconstructed fakeable leptons, which are not matched to prompt leptons on generator level, are considered in the FF computation.

G Measurement of τ_h misidentification rate

The rate with which quark and gluon jets passing the fakeable τ_h selection criteria pass the tight τ_h selection (*cf.* Table 6) is measured using $t\bar{t}$+jets events in which the two W bosons produced in the top quark decays decay to an electron plus muon pair. Electron and muon are required to pass the tight selection criteria defined in Tables 4 and 5, respectively, and to be of opposite charge. The "leading" lepton (lepton of higher p_T) is required to have $p_T > 25$ GeV, while the "subleading" lepton (lepton of lower p_T) is required to have $p_T > 15$ GeV. Selected events are further required to contain at least one τ_h candidate passing the fakeable τ_h selection and ≥ 2 jets of $p_T > 25$ GeV and $|\eta| < 2.4$, of which either ≥ 2 pass the loose WP of the DeepCSV b-tagging discriminant (*cf.* Section 3.3), or ≥ 1 passes the medium WP. The jets are required not to overlap, within $\Delta R = 0.3$, with the electron, the muon, and any τ_h candidate that passes the fakeable τ_h selection. The contribution of backgrounds with lepton pairs of low mass, which are not well modelled by the MC samples used in the analysis, are removed by requiring the mass of the electron plus muon pair to satisfy the condition $m_{e\mu} > 12$ GeV. The purity of the $t\bar{t}$+jets sample is enhanced by requiring the linear discriminant based on E_T^{miss} and H_T^{miss}, defined by Eq. (10), to satisfy the condition $E_T^{miss}\text{LD} > 0.2$. Events selected in data are required to pass either the triggers based on the presence of an electron plus muon pair, the single muon, or the single electron triggers (*cf.* Table 8).

Residual contributions of background processes in which the reconstructed τ_h are due to a genuine τ_h, an electron, or a muon are subtracted based on the MC simulation.

The measured probabilities for jets passing the fakeable τ_h selection to pass the loose WP of the τ_h identification discriminant applied in the $2\ell ss + 1\tau_h$ and $3\ell + 1\tau_h$ channels, and to pass the medium WP, which is applied in the $1\ell + 2\tau_h$ and $2\ell + 2\tau_h$ channels, are compared to the MC expectation in Fig. 16. The probabilities, denoted by f, are determined in bins of τ_h candidate p_T and separately for τ_h candidates reconstructed in the central ($|\eta| < 1.479$) and in the forward ($|\eta| > 1.479$) region of the detector. Numerical values of the probabilities are given in Table 29. The data-to-MC ratio of these probabilities is fitted by linear functions of the form $c_0 + c_1 \cdot p_T$. In case the statistical uncertainty of f amounts to less than 2% the statistical uncertainty is set to 2%, in order to avoid that bins at low p_T, with high event statistics, dominate the fit.

In order to reduce statistical fluctuations when using the measured probabilities for the purpose of obtaining a data-driven estimate of the fake τ_h background as described in Section 8, we do not directly use the probabilities measured in data, but instead use the MC expectation for these probabilities times the function obtained by fitting the data-to-MC ratio. Control plots of the fits are shown in Fig. 17.

The uncertainty on the parameters of the fit is considered as systematic uncertainty on the fake τ background estimate (*cf.* Section 9). Possible correlations between the parameters c_0 and c_1 are accounted for by determining the Eigenvectors, v_i, and Eigenvalues, λ_i, of the covariance matrix $cov(c_0, c_1)$ returned by the fit. The Eigenvectors yield transformed parameters, $k_i = v_{i0} \cdot c_0 + v_{i1} \cdot c_1$ ($k = 0, 1$), which are uncorrelated; the Eigenvalues represent the uncertainties on the transformed parameters. The systematic uncertainties on the fake τ background estimate are then obtained by using the functions $(1 + \lambda_i v_{i0})\ c_0 + (1 + \lambda_i v_{i1})\ c_1 \cdot p_T$ and $(1 - \lambda_i v_{i0})\ c_0 + (1 - \lambda_i v_{i1})\ c_1 \cdot p_T$ as well as the function $c_0 + c_1 \cdot p_T$ as input to the background estimation procedure described in Section 8 and taking the difference. The slope and offset coefficients of the linear functions obtained from the fit and of the functions used for estimating the systematic uncertainties are given in Table 30.

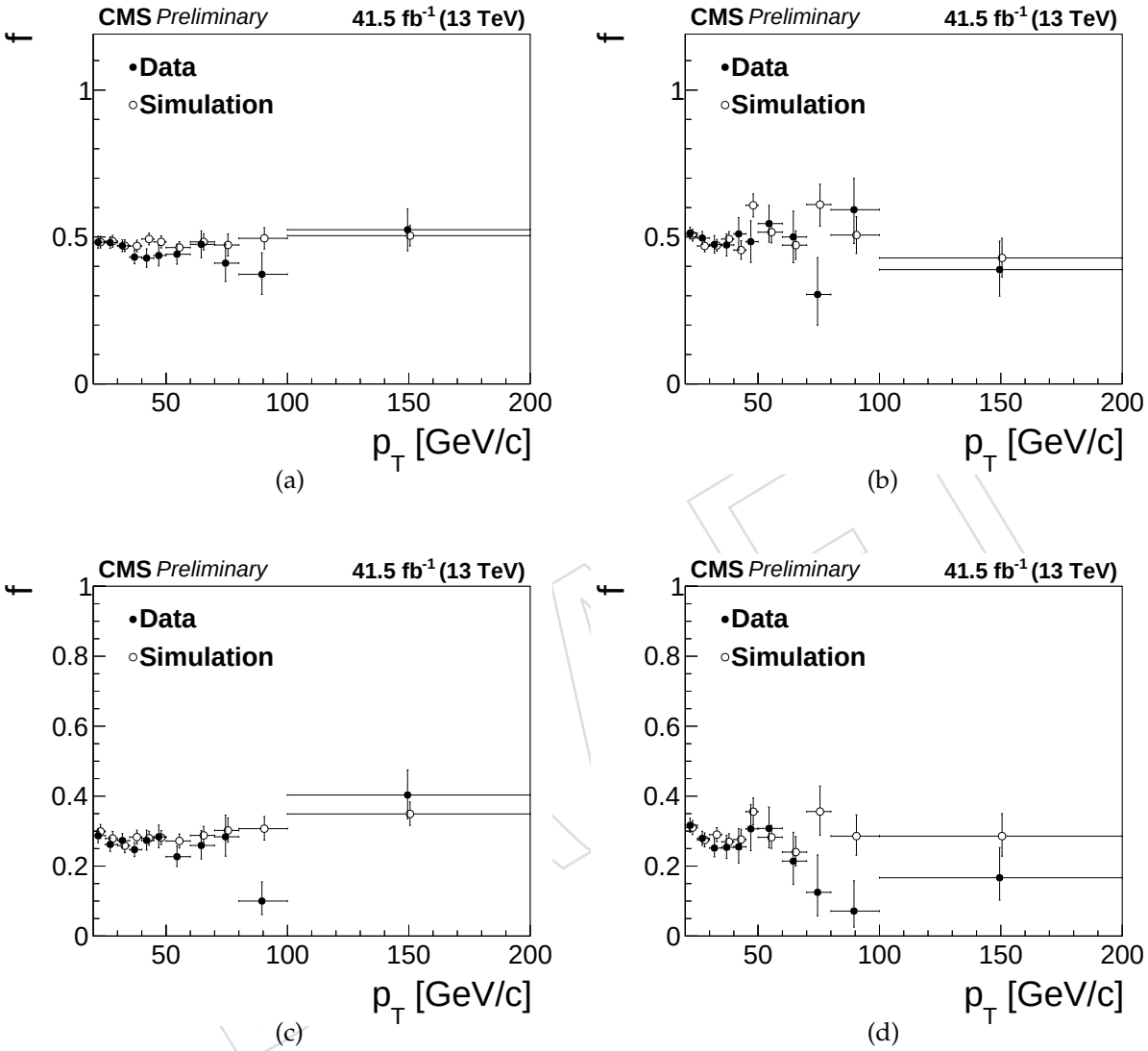
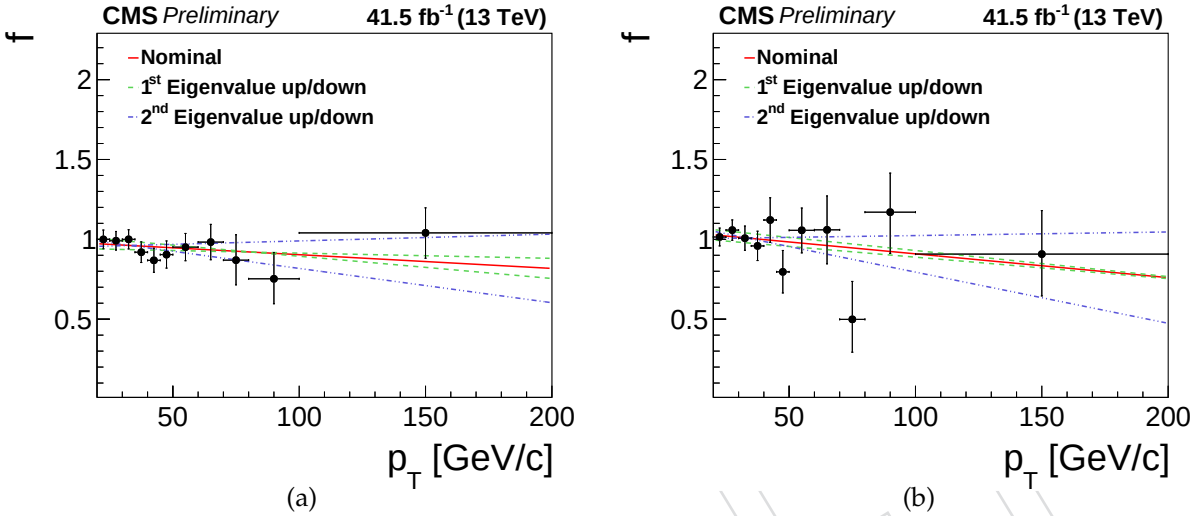


Figure 16: Probabilities f for jets passing the fakeable τ_h selection to pass the loose WP of the τ_h identification discriminant (a, b) applied in the $2\ell ss + 1\tau_h$ and $3\ell + 1\tau_h$ channels, and to pass the medium WP (c, d), which is applied in the $1\ell + 2\tau_h$ and $2\ell + 2\tau_h$ channels, separately for τ_h candidates reconstructed in the regions $|\eta| < 1.479$ (a, c) and $|\eta| > 1.479$ (b, d).

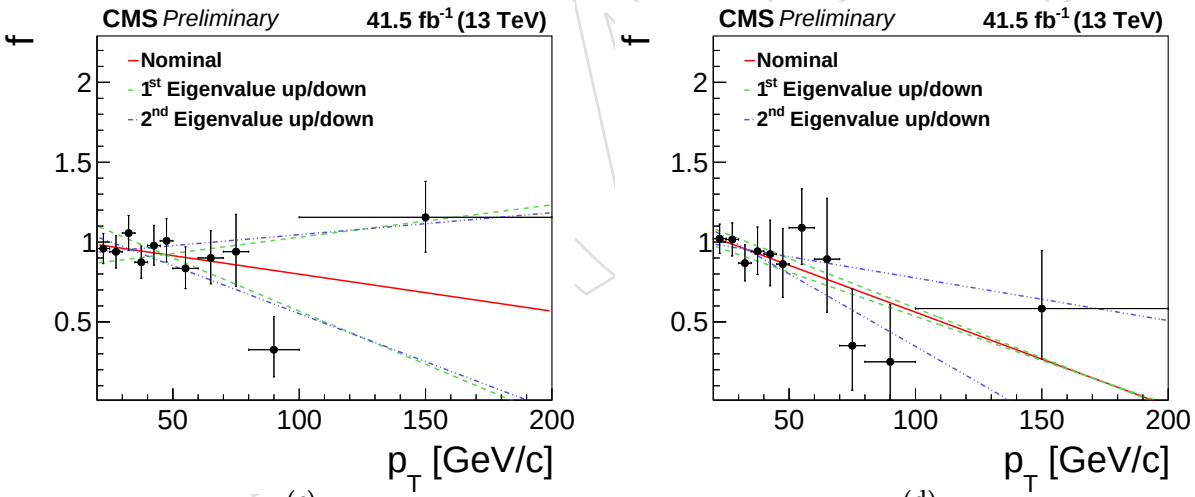
$2\ell ss + 1\tau_h$ and $3\ell + 1\tau_h$ channels				
Kinematic range		Data	Simulation	Data/simulation
$20 < p_T < 25$ GeV	$\|\eta\| < 1.479$	$0.4821^{+0.0200}_{-0.0200}$	$0.4823^{+0.0200}_{-0.0200}$	$0.9995^{+0.0283}_{-0.0283}$
$25 < p_T < 30$ GeV	$\|\eta\| < 1.479$	$0.4809^{+0.0200}_{-0.0200}$	$0.4852^{+0.0200}_{-0.0200}$	$0.9910^{+0.0283}_{-0.0283}$
$30 < p_T < 35$ GeV	$\|\eta\| < 1.479$	$0.4700^{+0.0200}_{-0.0200}$	$0.4700^{+0.0200}_{-0.0200}$	$0.9999^{+0.0283}_{-0.0283}$
$35 < p_T < 40$ GeV	$\|\eta\| < 1.479$	$0.4314^{+0.0230}_{-0.0228}$	$0.4693^{+0.0200}_{-0.0200}$	$0.9192^{+0.0305}_{-0.0303}$
$40 < p_T < 45$ GeV	$\|\eta\| < 1.479$	$0.4281^{+0.0317}_{-0.0311}$	$0.4929^{+0.0200}_{-0.0200}$	$0.8684^{+0.0375}_{-0.0370}$
$45 < p_T < 50$ GeV	$\|\eta\| < 1.479$	$0.4369^{+0.0358}_{-0.0352}$	$0.4832^{+0.0208}_{-0.0207}$	$0.9043^{+0.0413}_{-0.0409}$
$50 < p_T < 60$ GeV	$\|\eta\| < 1.479$	$0.4412^{+0.0345}_{-0.0340}$	$0.4640^{+0.0201}_{-0.0200}$	$0.9509^{+0.0399}_{-0.0395}$
$60 < p_T < 70$ GeV	$\|\eta\| < 1.479$	$0.4748^{+0.0460}_{-0.0455}$	$0.4832^{+0.0278}_{-0.0277}$	$0.9826^{+0.0537}_{-0.0534}$
$70 < p_T < 80$ GeV	$\|\eta\| < 1.479$	$0.4110^{+0.0653}_{-0.0626}$	$0.4726^{+0.0378}_{-0.0375}$	$0.8695^{+0.0753}_{-0.0731}$
$80 < p_T < 100$ GeV	$\|\eta\| < 1.479$	$0.3729^{+0.0732}_{-0.0684}$	$0.4954^{+0.0363}_{-0.0362}$	$0.7527^{+0.0817}_{-0.0774}$
$100 < p_T < 200$ GeV	$\|\eta\| < 1.479$	$0.5246^{+0.0710}_{-0.0719}$	$0.5043^{+0.0349}_{-0.0349}$	$1.0400^{+0.0792}_{-0.0799}$
$20 < p_T < 25$ GeV	$\|\eta\| > 1.479$	$0.5131^{+0.0200}_{-0.0200}$	$0.5058^{+0.0200}_{-0.0200}$	$1.0150^{+0.0283}_{-0.0283}$
$25 < p_T < 30$ GeV	$\|\eta\| > 1.479$	$0.4964^{+0.0223}_{-0.0222}$	$0.4693^{+0.0200}_{-0.0200}$	$1.0580^{+0.0299}_{-0.0299}$
$30 < p_T < 35$ GeV	$\|\eta\| > 1.479$	$0.4745^{+0.0298}_{-0.0296}$	$0.4712^{+0.0200}_{-0.0200}$	$1.0070^{+0.0359}_{-0.0358}$
$35 < p_T < 40$ GeV	$\|\eta\| > 1.479$	$0.4726^{+0.0378}_{-0.0375}$	$0.4930^{+0.0253}_{-0.0253}$	$0.9587^{+0.0454}_{-0.0452}$
$40 < p_T < 45$ GeV	$\|\eta\| > 1.479$	$0.5102^{+0.0552}_{-0.0554}$	$0.4552^{+0.0317}_{-0.0314}$	$1.1210^{+0.0635}_{-0.0638}$
$45 < p_T < 50$ GeV	$\|\eta\| > 1.479$	$0.4839^{+0.0712}_{-0.0706}$	$0.6077^{+0.0383}_{-0.0396}$	$0.7962^{+0.0814}_{-0.0803}$
$50 < p_T < 60$ GeV	$\|\eta\| > 1.479$	$0.5455^{+0.0621}_{-0.0634}$	$0.5163^{+0.0362}_{-0.0364}$	$1.0570^{+0.0720}_{-0.0731}$
$60 < p_T < 70$ GeV	$\|\eta\| > 1.479$	$0.5000^{+0.0879}_{-0.0879}$	$0.4720^{+0.0487}_{-0.0482}$	$1.0590^{+0.1002}_{-0.1004}$
$70 < p_T < 80$ GeV	$\|\eta\| > 1.479$	$0.3043^{+0.1250}_{-0.1047}$	$0.6102^{+0.0692}_{-0.0734}$	$0.4988^{+0.1449}_{-0.1255}$
$80 < p_T < 100$ GeV	$\|\eta\| > 1.479$	$0.5926^{+0.1067}_{-0.1147}$	$0.5065^{+0.0629}_{-0.0631}$	$1.1700^{+0.1239}_{-0.1308}$
$100 < p_T < 200$ GeV	$\|\eta\| > 1.479$	$0.3889^{+0.0972}_{-0.0901}$	$0.4286^{+0.0669}_{-0.0646}$	$0.9074^{+0.1167}_{-0.1122}$
$1\ell + 2\tau_h$ and $2\ell + 2\tau_h$ channels				
Kinematic range		Data	Simulation	Data/simulation
$20 < p_T < 25$ GeV	$\|\eta\| < 1.479$	$0.2869^{+0.0200}_{-0.0200}$	$0.2991^{+0.0200}_{-0.0200}$	$0.9594^{+0.0283}_{-0.0283}$
$25 < p_T < 30$ GeV	$\|\eta\| < 1.479$	$0.2619^{+0.0200}_{-0.0200}$	$0.2789^{+0.0200}_{-0.0200}$	$0.9391^{+0.0283}_{-0.0283}$
$30 < p_T < 35$ GeV	$\|\eta\| < 1.479$	$0.2729^{+0.0200}_{-0.0200}$	$0.2584^{+0.0200}_{-0.0200}$	$1.0560^{+0.0283}_{-0.0283}$
$35 < p_T < 40$ GeV	$\|\eta\| < 1.479$	$0.2471^{+0.0206}_{-0.0200}$	$0.2829^{+0.0200}_{-0.0200}$	$0.8734^{+0.0287}_{-0.0283}$
$40 < p_T < 45$ GeV	$\|\eta\| < 1.479$	$0.2734^{+0.0294}_{-0.0276}$	$0.2797^{+0.0200}_{-0.0200}$	$0.9776^{+0.0355}_{-0.0341}$
$45 < p_T < 50$ GeV	$\|\eta\| < 1.479$	$0.2838^{+0.0335}_{-0.0314}$	$0.2816^{+0.0200}_{-0.0200}$	$1.0080^{+0.0390}_{-0.0372}$
$50 < p_T < 60$ GeV	$\|\eta\| < 1.479$	$0.2269^{+0.0304}_{-0.0280}$	$0.2718^{+0.0200}_{-0.0200}$	$0.8349^{+0.0364}_{-0.0344}$
$60 < p_T < 70$ GeV	$\|\eta\| < 1.479$	$0.2590^{+0.0425}_{-0.0387}$	$0.2877^{+0.0259}_{-0.0247}$	$0.9002^{+0.0492}_{-0.0466}$
$70 < p_T < 80$ GeV	$\|\eta\| < 1.479$	$0.2838^{+0.0620}_{-0.0555}$	$0.3020^{+0.0358}_{-0.0336}$	$0.9397^{+0.0706}_{-0.0660}$
$80 < p_T < 100$ GeV	$\|\eta\| < 1.479$	$0.1000^{+0.0550}_{-0.0389}$	$0.3070^{+0.0347}_{-0.0327}$	$0.3258^{+0.0640}_{-0.0521}$
$100 < p_T < 200$ GeV	$\|\eta\| < 1.479$	$0.4032^{+0.0715}_{-0.0680}$	$0.3491^{+0.0341}_{-0.0327}$	$1.1550^{+0.0786}_{-0.0761}$
$20 < p_T < 25$ GeV	$\|\eta\| > 1.479$	$0.3166^{+0.0200}_{-0.0200}$	$0.3101^{+0.0200}_{-0.0200}$	$1.0210^{+0.0283}_{-0.0283}$
$25 < p_T < 30$ GeV	$\|\eta\| > 1.479$	$0.2792^{+0.0205}_{-0.0200}$	$0.2749^{+0.0200}_{-0.0200}$	$1.0160^{+0.0286}_{-0.0283}$
$30 < p_T < 35$ GeV	$\|\eta\| > 1.479$	$0.2516^{+0.0269}_{-0.0252}$	$0.2899^{+0.0200}_{-0.0200}$	$0.8678^{+0.0335}_{-0.0322}$
$35 < p_T < 40$ GeV	$\|\eta\| > 1.479$	$0.2537^{+0.0344}_{-0.0318}$	$0.2693^{+0.0232}_{-0.0221}$	$0.9421^{+0.0409}_{-0.0393}$
$40 < p_T < 45$ GeV	$\|\eta\| > 1.479$	$0.2551^{+0.0517}_{-0.0461}$	$0.2760^{+0.0294}_{-0.0277}$	$0.9243^{+0.0586}_{-0.0547}$
$45 < p_T < 50$ GeV	$\|\eta\| > 1.479$	$0.3065^{+0.0696}_{-0.0626}$	$0.3556^{+0.0392}_{-0.0375}$	$0.8619^{+0.0790}_{-0.0738}$
$50 < p_T < 60$ GeV	$\|\eta\| > 1.479$	$0.3077^{+0.0610}_{-0.0555}$	$0.2824^{+0.0340}_{-0.0318}$	$1.0900^{+0.0688}_{-0.0651}$
$60 < p_T < 70$ GeV	$\|\eta\| > 1.479$	$0.2143^{+0.0824}_{-0.0665}$	$0.2400^{+0.0443}_{-0.0398}$	$0.8929^{+0.0915}_{-0.0799}$
$70 < p_T < 80$ GeV	$\|\eta\| > 1.479$	$0.1250^{+0.1068}_{-0.0672}$	$0.3559^{+0.0729}_{-0.0674}$	$0.3512^{+0.1263}_{-0.0992}$
$80 < p_T < 100$ GeV	$\|\eta\| > 1.479$	$0.0714^{+0.0865}_{-0.0460}$	$0.2857^{+0.0607}_{-0.0545}$	$0.2500^{+0.1023}_{-0.0762}$
$100 < p_T < 200$ GeV	$\|\eta\| > 1.479$	$0.1667^{+0.0863}_{-0.0641}$	$0.2857^{+0.0641}_{-0.0573}$	$0.5833^{+0.1035}_{-0.0906}$

Table 29: Probabilities f for jets passing the fakeable τ_h selection to pass the tight τ_h selection criteria applied in the $2\ell ss + 1\tau_h$, $3\ell + 1\tau_h$, $1\ell + 2\tau_h$, and $2\ell + 2\tau_h$ channels, in bins of τ_h candidate p_T and η.



(a)

(b)



(c)

(d)

Figure 17: Fits of linear functions of the form $c_0 + c_1 \cdot p_T$ to the data-to-MC ratios of the probabilities f shown in Fig. 16.

$2\ell ss + 1\tau_h$ and $3\ell + 1\tau_h$ channels				
Function	$\lvert\eta\rvert < 1.479$		$\lvert\eta\rvert > 1.479$	
	Offset	Slope	Offset	Slope
Nominal	0.9614	-8.607×10^{-4}	1.0084	-1.489×10^{-3}
1st Eigenvalue "up"	0.9874	-1.395×10^{-3}	1.0390	-1.635×10^{-3}
1st Eigenvalue "down"	0.9355	-3.269×10^{-4}	0.9778	-1.344×10^{-3}
2nd Eigenvalue "up"	0.9614	$+4.261 \times 10^{-4}$	1.0084	$+2.254 \times 10^{-4}$
2nd Eigenvalue "down"	0.9614	-2.148×10^{-3}	1.0084	-3.204×10^{-3}
$1\ell + 2\tau_h$ and $2\ell + 2\tau_h$ channels				
Function	$\lvert\eta\rvert < 1.479$		$\lvert\eta\rvert > 1.479$	
	Offset	Slope	Offset	Slope
Nominal	0.9539	-2.317×10^{-3}	0.9527	-5.875×10^{-3}
1st Eigenvalue "up"	1.0134	-6.663×10^{-3}	1.0016	-6.222×10^{-3}
1st Eigenvalue "down"	0.8944	$+2.028 \times 10^{-3}$	0.9039	-5.528×10^{-3}
2nd Eigenvalue "up"	0.9542	$+1.376 \times 10^{-3}$	0.9527	-2.666×10^{-3}
2nd Eigenvalue "down"	0.9536	-6.010×10^{-3}	0.9527	-9.084×10^{-3}

Table 30: Slope and offset coefficients obtained by fitting linear functions of the form $c_0 + c_1 \cdot p_T$ to the data-to-MC ratio of the probabilities f shown in Fig. 16 ("nominal") and of the corresponding functions ("1st Eigenvalue up/down", "2nd Eigenvalue up/down") used for estimating systematic uncertainties (see text).

H Measurement of the lepton charge misidentification rate

The measurement of the lepton charge misidentification rate is important to estimate the charge "flip" background in the $2\ell ss$ and $2\ell ss + 1\tau_h$ categories, which originates from events with two prompt leptons of opposite charge, in which the charge of one of the two leptons is mismeasured.

The charge misidentification probability for electrons (muons) is determined from data, by measuring the ratio of $Z/\gamma^* \to ee$ ($Z/\gamma^* \to \mu\mu$) events with two leptons of the same charge (SS) leptons to $Z/\gamma^* \to ee$ ($Z/\gamma^* \to \mu\mu$) events with two leptons of opposite charge (OS). The measurement utilizes the fact that $Z/\gamma^* \to ee$ events exhibit a narrow peak at $m_Z = 91.2$ GeV [52], even in case the charge of an electron is mismeasured, which allows to separate $Z/\gamma^* \to ee$ events from backgrounds by means of performing a maximum likelihood fit of the distribution in mass of the electron pair. The measurement is performed in bins of p_T and η of the leptons.

Events selected in the control region used to measure the charge misidentification rate are required to contain two leptons passing the "tight" lepton selection described in Section 3.1.4. Lepton pairs of SS and of OS charge are selected in separate control regions.

Samples of simulated events are used to obtain the shape templates for the maximum likelihood fit, as well as a prediction for the rate of $Z/\gamma^* \to \ell\ell$ ($\ell = e, \mu$) signal events and for the rate of background events in the SS and OS regions. The list of samples utilized in the measurement of the lepton charge misidentification rate is given in Table 2. The $Z/\gamma^* \to \ell\ell$ events are considered as "signal" in case both reconstructed leptons are matched to leptons of the same flavor (but possibly different charge) on generator level. The contribution of $Z/\gamma^* \to \ell\ell$ events in which one or both of the leptons are not matched to generator-level leptons and of the other processes given in Table 2 is considered as background.

H.1 Electrons

The electron charge misidentification rates are measured in 6 bins in p_T and η, given in Table 31. A complication arises from the fact that events selected in the SS region contain two electrons, in general of different p_T and η, and we do not know the charge of which of the two electrons is mismeasured. Thus, the measurement is performed in event categories that are based on p_T and η of both electrons. The categories are given in Table 32. We denote by e_1 the electron of higher p_T and by e_2 the electron of lower p_T. For example, the category that contains events in which e_1 is in the endcap and has a high p_T (**EH**), while e_2 is in the barrel and has medium p_T (**BM**) is denoted **EH_BM**. Making the assumption that the charge misidentification rates for both electrons are independent allows to combine the categories **EH_BM** and **BM_EH** into a single category. Performing such a combination for all categories reduces the number of categories from $6^2 = 36$ to 21.

p_T / $\lvert\eta\rvert$	$\lvert\eta\rvert \leq 1.479$	$1.479 < \lvert\eta\rvert < 2.5$
$10 < p_T < 25$ GeV	**BL**	**EL**
$25 < p_T < 50$ GeV	**BM**	**EM**
$p_T > 50$ GeV	**BH**	**EH**

Table 31: Bins in p_T and η in which the electron charge misidentification rate is measured.

The measured ratios, $r = \frac{N_{SS}}{N_{SS}+N_{OS}}$, of $Z/\gamma^* \to ee$ events in the SS and OS regions of the 21 events categories are related to the charge misidentification rates p in the 6 bins of electron p_T and η by the set of expressions given by Eq. (13). The expressions assume that the probabilities

e_1 e_2	EH	BH	EM	BM	EL	BL
EH	**EH_EH**	**EH_BH**	**EH_EM**	**EH_BM**	**EH_EL**	**EH_BL**
BH	-	**BH_BH**	**BH_EM**	**BH_BM**	**BH_EL**	**BH_BL**
EM	-	-	**EM_EM**	**EM_BM**	**EM_EL**	**EM_BL**
BM	-	-	-	**BM_BM**	**BM_EL**	**BM_BL**
EL	-	-	-	-	**EL_EL**	**EL_BL**
BL	-	-	-	-	-	**BL_BL**

Table 32: Event categories, based on p_T and η of e_1 and e_2, used for the measurement of the electron charge misidentification rate.

for electron charge misidentification are small, i.e. $p \cdot p \ll p$.

$$
\begin{aligned}
r_{\text{BL\_BL}} &= 2\,p_{\text{BL}} & r_{\text{BH\_EH}} &= p_{\text{BH}} + p_{\text{EH}} \\
r_{\text{BM\_BL}} &= p_{\text{BM}} + p_{\text{BL}} & r_{\text{BL\_EL}} &= p_{\text{BL}} + p_{\text{EL}} \\
r_{\text{BM\_BM}} &= 2\,p_{\text{BM}} & r_{\text{BM\_EL}} &= p_{\text{BM}} + p_{\text{EL}} \\
r_{\text{BH\_BL}} &= p_{\text{BH}} + p_{\text{BL}} & r_{\text{EM\_BL}} &= p_{\text{EM}} + p_{\text{BL}} \\
r_{\text{BH\_BM}} &= p_{\text{BH}} + p_{\text{BM}} & r_{\text{BM\_EM}} &= p_{\text{BM}} + p_{\text{EM}} \\
r_{\text{BH\_BH}} &= 2\,p_{\text{BH}} & r_{\text{BH\_EL}} &= p_{\text{BH}} + p_{\text{EL}} \\
r_{\text{EL\_EL}} &= 2\,p_{\text{EL}} & r_{\text{EH\_BL}} &= p_{\text{EH}} + p_{\text{BL}} \\
r_{\text{EM\_EL}} &= p_{\text{EM}} + p_{\text{EL}} & r_{\text{BH\_EM}} &= p_{\text{BH}} + p_{\text{EM}} \\
r_{\text{EM\_EM}} &= 2\,p_{\text{EM}} & r_{\text{EH\_BM}} &= p_{\text{EH}} + p_{\text{BM}} \\
r_{\text{EH\_EL}} &= p_{\text{EH}} + p_{\text{EL}} & r_{\text{BH\_EH}} &= p_{\text{BH}} + p_{\text{EH}} \\
r_{\text{EH\_EM}} &= p_{\text{EH}} + p_{\text{EM}}
\end{aligned}
\tag{13}
$$

The electron charge misidentification rates p in the 6 bins of p_T and η, and the corresponding uncertainties, are obtained from the ratios r measured in the 21 event categories by using matrix algebra [113] to solve the overconstrained system of linear equations given by Eq. (13).

The ratios r on the left-hand-side of Eq. (13) are determined by maximum likelihood fits of the distributions in mass, m_{ee}, if the electron pairs within the range $60 < m_{ee} < 120$ GeV. The mass distributions for different event categories are fitted independently. In each category, the SS and OS regions are fitted simultaneously. Systematic uncertainties are represented by nuisance parameters, which are correlated between the SS (low statistics) and OS (high statistics) regions. Three different approaches were tried to perform the fit.

The first approach uses analytic functions to model the $Z/\gamma^* \to ee$ signal and to model the backgrounds in the SS and OS regions, following Refs. [53, 76, 84]. The $Z/\gamma^* \to ee$ signal is modeled by a convolution of a crystal-ball and a Breit-Wigner function. The `CMSShape` function is used to model the background. It consists of an exponential tail at high mass and a sigmoid function, to model threshold effects, at low mass. The advantage of the analytic approach is that it does not require high statistics MC samples. Its disadvantage is that it relies on the assumption that the signal and background distributions in each of the 21 event categories are well modeled by the chosen analytic functions.

It turns-out, however, that the analytic functions do not model well the high and low mass tails of the m_{ee} distribution for signal and background, causing the fit to be biased. To mitigate the issue, a "hybrid" approach of analytic and simulation based modeling was attempted, in which an analytic function was used to model the background in the SS region, which suffers the most

Category	r
BL_BL	0.047 ± 0.022
BM_BL	0.037 ± 0.004
BM_BM	0.016 ± 0.001
BH_BL	0.016 ± 0.009
BH_BM	0.020 ± 0.001
BH_BH	0.026 ± 0.014
EL_EL	0.000 ± 0.000
EM_EL	0.000 ± 0.000
EM_EM	0.181 ± 0.008
EH_EL	0.340 ± 0.070
EH_EM	0.191 ± 0.009

Category	r
EH_EH	0.337 ± 0.037
BL_EL	0.028 ± 0.017
BM_EL	0.042 ± 0.006
EM_BL	0.069 ± 0.007
BM_EM	0.078 ± 0.003
BH_EL	0.046 ± 0.012
EH_BL	0.198 ± 0.018
BH_EM	0.088 ± 0.007
EH_BM	0.119 ± 0.007
BH_EH	0.204 ± 0.022

Table 33: Ratios r measured in the 21 event categories with the fit approach based on template histograms obtained from the MC simulation.

from insufficient MC statistics, while template histograms obtained from the simulation were used to model the $Z/\gamma^* \to ee$ signal and the background in the OS region. Comparing data to the sum of fitted signal plus background shapes, we still find a bias for some of the 21 event categories, however.

We find that the best performing approach, with the least bias, is to use template histograms obtained from the simulation to model the signal and background shapes in the SS and OS regions, even though the MC statistics in some of the 21 event categories is severely limited. Systematic uncertainties on the electron energy scale (eES) and resolution (eER) are represented by nuisance parameters in the fit. The eES uncertainty amounts to 1% for electrons in the barrel and to 2.5% for electrons in the endcap. The eER is allowed to vary from 0.75 to 1.25 times the nominal eER, obtained from the simulation. Statistical uncertainties are taken into account by the technique described in Refs. [95, 96]. The fits are performed using the `combine` tool [99] developed by the Higgs combination group.

The electron charge misidentification rates measured with different approaches differ by up to 30% with respect to each other and with respect to the MC truth in a closure test that we performed. We take the differences into account as systematic uncertainty when applying the measured charge misidentification rates to estimate the charge "flip" background in the in the $2\ell ss$ and $2\ell ss + 1\tau_h$ categories (*cf.* Sections 8.2 and 9).

The ratios r measured in the 21 categories with the template histogram approach are given in Table 33.

For illustration, the m_{ee} distributions observed in data are compared to the results of the fit for the SS and OS regions of the **BM_BM**, **EM_BM**, and **EM_EM** categories in Fig. **??**.

The electron charge misidentification rates measured in the 6 bins of p_T and η, obtained by solving Eq. (13) for the numbers shown in Table 33, are given in Table 34.

H.2 Muons

The muon charge misidentification rates have been measured in data recorded in 2016 and found to be negligible [114].

		$15 \leq p_{\mathrm{T}} < 25$ GeV	$25 \leq p_{\mathrm{T}} < 50$ GeV	$p_{\mathrm{T}} \geq 50$ GeV
MC truth	$\lvert\eta\rvert < 1.479$	0.0081 ± 0.0015	0.0053 ± 0.0003	0.0071 ± 0.0009
	$1.479 \leq \lvert\eta\rvert < 2.5$	0.0579 ± 0.0068	0.0507 ± 0.0016	0.0915 ± 0.0061
Pseudodata	$\lvert\eta\rvert < 1.479$	0.0009 ± 0.0024	0.0096 ± 0.0027	0.0132 ± 0.0019
	$1.479 \leq \lvert\eta\rvert < 2.5$	0.0095 ± 0.0045	0.0507 ± 0.0028	0.0907 ± 0.0067
Data	$\lvert\eta\rvert < 1.479$	0.0134 ± 0.0041	0.0224 ± 0.0041	0.0228 ± 0.0053
	$1.479 \leq \lvert\eta\rvert < 2.5$	0.0199 ± 0.0070	0.0560 ± 0.0041	0.1387 ± 0.0087

Table 34: Charge misidentification rates (in units of percent) for electrons of different p_{T} and η, measured with the fit approach based on template histograms obtained from the MC simulation. The difference between MC truth and pseudodata quantifies the bias of the fit model. The m_{ee} distribution that is used as input in the pseudodata case is constructed by setting the "data" in each bin to a random variable, drawn from a Poisson distribution the average of which is given by the sum of $\mathrm{Z}/\gamma^{*} \rightarrow \mathrm{ee}$ signal plus background expected in that bin.

I Background control regions

The modeling of irreducible background contributions in the SR, which are obtained from the MC simulation, is validated in three CRs, targeting one of the dominant irreducible background processes $t\bar{t}W$, $t\bar{t}Z$, and WZ+jets each. Events in these CRs are selected as described in Sections I.1, I.2, and I.3.

The number of events observed in the CRs is compared to the expected yields of irreducible and reducible backgrounds and of the $t\bar{t}H$ signal in Tables 40 and 41. Background contributions to the control regions, which arise from the misidentification of a non-prompt lepton or hadron as prompt lepton or from the mismeasurement of the lepton charge are estimated from data, using the procedures described in Sections 8.1 and 8.2. The yield expected for the $t\bar{t}H$ signal is given for the SM cross section.

Table 35: Number of events events selected in CRs enriched in the contributions of $t\bar{t}W$, $t\bar{t}Z$, and WZ+jets backgrounds, respectively. Quoted uncertainties represent statistical uncertainties only.

Process	WZ+jets control region
$t\bar{t}Z$	$XX.XX \pm X.XX$
$t\bar{t}W$	$XX.XX \pm X.XX$
WZ+jets	$XX.XX \pm X.XX$
Other electroweak	$XX.XX \pm X.XX$
Fake	$XX.XX \pm X.XX$
Charge flip	$XX.XX \pm X.XX$
Conversions	$XX.XX \pm X.XX$
Other	$XX.XX \pm X.XX$
$t\bar{t}H$	$XX.XX \pm X.XX$
SM expectation	$XX.XX \pm X.XX$
Observed data	XX

Distributions in selected kinematic observables are shown in Figs. 18, 19, and 20.

Good agreement between data and simulation is observed in modeling the yields of events selected in the $t\bar{t}W$, $t\bar{t}Z$, and WZ+jets CRs as well as in modeling the shape of distributions.

I.1 $t\bar{t}W$

Events in the CR enriched in the contribution of the $t\bar{t}W$ background are selected by applying the event selection criteria for the SR of the $2\ell ss$ category, described in Section 5.4, but requiring exactly three jets of $p_T > 25$ GeV and $|\eta| < 2.4$ for events selected in the CR.

I.2 $t\bar{t}Z$

Events in the CR enriched in the contribution of the $t\bar{t}Z$ background are selected by applying the event selection criteria for the SR of the 3ℓ category, described in Section 5.6, except that events selected in the CR are required to contain a pair of fakeable leptons of the same flavor, opposite charge, and mass $|m_{\ell\ell} - m_Z| < 10$ GeV.

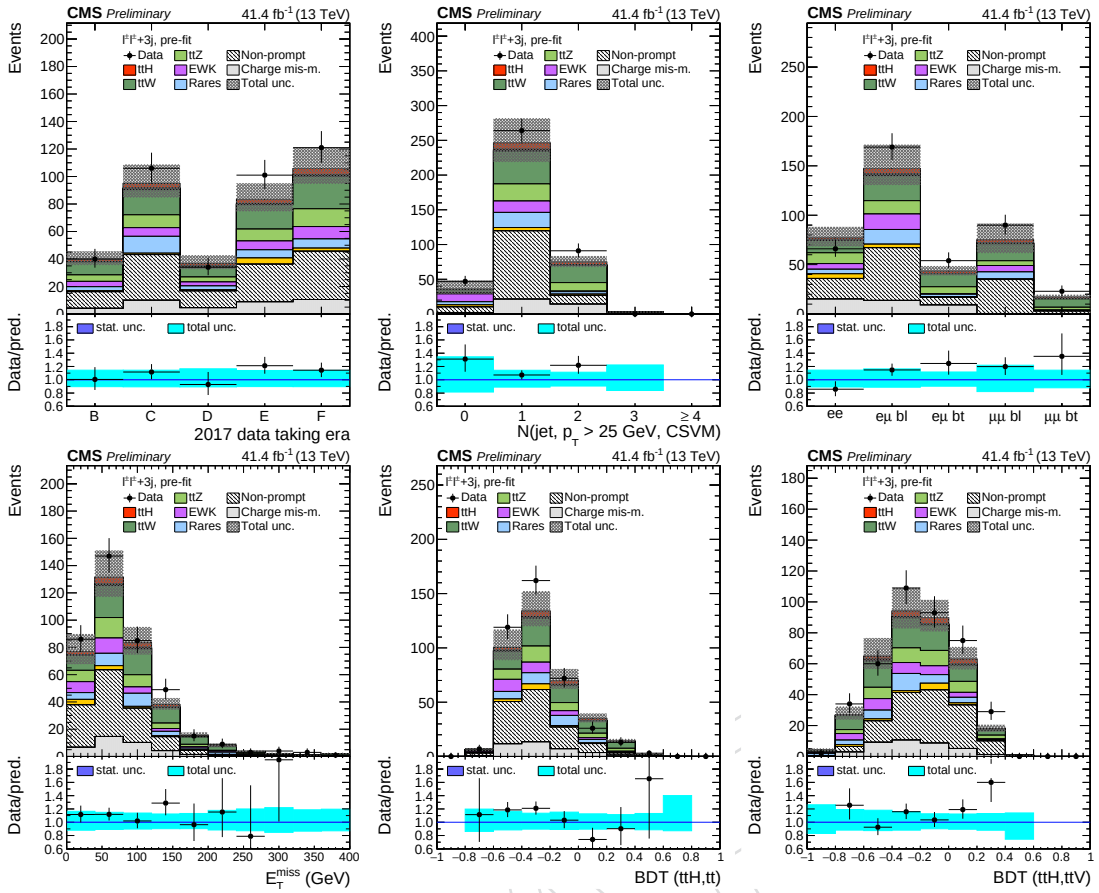


Figure 18: Distributions in selected kinematic observables in events selected in the CR enriched in the contribution of $t\bar{t}W$ background.

I.3 WZ+jets

Events in the CR enriched in the contribution of the $t\bar{t}Z$ background are selected by applying the event selection criteria for the SR of the 3ℓ category, but requiring the events selected in the CR to contain a pair of fakeable leptons of the same flavor, opposite charge, and mass $|m_{\ell\ell} - m_{Z}| < 10$ GeV and no jets of $p_{T} > 25$ GeV, $|\eta| < 2.4$, and passing the medium WP of the DeepCSV discriminant.

I.4 "Fake" background control plots

I.5 $1\ell + 2\text{ss}\tau_{h}$

Events in this control region are selected similarly as $1\ell + 2\tau_{h}$ signal region except the two τ_{h} are required to be of same charge. The "fake" background in this CR is estimated with FF method described in 8.1 from an application region that is the same as AR used to estimate "fake" background in $1\ell + 2\tau_{h}$ SR, but with the requirement that the two τ_{h} are same sign instead of opposite sign.

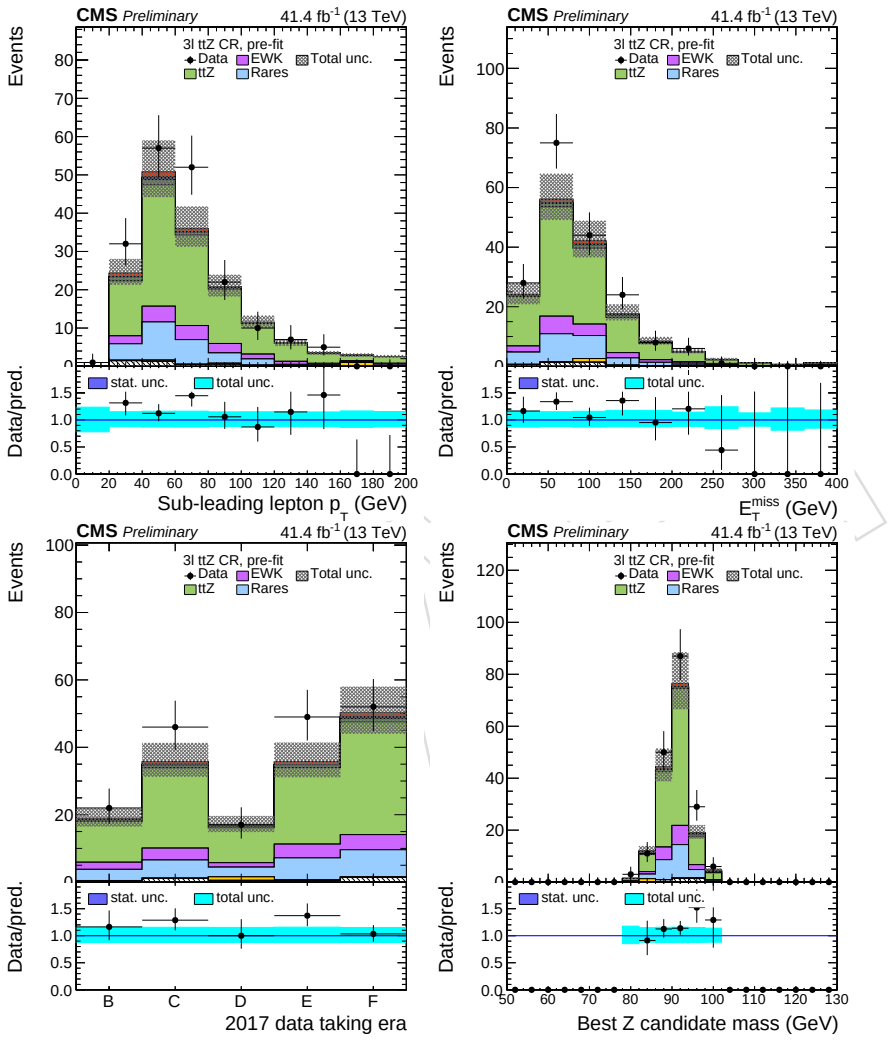


Figure 19: Distributions in selected kinematic observables in events selected in the CR enriched in the contribution of $t\bar{t}Z$ background.

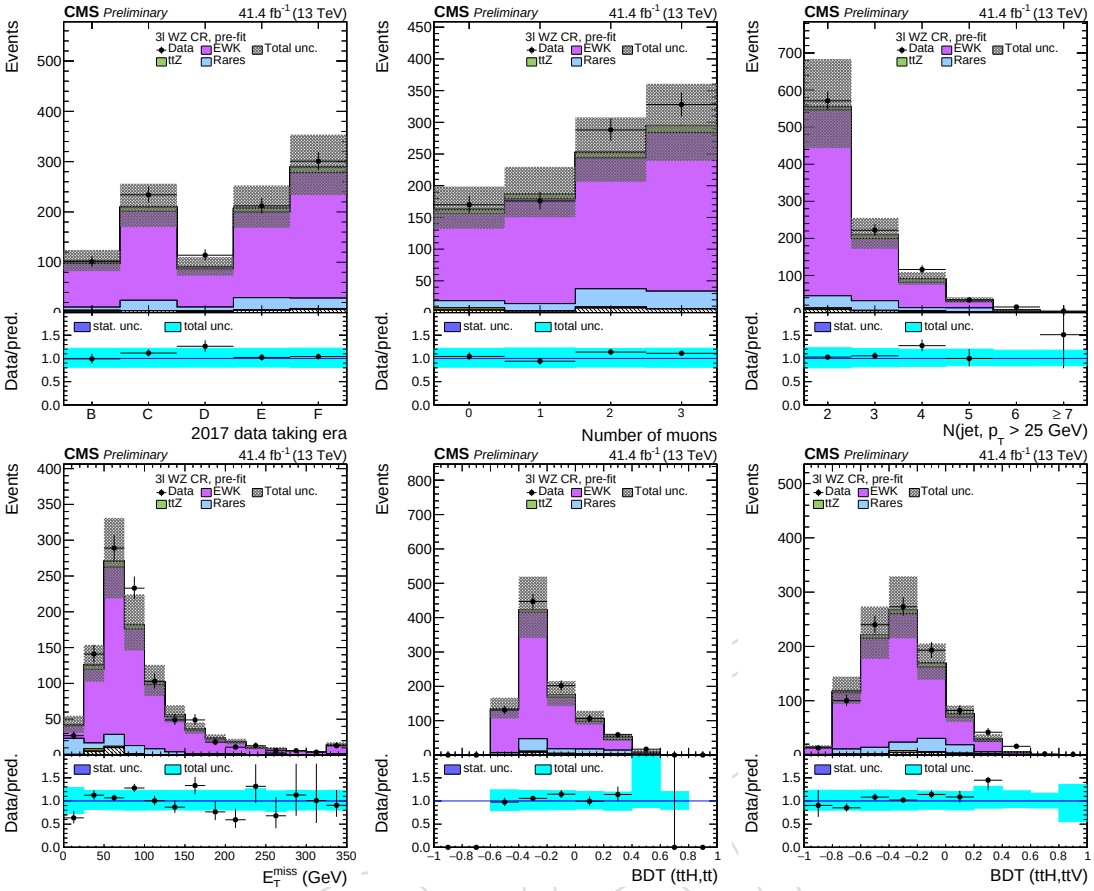


Figure 20: Distributions in selected kinematic observables in events selected in the CR enriched in the contribution of WZ+jets background.

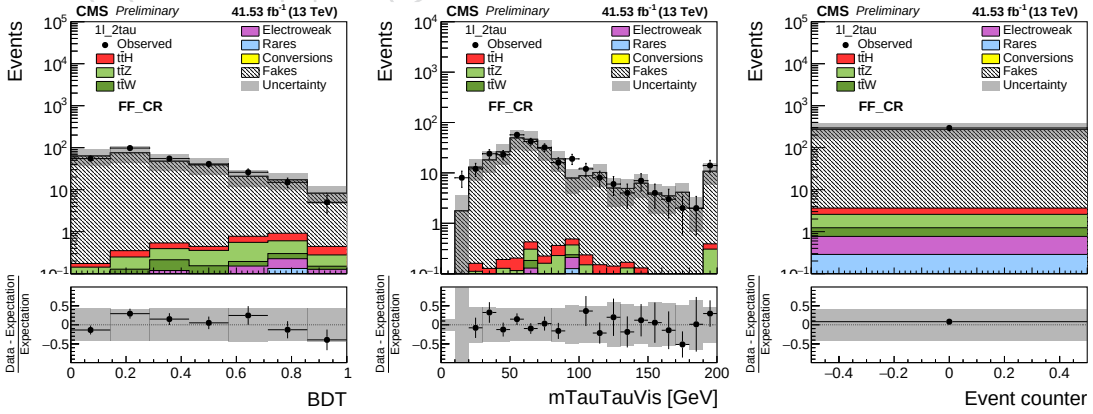
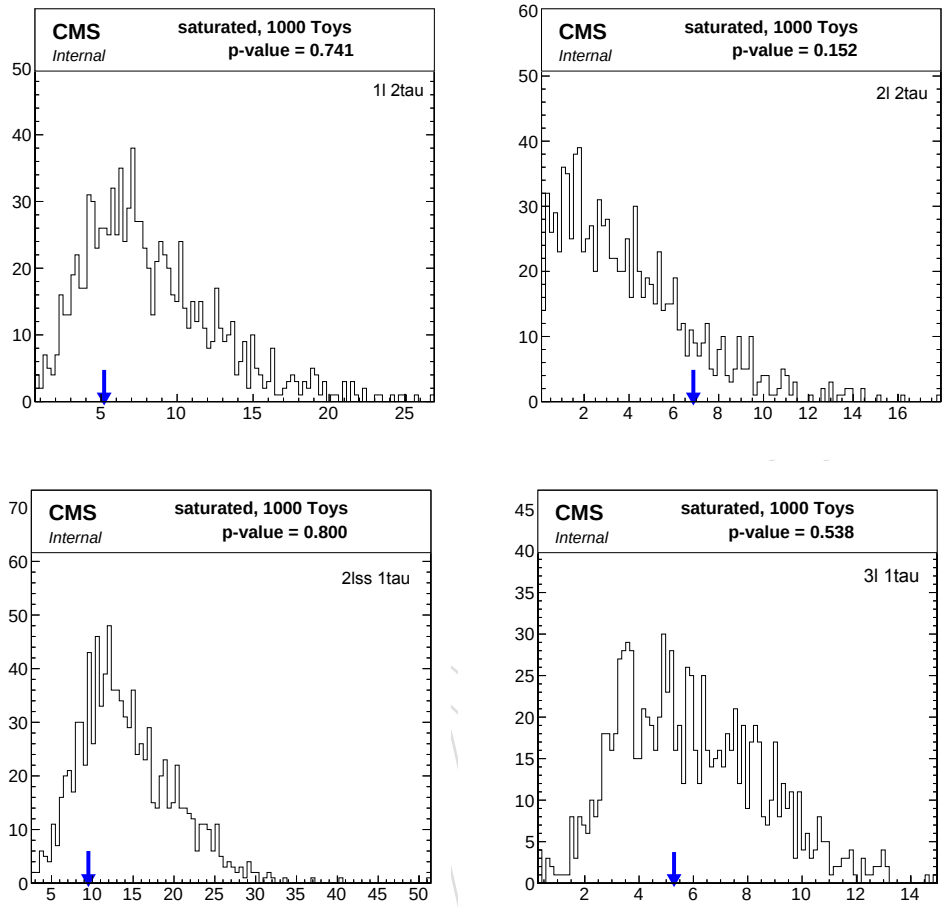


Figure 21: Distributions in the BDT output and visible mass of the τ_h pair, as well as the event count in $1\ell + 2ss\tau_h$ control region.

J Fit diagnosis on categories with taus

Figure 22: Goodness of fit for the fit on the individual BDT input variables for the$1\ell + 2\tau_h$ (top left), $2\ell ss + 1\tau_h$ (top right), $2\ell + 2\tau_h$ (bottom left), and $3\ell + 1\tau_h$ (bottom right) categories.



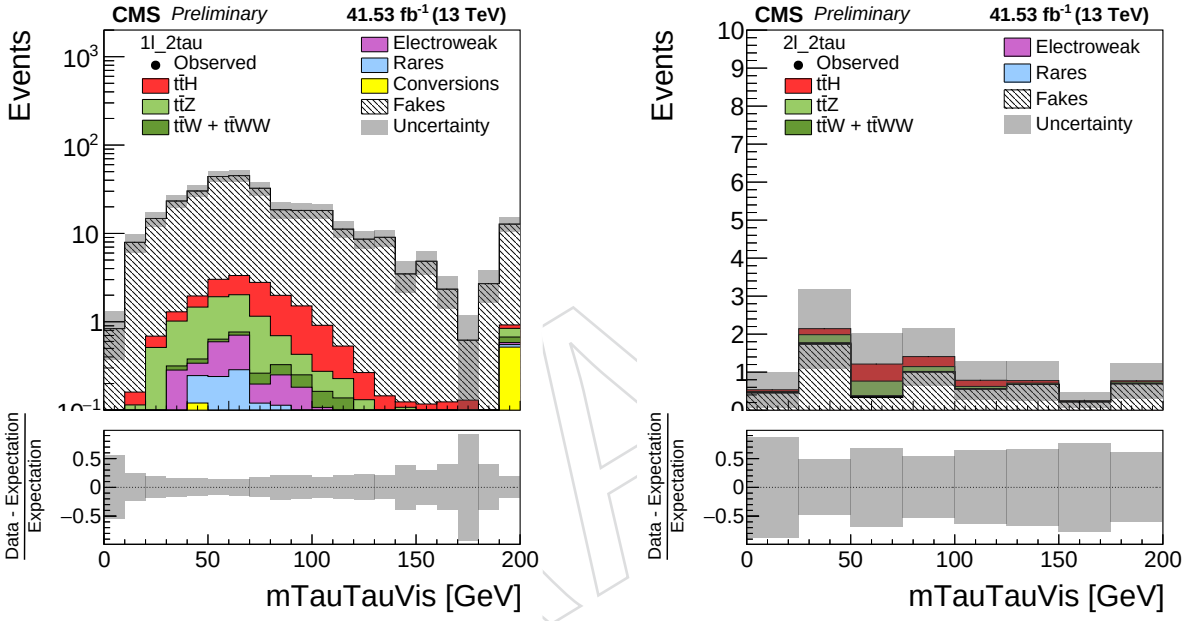
K Alternative on categories with taus

We make an alternative signal extraction using the reconstructed visible mass of the H in each of the categigoes. The H visible mass for each one of the categories with taus are:

- $1\ell + 2\tau_h$, Mass of leading τ_h + trailing τ_h
- $2\ell ss + 1\tau_h$, Mass of leading ℓ+ leading τ_h
- $2\ell + 2\tau_h$, Mass of leading τ_h + trailing τ_h
- $3\ell + 1\tau_h$, Mass of leading ℓ+ leading τ_h for the collection of ℓ's that has opposite charge to the τ_h

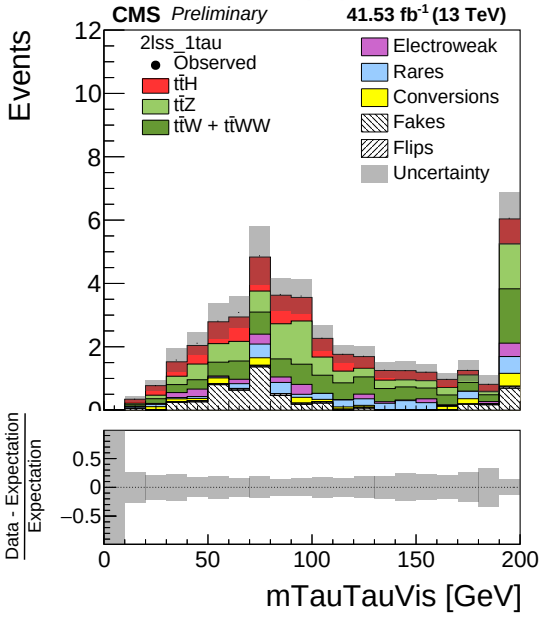
On Fig. 23 we show the post-fit plots for the fits and on Tab. 36 the resultant expected limits.

Figure 23: Distributions in the discriminating observable used for alternative signal extraction (the H visible mass) in (a) the $1\ell + 2\tau_h$ category, (b) the $2\ell + 2\tau_h$ category, (c) the $2\ell ss + 1\tau_h$ and (d) $3\ell + 1\tau_h$ compared to the SM expectation for the $t\bar{t}H$ signal and for background processes. The distributions are obtained via. category-wise ML fit to asimov dataset (made under the SM signal + SM background hypothesis with signal strength (μ) fixed to 1 times the SM expectation). NOTE: data is still blinded!

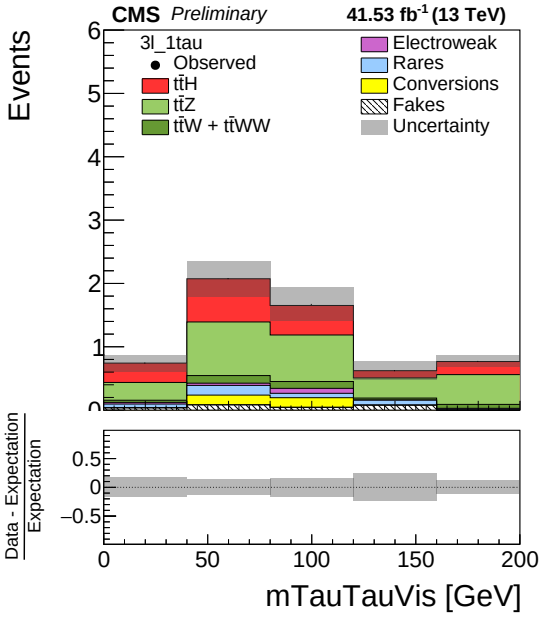


(a)

(b)



(c)



(d)

Figure 24: Goodness of fit on the fit of the discriminating observable used for alternative signal extraction (the H visible mass) in (a) the $1\ell + 2\tau_h$ category, (b) the $2\ell + 2\tau_h$ category, (c) the $2\ell ss + 1\tau_h$ and (d) $3\ell + 1\tau_h$ compared to the SM expectation for the $t\bar{t}H$ signal and for background processes.

Table 36: Expected limits on the categories with taus when the signal is extracted by the H visible mass.

Category	Central	-2σ	-1σ	1σ	2σ	Central without systematics
1l_2tau	7.19	3.82	5.11	10.37	14.63	3.39
2lss_1tau	1.56	0.79	1.08	2.33	3.38	1.27
2l_2tau	5.42	2.52	3.56	8.62	13.41	4.61
3l_1tau	3.16	1.45	2.07	5.04	7.85	2.88

L Yields prefit

In this section we show the pre-fit yields for the different processes on the different categories used for the fit.

Table 37: The selected number of events in the $2\ell ss$ (sub)categories. The rates are adjusted using the maximum-likelihood fit (pre-fit). Smaller uncertainties than 0.01 are denoted as 0.00.

Category	$2\ell ss$									
b-tag	no req.		Loose				Tight			
Leptons	ee		$e\mu$		$\mu\mu$		$e\mu$		$\mu\mu$	
Charge	$-$	$+$	$-$	$+$	$-$	$+$	$-$	$+$	$-$	$+$
$t\bar{t}H$, $H \to ZZ$	0.20 ± 0.03	0.17 ± 0.02	0.30 ± 0.04	0.21 ± 0.02	0.14 ± 0.02	0.18 ± 0.03	0.32 ± 0.04	0.21 ± 0.04	0.10 ± 0.02	0.10 ± 0.02
$t\bar{t}H$, $H \to WW$	3.58 ± 0.50	3.79 ± 0.49	7.27 ± 0.80	6.84 ± 0.79	4.76 ± 0.53	4.90 ± 0.55	5.07 ± 0.75	5.60 ± 0.84	3.65 ± 0.52	3.22 ± 0.46
$t\bar{t}H$, $H \to \tau\tau$	0.85 ± 0.12	0.82 ± 0.11	1.29 ± 0.14	1.38 ± 0.16	0.78 ± 0.08	0.88 ± 0.09	1.08 ± 0.15	0.91 ± 0.12	0.65 ± 0.09	0.56 ± 0.08
$t\bar{t}H$, $H \to \mu\mu$	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05
$t\bar{t}H$, $H \to Z\gamma$	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05
ttH (sum)	4.64 ± 0.52	4.81 ± 0.50	8.91 ± 0.81	8.47 ± 0.80	5.70 ± 0.54	5.98 ± 0.56	6.52 ± 0.77	6.75 ± 0.85	4.41 ± 0.53	3.89 ± 0.47
$t\bar{t}W$	9.32 ± 1.50	14.77 ± 2.39	16.00 ± 2.40	28.49 ± 4.27	10.00 ± 1.52	17.61 ± 2.64	11.13 ± 1.88	22.55 ± 3.86	7.98 ± 1.37	13.18 ± 2.21
$t\bar{t}WW$	0.68 ± 0.09	0.52 ± 0.07	1.36 ± 0.18	0.99 ± 0.13	0.67 ± 0.09	0.68 ± 0.10	1.03 ± 0.15	0.86 ± 0.13	0.51 ± 0.07	0.58 ± 0.10
ttW + ttWW	10.00 ± 1.50	15.29 ± 2.39	17.36 ± 2.41	29.48 ± 4.27	10.67 ± 1.53	18.29 ± 2.65	12.16 ± 1.89	23.41 ± 3.86	8.50 ± 1.38	13.75 ± 2.21
tHq $H \to \tau\tau$	< 0.05	0.09 ± 0.02	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05
tHW $H \to \tau\tau$	< 0.05	< 0.05	< 0.05	0.12 ± 0.02	< 0.05	0.13 ± 0.03	< 0.05	< 0.05	< 0.05	< 0.05
tHq $H \to WW$	0.06 ± 0.01	< 0.05	0.10 ± 0.02	0.15 ± 0.02	0.07 ± 0.01	0.14 ± 0.01	0.11 ± 0.02	< 0.05	< 0.05	< 0.05
tHW $H \to WW$	< 0.05	< 0.05	0.19 ± 0.04	0.12 ± 0.01	0.24 ± 0.07	0.14 ± 0.03	< 0.05	0.23 ± 0.05	< 0.05	0.13 ± 0.03
tHq $H \to ZZ$	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05
tHW $H \to ZZ$	< 0.05	0.10 ± 0.03	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05
tH (sum)	0.12 ± 0.02	0.25 ± 0.03	0.31 ± 0.05	0.42 ± 0.03	0.32 ± 0.07	0.46 ± 0.05	0.17 ± 0.02	0.25 ± 0.05	0.05 ± 0.01	0.13 ± 0.03
ZZ	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05
WZ	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05
WZ + ZZ	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05
$t\bar{t}Z$	11.33 ± 1.81	11.11 ± 1.73	12.75 ± 1.78	14.37 ± 1.91	4.61 ± 0.59	5.29 ± 0.73	9.32 ± 1.55	10.90 ± 1.76	3.71 ± 0.57	3.66 ± 0.55
Misidentified	9.82 ± 3.32	12.58 ± 3.81	29.71 ± 7.33	31.21 ± 7.85	20.20 ± 7.16	17.14 ± 6.43	6.08 ± 2.16	6.88 ± 2.23	2.91 ± 1.37	3.72 ± 1.64
Conversions	1.73 ± 0.53	4.27 ± 2.58	2.69 ± 0.80	6.34 ± 3.41	< 0.05	< 0.05	1.03 ± 0.35	1.19 ± 0.36	< 0.05	< 0.05
signal flip	4.82 ± 1.28	4.89 ± 1.30	3.95 ± 1.05	4.03 ± 1.07	< 0.05	< 0.05	3.30 ± 0.88	3.27 ± 0.87	< 0.05	< 0.05
Other	1.59 ± 0.69	2.77 ± 1.22	5.32 ± 2.42	6.52 ± 2.88	1.97 ± 0.89	3.69 ± 1.62	2.43 ± 1.06	2.64 ± 1.15	1.19 ± 0.54	1.94 ± 0.84
SM expectation	44.05 ± 4.38	55.97 ± 5.77	81.00 ± 8.42	100.84 ± 10.26	43.45 ± 7.42	50.86 ± 7.20	41.02 ± 3.64	55.29 ± 5.08	20.78 ± 2.16	27.09 ± 2.97
Observed data	54	67	86	108	41	79	50	70	32	27

Table 38: The selected number of events in the 3ℓ and 4ℓ categories. The rates are adjusted using the maximum-likelihood fit (pre-fit). Smaller uncertainties than 0.01 are denoted as 0.00.

Category	3ℓ				4ℓ
b-tag	Loose		Tight		no req.
Total charge	$-$	$+$	$-$	$+$	no req.
$t\bar{t}H$, $H \to ZZ$	0.47 ± 0.12	0.17 ± 0.02	0.21 ± 0.03	0.19 ± 0.04	0.09 ± 0.01
$t\bar{t}H$, $H \to WW$	7.87 ± 1.02	7.03 ± 0.92	3.24 ± 0.50	2.84 ± 0.49	0.65 ± 0.09
$t\bar{t}H$, $H \to \tau\tau$	1.34 ± 0.22	0.54 ± 0.09	0.72 ± 0.11	1.00 ± 0.13	0.19 ± 0.03
$t\bar{t}H$, $H \to \mu\mu$	0.09 ± 0.02	0.24 ± 0.03	0.07 ± 0.02	0.17 ± 0.03	0.07 ± 0.01
$t\bar{t}H$, $H \to Z\gamma$	< 0.05	< 0.05	< 0.05	0.04 ± 0.01	< 0.05
ttH (sum)	9.78 ± 1.05	7.98 ± 0.93	4.24 ± 0.51	4.24 ± 0.51	0.99 ± 0.10
$t\bar{t}W$	9.04 ± 1.56	14.42 ± 2.36	4.31 ± 0.89	6.50 ± 1.44	< 0.05
$t\bar{t}WW$	0.88 ± 0.11	0.97 ± 0.13	0.57 ± 0.09	0.52 ± 0.08	0.09 ± 0.01
ttW + ttWW	9.91 ± 1.57	15.39 ± 2.36	4.89 ± 0.89	7.01 ± 1.44	0.09 ± 0.01
tHq $H \to \tau\tau$	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05
tHW $H \to \tau\tau$	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05
tHq $H \to WW$	0.08 ± 0.01	0.28 ± 0.04	0.34 ± 0.13	< 0.05	< 0.05
tHW $H \to WW$	0.06 ± 0.01	0.04 ± 0.01	< 0.05	< 0.05	< 0.05
tHq $H \to ZZ$	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05
tHW $H \to ZZ$	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05
tH (sum)	0.16 ± 0.02	0.36 ± 0.04	0.36 ± 0.13	0.09 ± 0.01	< 0.05
ZZ	0.21 ± 0.03	0.18 ± 0.04	< 0.05	< 0.05	0.09 ± 0.05
WZ	3.96 ± 0.71	7.41 ± 1.30	0.24 ± 0.06	0.45 ± 0.06	< 0.05
WZ + ZZ	4.17 ± 0.71	7.59 ± 1.30	0.24 ± 0.06	0.45 ± 0.06	0.09 ± 0.05
$t\bar{t}Z$	16.37 ± 2.59	13.65 ± 2.32	9.25 ± 1.79	8.28 ± 1.55	2.46 ± 0.37
Misidentified	13.24 ± 4.11	11.39 ± 4.07	1.79 ± 0.73	2.37 ± 0.96	< 0.05
Conversions	1.84 ± 0.79	2.21 ± 0.80	2.69 ± 1.11	0.96 ± 0.36	< 0.05
signal flip	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05
Other	2.70 ± 1.16	4.33 ± 1.89	1.77 ± 0.77	2.05 ± 0.89	0.10 ± 0.07
SM expectation	58.17 ± 5.44	62.90 ± 5.86	25.23 ± 2.58	25.45 ± 2.57	3.78 ± 0.39
Observed data	55	84	23	33	6

Table 39: The selected number of events in the $1\ell + 2\tau_h$, $2\ell ss + 1\tau_h$, $2\ell + 2\tau_h$, and $3\ell + 1\tau_h$ categories. The rates are adjusted using the maximum-likelihood fit (pre-fit). Smaller uncertainties than 0.01 are denoted as 0.00.

Process	$1\ell + 2\tau_h$	$2\ell + 2\tau_h$	$3\ell + 1\tau_h$	$2\ell ss + 1\tau_h$
$t\bar{t}H$, $H \to ZZ$	0.10 ± 0.02	< 0.05	0.05 ± 0.01	0.09 ± 0.01
$t\bar{t}H$, $H \to WW$	0.46 ± 0.07	0.09 ± 0.03	0.54 ± 0.08	2.34 ± 0.30
$t\bar{t}H$, $H \to \tau\tau$	8.27 ± 1.24	1.08 ± 0.20	0.74 ± 0.11	4.52 ± 0.58
$t\bar{t}H$, $H \to \mu\mu$	< 0.05	< 0.05	< 0.05	< 0.05
$t\bar{t}H$, $H \to Z\gamma$	< 0.05	< 0.05	< 0.05	< 0.05
ttH (sum)	8.83 ± 1.24	1.18 ± 0.20	1.36 ± 0.14	6.96 ± 0.65
$t\bar{t}W$	0.73 ± 0.14	< 0.05	0.25 ± 0.07	8.49 ± 1.49
$t\bar{t}WW$	0.07 ± 0.01	< 0.05	0.10 ± 0.01	0.49 ± 0.07
ttW + ttWW	0.80 ± 0.14	< 0.05	0.35 ± 0.08	8.98 ± 1.49
tHq	< 0.05	< 0.05	< 0.05	0.32 ± 0.03
tHW	0.23 ± 0.06	< 0.05	< 0.05	0.13 ± 0.01
tH (sum)	0.26 ± 0.06	< 0.05	< 0.05	0.45 ± 0.03
$WZ + ZZ$	1.58 ± 0.73	0.05 ± 0.02	0.13 ± 0.03	2.17 ± 1.46
$t\bar{t}Z$	7.07 ± 1.24	0.94 ± 0.19	2.69 ± 0.33	9.21 ± 0.95
Misidentified	287.30 ± 120.82	5.27 ± 2.56	0.22 ± 0.17	5.56 ± 2.38
Conversions	0.27 ± 0.13	< 0.05	< 0.05	< 0.05
signal flip	< 0.05	< 0.05	< 0.05	< 0.05
Other	1.24 ± 0.54	0.06 ± 0.03	0.34 ± 0.16	3.33 ± 1.12
SM expectation	307.35 ± 120.83	7.52 ± 2.58	5.11 ± 0.43	36.66 ± 3.55
Observed data	324	7	4	53

Table 40: The selected number of events in the $2\ell ss$ (3j) control regions. The rates are adjusted using the maximum-likelihood fit (post-fit). Smaller uncertainties than 0.01 are denoted as 0.00.

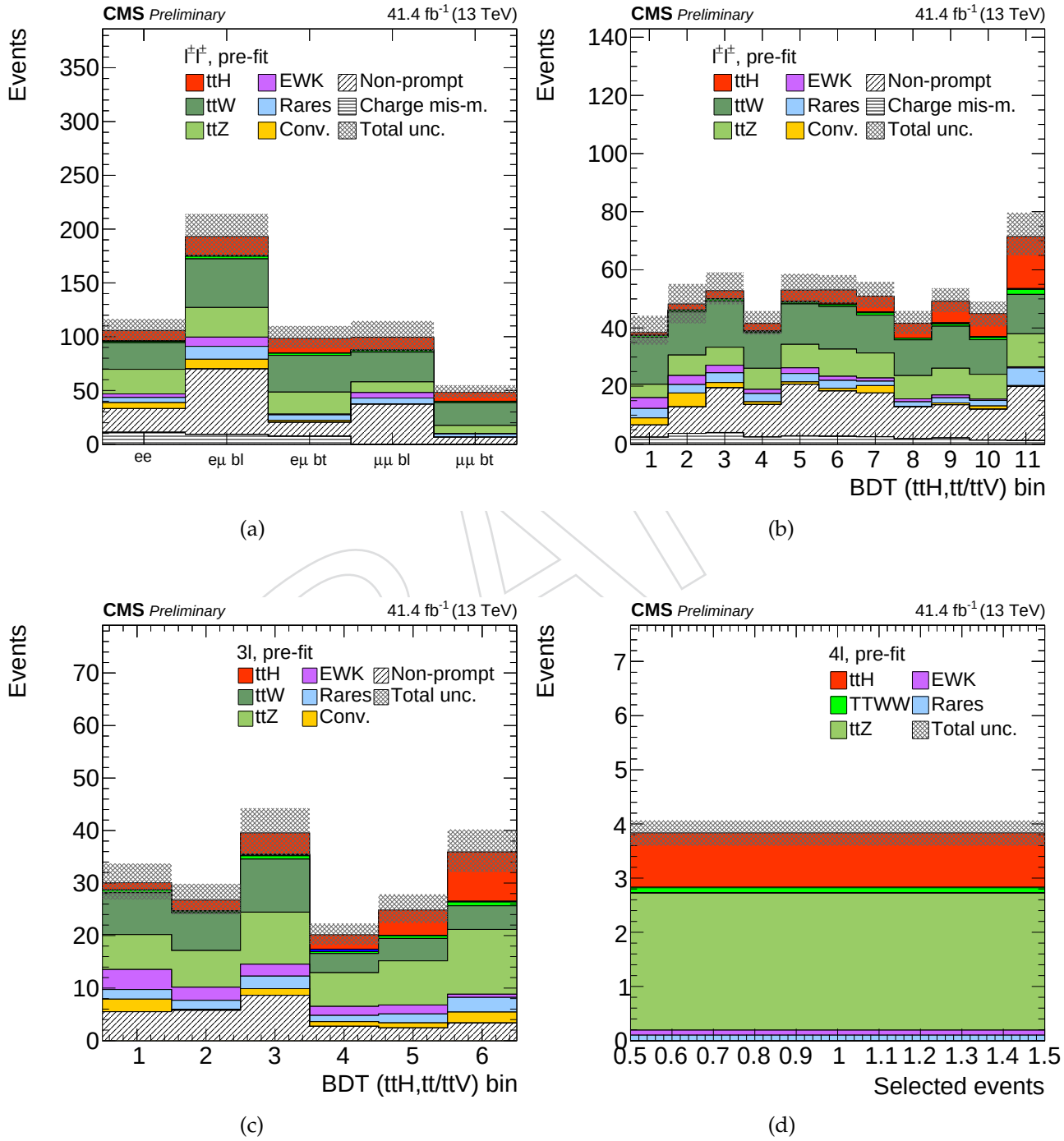
Category	$2\ell ss$									
b-tag	no req.		Loose				Tight			
Leptons	ee		$e\mu$		$\mu\mu$		$e\mu$		$\mu\mu$	
Charge	$-$	$+$	$-$	$+$	$-$	$+$	$-$	$+$	$-$	$+$
$t\bar{t}H$, H $\to$ ZZ	< 0.05	< 0.05	0.05 ± 0.01	< 0.05	0.05 ± 0.01	< 0.05	< 0.05	0.05 ± 0.01	< 0.05	< 0.05
$t\bar{t}H$, H $\to$ WW	0.76 ± 0.11	0.82 ± 0.10	1.97 ± 0.23	2.00 ± 0.23	1.36 ± 0.18	1.33 ± 0.15	0.86 ± 0.12	0.96 ± 0.14	0.60 ± 0.08	0.50 ± 0.07
$t\bar{t}H$, H $\to \tau\tau$	0.23 ± 0.04	0.24 ± 0.03	0.47 ± 0.06	0.44 ± 0.05	0.26 ± 0.03	0.30 ± 0.03	0.26 ± 0.03	0.14 ± 0.02	0.19 ± 0.03	0.21 ± 0.03
$t\bar{t}H$, H $\to \mu\mu$	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05
$t\bar{t}H$, H $\to Z\gamma$	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05
ttH (sum)	1.04 ± 0.12	1.09 ± 0.11	2.52 ± 0.24	2.50 ± 0.24	1.68 ± 0.18	1.68 ± 0.15	1.15 ± 0.12	1.15 ± 0.14	0.82 ± 0.09	0.73 ± 0.07
$t\bar{t}W$	4.37 ± 0.69	7.83 ± 1.23	8.72 ± 1.29	16.32 ± 2.32	5.59 ± 0.84	11.63 ± 1.65	4.53 ± 0.76	8.46 ± 1.38	2.89 ± 0.49	5.63 ± 0.91
$t\bar{t}WW$	0.04 ± 0.01	0.17 ± 0.02	0.27 ± 0.04	0.17 ± 0.02	0.17 ± 0.04	0.19 ± 0.03	0.07 ± 0.01	0.10 ± 0.02	0.05 ± 0.01	< 0.05
ttW + ttWW	4.41 ± 0.69	7.99 ± 1.23	8.99 ± 1.29	16.49 ± 2.32	5.76 ± 0.84	11.82 ± 1.65	4.61 ± 0.76	8.57 ± 1.38	2.94 ± 0.49	5.67 ± 0.91
tHq H $\to \tau\tau$	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05
tHW H $\to \tau\tau$	< 0.05	0.04 ± 0.01	< 0.05	0.19 ± 0.04	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05
tHq H $\to$ WW	< 0.05	< 0.05	0.16 ± 0.02	0.18 ± 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05
tHW H $\to$ WW	< 0.05	< 0.05	0.40 ± 0.09	0.55 ± 0.08	< 0.05	0.18 ± 0.05	< 0.05	0.11 ± 0.01	< 0.05	< 0.05
tHq H $\to$ ZZ	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05
tHW H $\to$ ZZ	< 0.05	< 0.05	0.06 ± 0.01	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05
tH (sum)	0.06 ± 0.00	0.06 ± 0.01	0.65 ± 0.09	0.93 ± 0.10	0.06 ± 0.01	0.21 ± 0.05	< 0.05	0.11 ± 0.01	< 0.05	< 0.05
ZZ	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05
WZ	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05
WZ + ZZ	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05
$t\bar{t}Z$	5.18 ± 0.87	5.40 ± 0.81	6.11 ± 0.87	7.11 ± 0.96	2.38 ± 0.31	2.38 ± 0.31	3.90 ± 0.74	3.21 ± 0.61	1.52 ± 0.23	1.35 ± 0.19
Misidentified	9.49 ± 3.10	11.52 ± 3.44	28.45 ± 6.98	25.00 ± 6.34	18.58 ± 6.83	16.72 ± 6.72	2.74 ± 1.01	4.87 ± 1.56	1.60 ± 0.70	1.57 ± 0.76
Conversions	1.59 ± 0.68	3.32 ± 2.36	1.57 ± 0.44	1.85 ± 0.91	< 0.05	< 0.05	0.59 ± 0.20	0.63 ± 0.19	< 0.05	< 0.05
signal flip	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05
Other	1.89 ± 0.83	2.60 ± 1.13	4.14 ± 1.80	6.43 ± 2.79	2.92 ± 1.32	4.47 ± 2.00	1.01 ± 0.44	1.44 ± 0.79	0.13 ± 0.06	0.58 ± 0.26
SM expectation	23.66 ± 3.47	31.98 ± 4.57	52.44 ± 7.39	60.29 ± 7.44	31.39 ± 7.02	37.29 ± 7.21	14.03 ± 1.55	19.98 ± 2.32	7.01 ± 0.89	9.95 ± 1.23
Observed data	33	33	64	105	49	41	27	27	15	8

Table 41: The selected number of events in the 3ℓ and 4ℓ control regions. The rates are adjusted using the maximum-likelihood fit (prefit-fit). Smaller uncertainties than 0.01 are denoted as 0.00.

Category	3ℓ					4ℓ
CR	Z-peak ($t\bar{t}Z$ enrich.)				WZ enrich.	ZZ enrich.
b-tag	Loose		Tight		no req.	
Total charge	$-$	$+$	$-$	$+$	no req.	
$t\bar{t}H$, $H \to ZZ$	0.84 ± 0.11	0.73 ± 0.09	0.37 ± 0.07	0.33 ± 0.05	0.23 ± 0.18	0.08 ± 0.06
$t\bar{t}H$, $H \to WW$	0.62 ± 0.08	1.04 ± 0.13	0.76 ± 0.13	0.29 ± 0.05	0.47 ± 0.37	< 0.05
$t\bar{t}H$, $H \to \tau\tau$	0.26 ± 0.03	0.19 ± 0.03	0.14 ± 0.04	< 0.05	0.06 ± 0.05	< 0.05
$t\bar{t}H$, $H \to \mu\mu$	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05
$t\bar{t}H$, $H \to Z\gamma$	0.05 ± 0.01	0.10 ± 0.02	< 0.05	< 0.05	< 0.05	< 0.05
ttH (sum)	1.77 ± 0.14	2.06 ± 0.16	1.31 ± 0.15	0.69 ± 0.07	0.79 ± 0.41	0.08 ± 0.06
$t\bar{t}W$	1.19 ± 0.19	1.87 ± 0.29	0.74 ± 0.13	1.31 ± 0.24	0.63 ± 0.49	< 0.05
$t\bar{t}WW$	0.17 ± 0.03	0.14 ± 0.09	0.14 ± 0.03	< 0.05	0.05 ± 0.03	< 0.05
ttW + ttWW	1.35 ± 0.19	2.02 ± 0.30	0.87 ± 0.13	1.33 ± 0.24	0.68 ± 0.49	< 0.05
tHq $H \to \tau\tau$	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05
tHW $H \to \tau\tau$	< 0.05	0.17 ± 0.02	< 0.05	< 0.05	< 0.05	< 0.05
tHq $H \to WW$	< 0.05	< 0.05	0.08 ± 0.01	0.34 ± 0.05	< 0.05	< 0.05
tHW $H \to WW$	< 0.05	< 0.05	< 0.05	0.05 ± 0.01	< 0.05	< 0.05
tHq $H \to ZZ$	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05
tHW $H \to ZZ$	< 0.05	< 0.05	< 0.05	< 0.05	0.06 ± 0.05	< 0.05
tH (sum)	0.05 ± 0.01	0.19 ± 0.02	0.10 ± 0.01	0.40 ± 0.05	0.08 ± 0.05	< 0.05
ZZ	2.62 ± 0.35	2.56 ± 0.43	0.37 ± 0.07	0.19 ± 0.08	49.82 ± 39.21	31.21 ± 3.56
WZ	22.80 ± 4.82	31.72 ± 6.56	1.64 ± 0.35	1.22 ± 0.28	688.45 ± 63.60	< 0.05
WZ + ZZ	25.41 ± 4.83	34.28 ± 6.57	2.01 ± 0.35	1.41 ± 0.29	738.27 ± 74.72	31.21 ± 3.56
$t\bar{t}Z$	45.19 ± 8.10	42.25 ± 7.59	36.98 ± 7.02	27.74 ± 5.55	25.05 ± 19.27	1.99 ± 1.54
Misidentified	6.01 ± 2.20	6.61 ± 2.03	1.46 ± 0.60	1.11 ± 0.46	18.90 ± 15.45	< 0.05
Conversions	0.10 ± 0.04	0.83 ± 0.28	0.98 ± 0.29	0.07 ± 0.03	4.34 ± 3.50	< 0.05
signal flip	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05
Other	18.09 ± 7.83	27.18 ± 11.65	6.13 ± 2.68	8.01 ± 3.47	56.25 ± 48.50	2.68 ± 2.33
SM expectation	97.97 ± 12.46	115.42 ± 15.52	49.85 ± 7.55	40.77 ± 6.57	844.37 ± 92.51	35.96 ± 4.52
Observed data	123	154	54	49	919	61

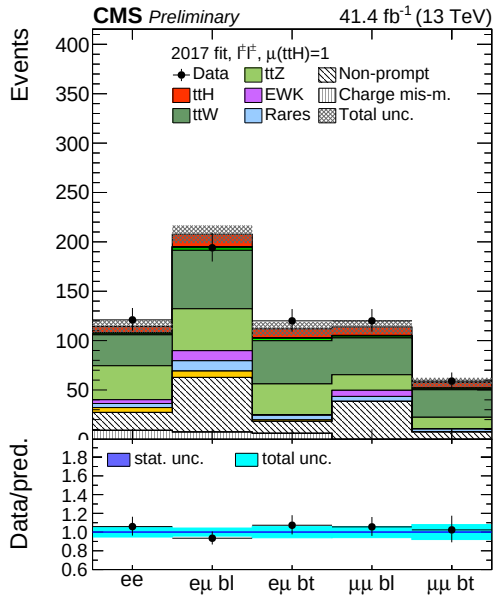
M Pre-fit Distributions

Figure 25: Distributions in the discriminating observables used for the signal extraction in the categories without taus. The SM expectations for the $t\bar{t}H$ signal and for the background processes are presented in colors. The SM (pre-fit) rates and uncertainties are used.

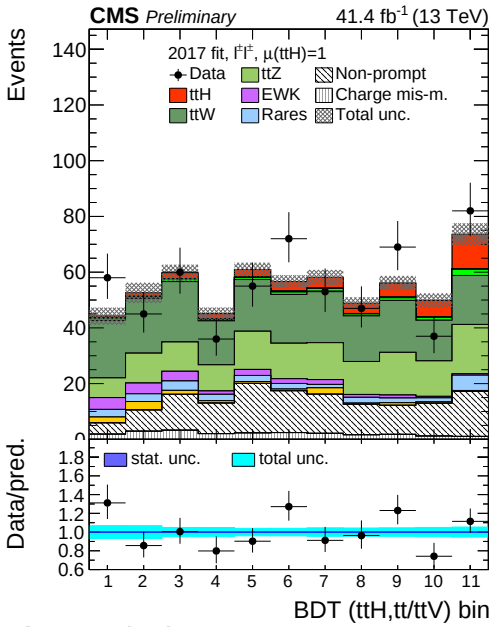


N Post-fit, fixed SM Higgs Distributions

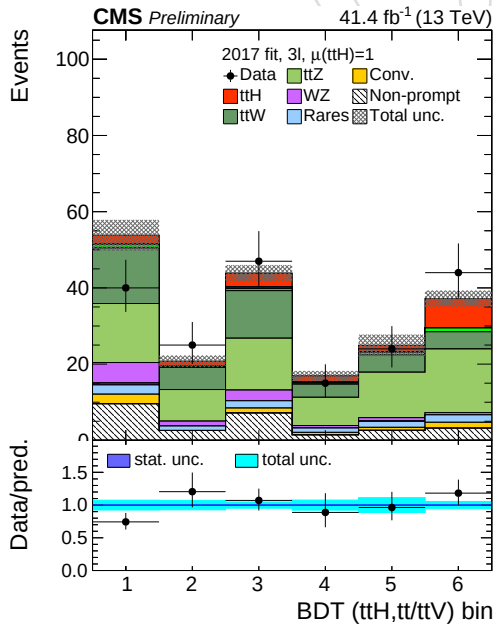
Figure 26: Distributions in the discriminating observables used for the signal extraction in the categories without taus. The SM expectations for the $t\bar{t}H$ signal and for the background processes are presented in colors. The post-fit rates ($\mu_{t\bar{t}H} = 1$) and uncertainties are used. Figures for subcategories are found in appendix O.



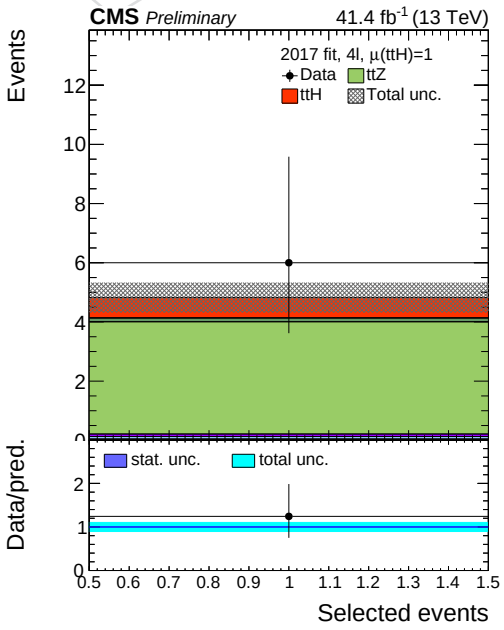
(a)



(b)



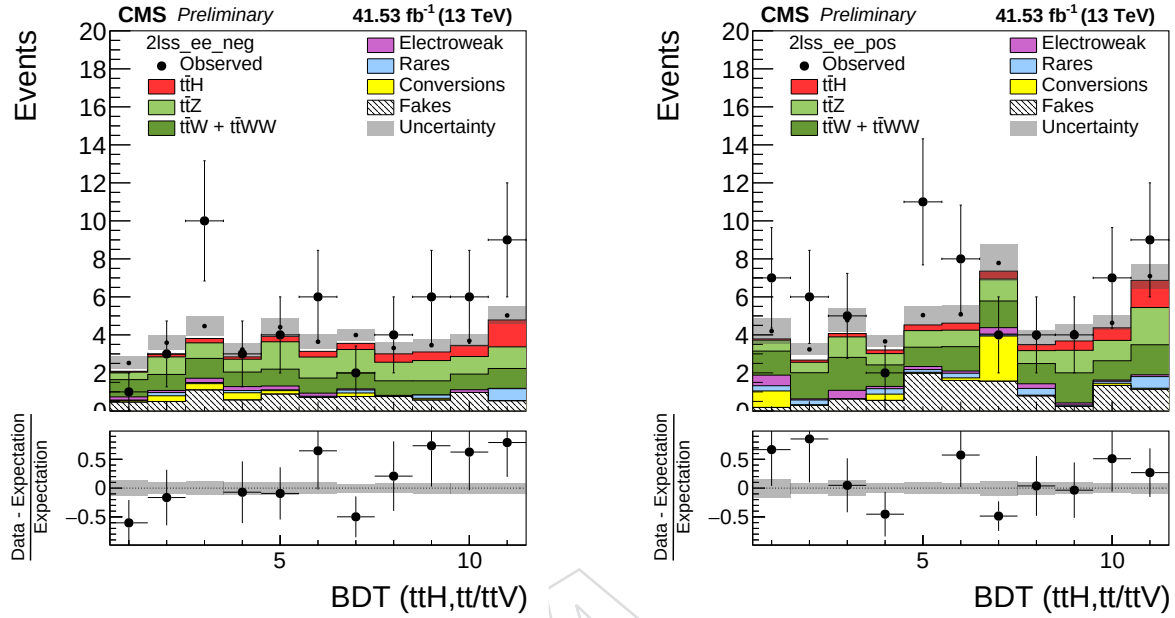
(c)



(d)

O Post-fit distributions for multilepton subcategories

Figure 27: Distributions in the discriminating observables used for the signal extraction in the categories without taus ans 2lss ($e\mu$). The SM expectations for the $t\bar{t}H$ signal and for the background processes are presented in colors.



(a) (b)

Figure 28: Distributions in the discriminating observables used for the signal extraction in the categories without taus ans 2lss ($e\mu$). The SM expectations for the $t\bar{t}H$ signal and for the background processes are presented in colors.

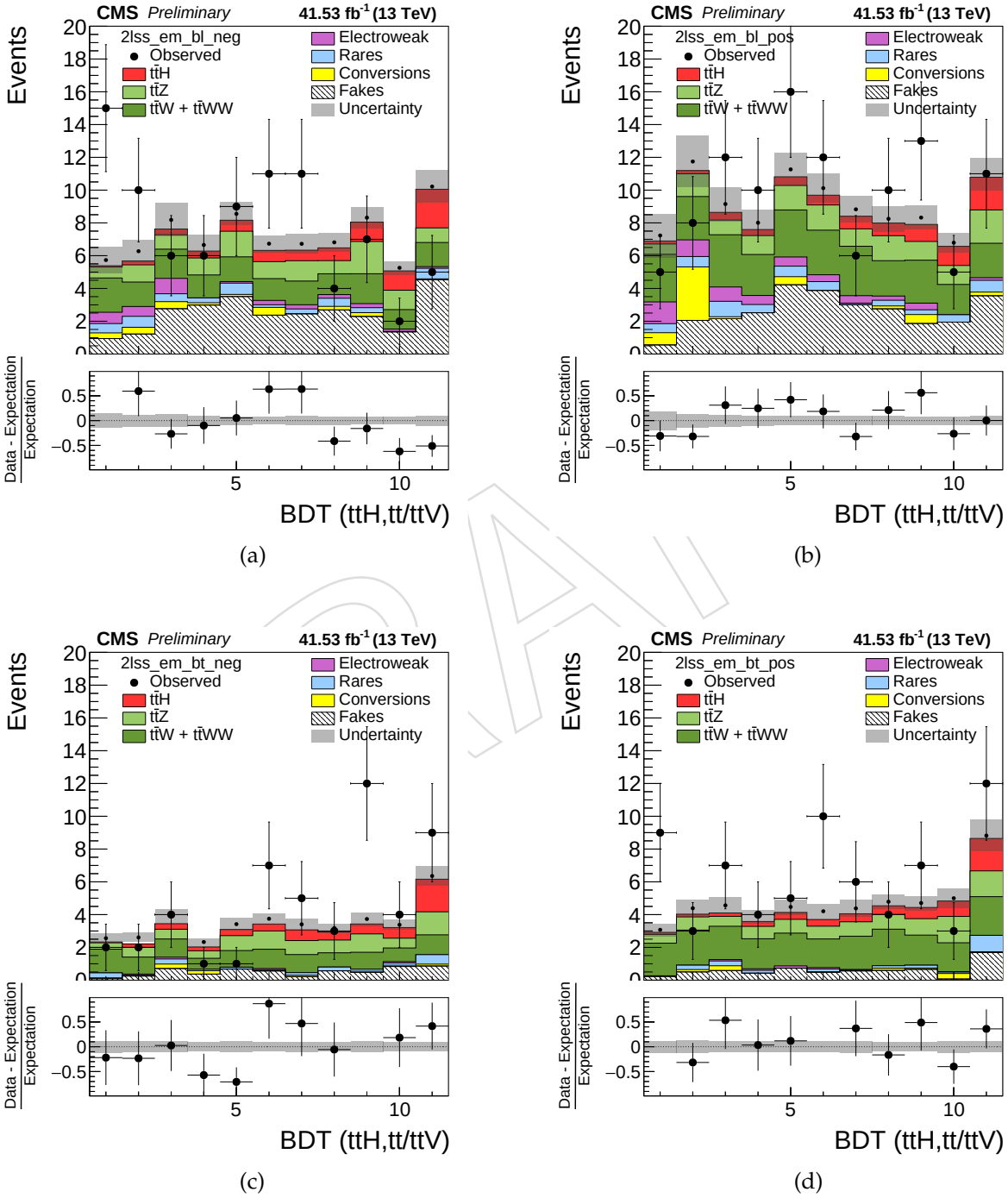


Figure 29: Distributions in the discriminating observables used for the signal extraction in the categories without taus ans 2lss ($\mu\mu$). The SM expectations for the $t\bar{t}H$ signal and for the background processes are presented in colors.

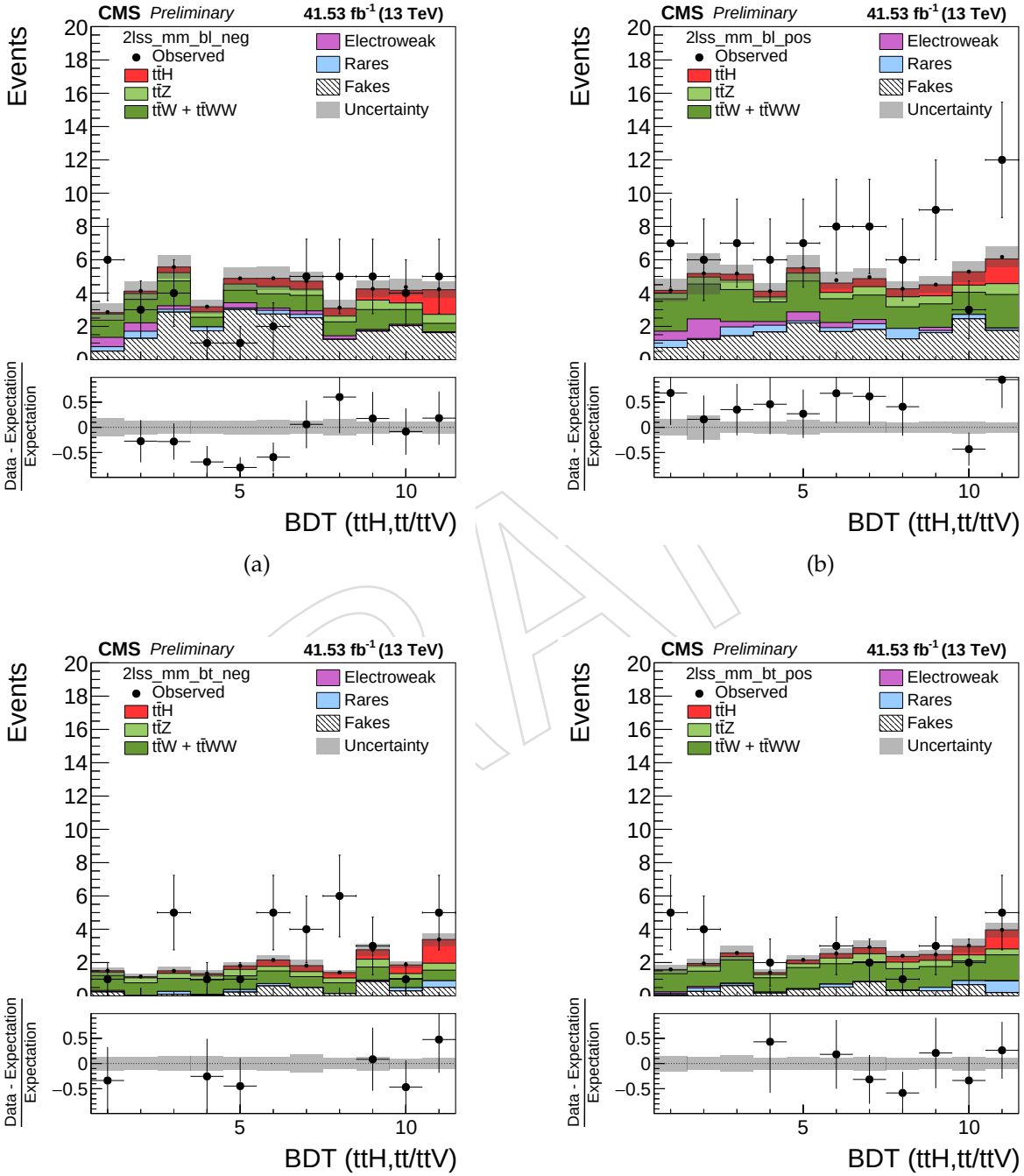


Figure 30: Distributions in the discriminating observables used for the signal extraction in the categories without taus ans 3l ($\mu\mu$). The SM expectations for the $t\bar{t}H$ signal and for the background processes are presented in colors.

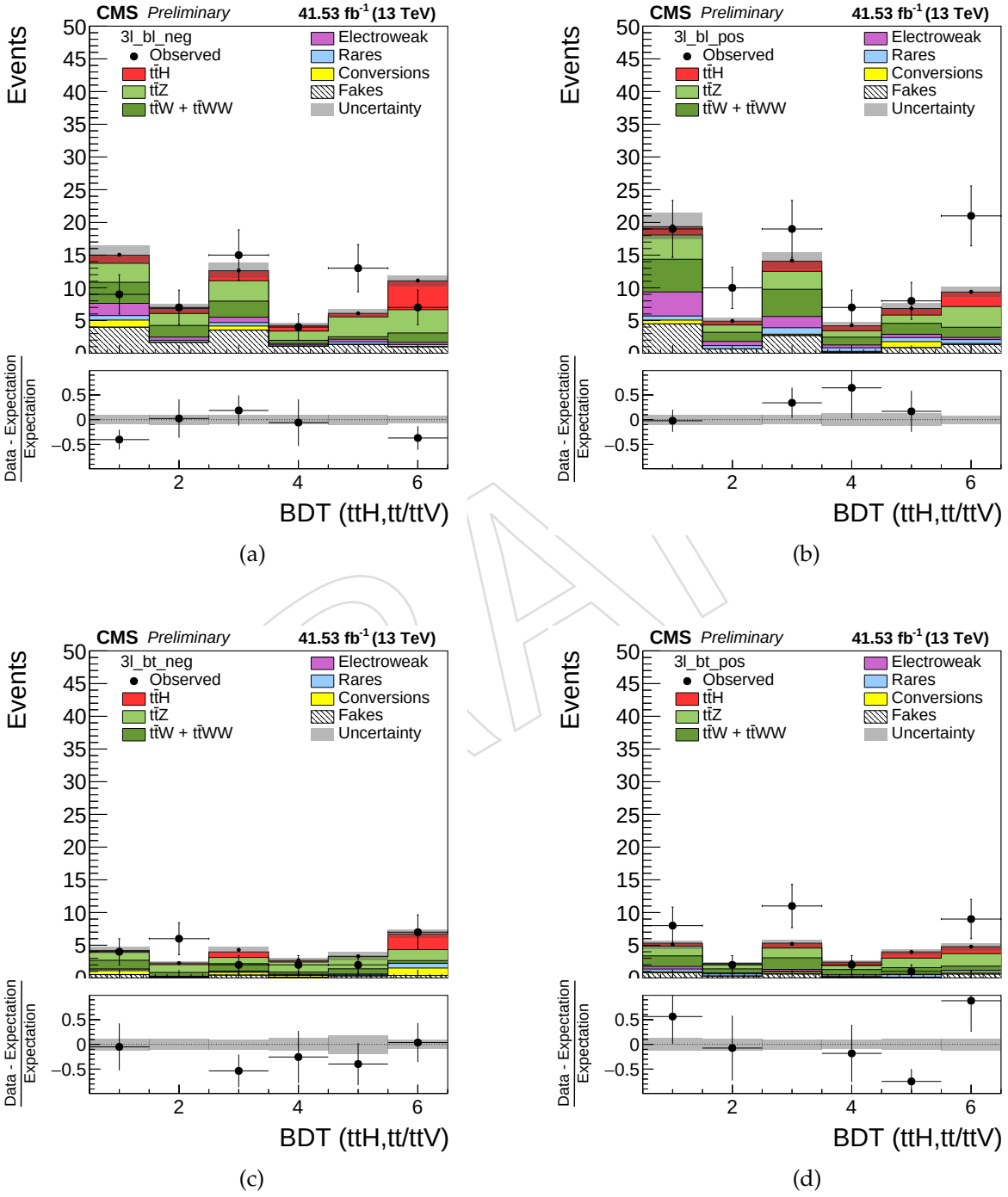


Figure 31: Distributions in the discriminating observables used for the signal extraction in the categories without taus ans 2lss ($e\mu$). The SM expectations for the $t\bar{t}H$ signal and for the background processes are presented in colors.

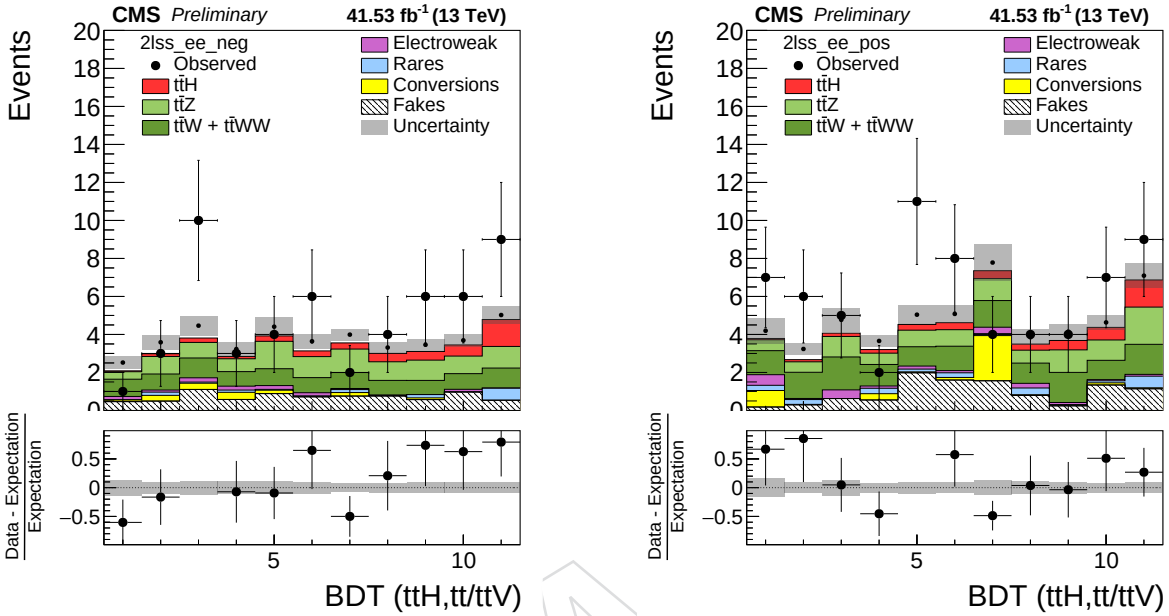


Figure 32: Distributions in the discriminating observables used for the signal extraction in the control region categories without taus ans 2lss ($e\mu$) -(3j) ttW enriched. The SM expectations for the $t\bar{t}H$ signal and for the background processes are presented in colors.

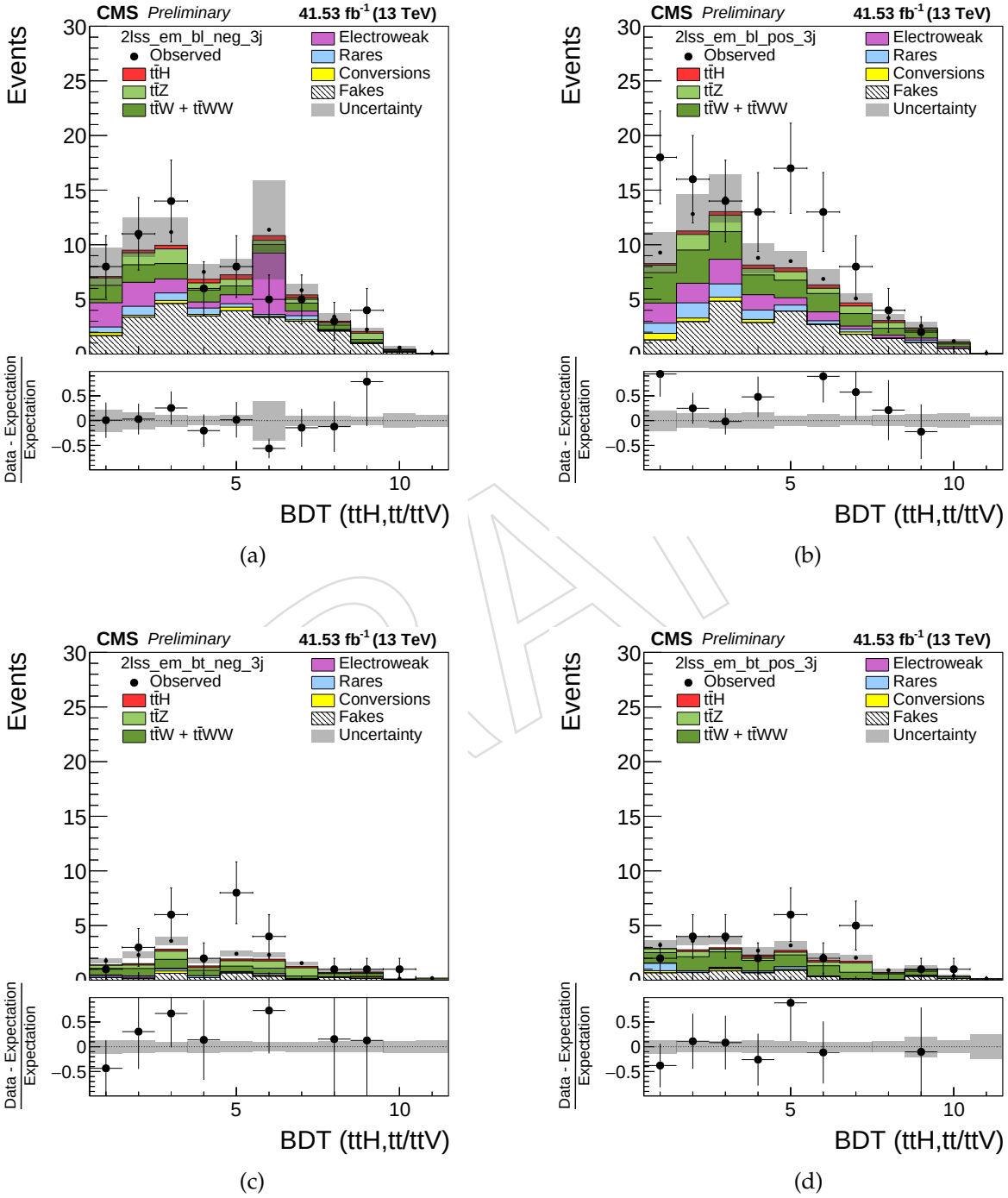


Figure 33: Distributions in the discriminating observables used for the signal extraction in the contril region categories without taus ans 2lss ($\mu\mu$) -(3j) ttW enriched. The SM expectations for the $t\bar{t}H$ signal and for the background processes are presented in colors.

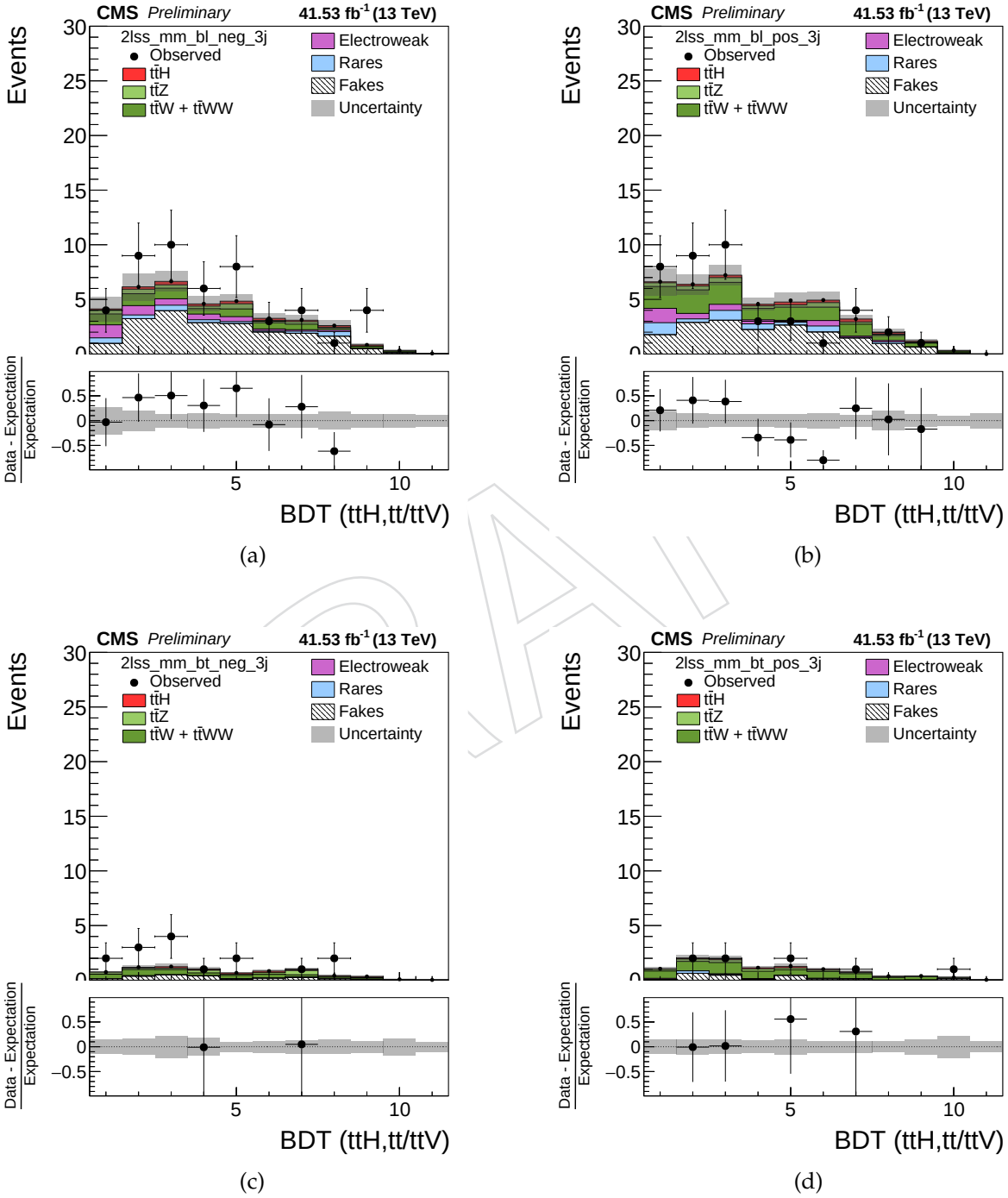


Figure 34: Distributions in the discriminating observables used for the signal extraction in the control region categories without taus ans 3l ($\mu\mu$) -(zpeak) ttZ enriched. The SM expectations for the $t\bar{t}H$ signal and for the background processes are presented in colors.

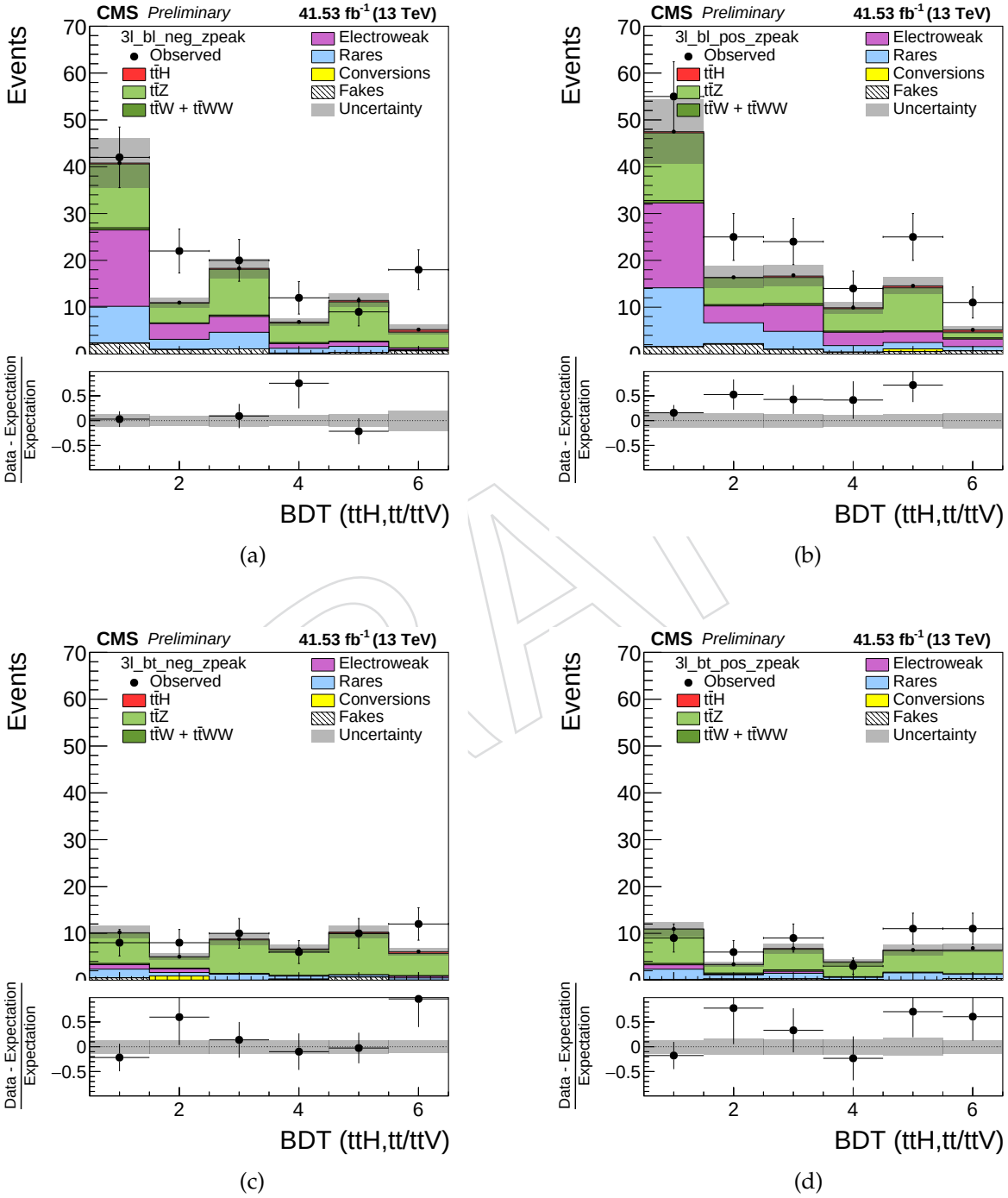
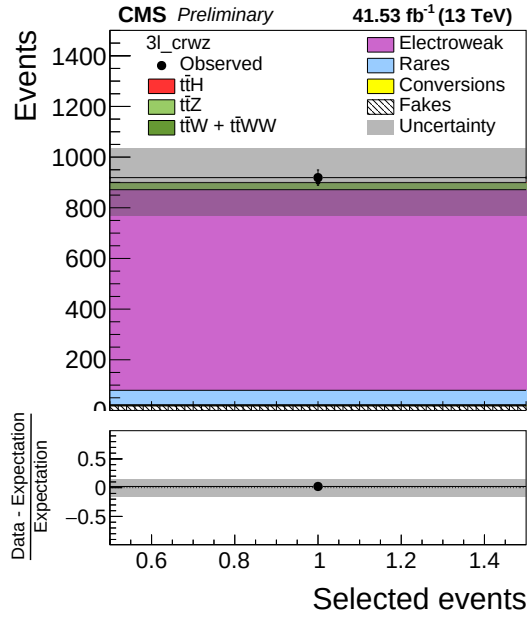


Figure 35: Distributions in the discriminating observables used for the signal extraction in the categories WW and WZ enriched. The SM expectations for the $t\bar{t}H$ signal and for the background processes are presented in colors.



(a)

References

[1] CMS Collaboration, "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC", *Phys. Lett.* **B 716** (2012) 30, 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.* **B 716** (2012) 1, doi:10.1016/j.physletb.2012.08.020, arXiv:1207.7214.

[3] ATLAS, CMS Collaboration, "Combined measurement of the Higgs boson mass in pp collisions at $\sqrt{s} = 7$ and 8 TeV with the ATLAS and CMS experiments", *Phys. Rev. Lett.* **114** (2015) 191803, doi:10.1103/PhysRevLett.114.191803, arXiv:1503.07589.

[4] S. Dawson et al., "Working Group Report: Higgs Boson", in *Proceedings, 2013 Community Summer Study on the Future of U.S. Particle Physics: Snowmass on the Mississippi (CSS2013): Minneapolis, MN, USA, July 29-August 6, 2013*. 2013. arXiv:1310.8361.

[5] K. Kondo, "Dynamical likelihood method for reconstruction of events with missing momentum. 1: method and toy models", *J. Phys. Soc. Jap.* **57** (1988) 4126–4140, doi:10.1143/JPSJ.57.4126.

[6] K. Kondo, "Dynamical likelihood method for reconstruction of events with missing momentum. 2: mass spectra for $2 \to 2$ processes", *J. Phys. Soc. Jap.* **60** (1991) 836–844, doi:10.1143/JPSJ.60.836.

[7] A. Hocker et al., "TMVA - toolkit for multivariate data analysis", *PoS* **ACAT** (2007) 040, arXiv:physics/0703039.

[8] F. Pedregosa et al., "Scikit-learn: Machine learning in python", *Journal of Machine Learning Research* **12** (2011) 2825.

[9] CMS Collaboration, "Evidence for associated production of a Higgs boson with a top quark pair in final states with electrons, muons, and hadronically decaying τ leptons at $\sqrt{s} = 13$ TeV", arXiv:1803.05485.

[10] Software guide PhysicsTools Collaboration. https://twiki.cern.ch/twiki/bin/view/CMSPublic/SWGuideGoodLumiSectionsJSONFile.

[11] J. Alwall et al., "The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations", *JHEP* **07** (2014) 079, doi:10.1007/JHEP07(2014)079, arXiv:1405.0301.

[12] P. Nason, "A new method for combining NLO QCD with shower Monte Carlo algorithms", *JHEP* **11** (2004) 040, doi:10.1088/1126-6708/2004/11/040, arXiv:hep-ph/0409146.

[13] S. Frixione, P. Nason, and C. Oleari, "Matching NLO QCD computations with parton shower simulations: the POWHEG method", *JHEP* **11** (2007) 070, doi:10.1088/1126-6708/2007/11/070, arXiv:0709.2092.

[14] S. Alioli, P. Nason, C. Oleari, and E. Re, "A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX", *JHEP* **06** (2010) 043, doi:10.1007/JHEP06(2010)043, arXiv:1002.2581.

[15] NNPDF Collaboration, "Parton distributions with QED corrections", *Nucl. Phys.* **B 877** (2013) 290, doi:10.1016/j.nuclphysb.2013.10.010, arXiv:1308.0598.

[16] NNPDF Collaboration, "Unbiased global determination of parton distributions and their uncertainties at NNLO and at LO", *Nucl. Phys.* **B 855** (2012) 153, doi:10.1016/j.nuclphysb.2011.09.024, arXiv:1107.2652.

[17] NNPDF Collaboration, "Parton distributions for the LHC Run II", *JHEP* **04** (2015) 040, doi:10.1007/JHEP04(2015)040, arXiv:1410.8849.

[18] CMS Collaboration, "Event generator tunes obtained from underlying event and multiparton scattering measurements", arXiv:1512.00815.

[19] P. Skands, S. Carrazza, and J. Rojo, "Tuning PYTHIA 8.1: the Monash 2013 tune", *Eur. Phys. J.* **C 74** (2014), no. 8, 3024, doi:10.1140/epjc/s10052-014-3024-y, arXiv:1404.5630.

[20] K. Melnikov and F. Petriello, "Electroweak gauge boson production at hadron colliders through $O(\alpha_s^2)$", *Phys. Rev.* **D 74** (2006) 114017, doi:10.1103/PhysRevD.74.114017, arXiv:hep-ph/0609070.

[21] M. Czakon and A. Mitov, "Top++: A program for the calculation of the top-pair cross section at hadron colliders", *Comput. Phys. Commun.* **185** (2014) 2930, doi:10.1016/j.cpc.2014.06.021, arXiv:1112.5675.

[22] Kant, P. and Kind, O. M. and Kintscher, T. and Lohse, T. and Martini, T. and Mölbitz, S. and Rieck, P. and Uwer, P., "HATHOR for single top-quark production: Updated predictions and uncertainty estimates for single top-quark production in hadronic collisions", *Comput. Phys. Commun.* **191** (2015) 74, doi:10.1016/j.cpc.2015.02.001, arXiv:1406.4403.

[23] M. Aliev et al., "HATHOR: HAdronic Top and Heavy quarks crOss section calculatoR", *Comput. Phys. Commun.* **182** (2011) 1034–1046, doi:10.1016/j.cpc.2010.12.040, arXiv:1007.1327.

[24] N. Kidonakis, "Two-loop soft anomalous dimensions for single top quark associated production with a W^- or H^-", *Phys. Rev.* **D 82** (2010) 054018, doi:10.1103/PhysRevD.82.054018, arXiv:1005.4451.

[25] J. M. Campbell, R. K. Ellis, and C. Williams, "Vector boson pair production at the LHC", *JHEP* **07** (2011) 018, doi:10.1007/JHEP07(2011)018, arXiv:1105.0020.

[26] CMS Collaboration. https://twiki.cern.ch/twiki/bin/view/LHCPhysics/CERNYellowReportPageBR.

[27] CMS Collaboration. https://twiki.cern.ch/twiki/bin/view/LHCPhysics/CERNYellowReportPageAt13TeV.

[28] CMS Collaboration. https://twiki.cern.ch/twiki/bin/view/CMS/SummaryTable1G25ns.

[29] CMS Collaboration. https://cms-gen-dev.cern.ch/xsdb/?searchQuery=DAS=TTWW_TuneCUETP8M2T4_13TeV-madgraph-pythia8.

[30] CMS Collaboration. https://twiki.cern.ch/twiki/bin/view/CMS/IHEPTopAnalyses.

[31] CMS Collaboration. http://pdglive.lbl.gov/BranchingRatio.action?desig=8&parCode=S043.

[32] CMS Collaboration. https://cms-gen-dev.cern.ch/xsdb/?searchQuery=DAS=TTTT_TuneCP5_13TeV-amcatnlo-pythia8.

[33] CMS Collaboration. https://cms-gen-dev.cern.ch/xsdb/?searchQuery=DAS=TTTo2L2Nu_TuneCP5_PSweights_13TeV-powheg-pythia8.

[34] CMS Collaboration. https://cms-gen-dev.cern.ch/xsdb/?searchQuery=DAS=TTToSemiLeptonic_TuneCP5_PSweights_13TeV-powheg-pythia8.

[35] CMS Collaboration. https://cms-gen-dev.cern.ch/xsdb/?searchQuery=DAS=TTToHadronic_TuneCP5_PSweights_13TeV-powheg-pythia8.

[36] CMS Collaboration. https://twiki.cern.ch/twiki/bin/view/LHCPhysics/SingleTopRefXsec.

[37] GEANT4 Collaboration, "GEANT4—a simulation toolkit", *Nucl. Instrum. Meth.* **A 506** (2003) 250, doi:10.1016/S0168-9002(03)01368-8.

[38] CMS Collaboration, "Particle-flow event reconstruction in CMS and performance for jets, taus, and E_T^{miss}", CMS Physics Analysis Summary CMS-PAS-PFT-09-001, 2009.

[39] CMS Collaboration, "Commissioning of the particle-flow event reconstruction with the first LHC collisions recorded in the CMS detector", CMS Physics Analysis Summary CMS-PAS-PFT-10-001, 2010.

[40] CMS Collaboration, "Commissioning of the particle-flow reconstruction in minimum-bias and jet events from pp collisions at 7 TeV", CMS Physics Analysis Summary CMS-PAS-PFT-10-002, 2010.

[41] CMS Collaboration, "Particle-flow commissioning with muons and electrons from J/ψ(1S) and W events at 7 TeV", CMS Physics Analysis Summary CMS-PAS-PFT-10-003, 2010.

[42] CMS Collaboration, "Particle-flow reconstruction and global event description with the CMS detector", *JINST* **12** (2017), no. 10, P10003, doi:10.1088/1748-0221/12/10/P10003, arXiv:1706.04965.

[43] CMS Collaboration, "Performance of electron reconstruction and selection with the CMS detector in pp collisions at $\sqrt{s} = 8$ TeV", *JINST* **10** (2015) P06005, doi:10.1088/1748-0221/10/06/P06005, arXiv:1502.02701.

[44] EGamma POG Collaboration. https://twiki.cern.ch/twiki/bin/view/CMS/MultivariateElectronIdentificationRun2#Recommended

[45] CMS Collaboration, "Performance of CMS muon reconstruction in pp collision events at $\sqrt{s} = 7$ TeV", *JINST* **7** (2012) P10002, doi:10.1088/1748-0221/7/10/P10002, arXiv:1206.4071.

[46] Muon POG Collaboration.
https://twiki.cern.ch/twiki/bin/view/CMS/SWGuideMuonIdRun2#Medium_Muon.

[47] M. Cacciari, G. P. Salam, and G. Soyez, "The catchment area of jets", *JHEP* **04** (2008) 005, doi:10.1088/1126-6708/2008/04/005, arXiv:0802.1188.

[48] M. Cacciari and G. P. Salam, "Pileup subtraction using jet areas", *Phys. Lett.* **B 659** (2008) 119, doi:10.1016/j.physletb.2007.09.077, arXiv:0707.1378.

[49] CMS Collaboration, "Search for $t\bar{t}H$ production in multilepton final states at $\sqrt{s} = 13$ TeV", CMS Physics Analysis Summary CMS-PAS-HIG-15-008, 2015.

[50] CMS Collaboration, "Search for associated production of Higgs bosons and top quarks in multilepton final states at $\sqrt{s} = 13$ TeV", CMS Physics Analysis Summary CMS-PAS-HIG-16-022, 2016.

[51] C. Botta, M. Peruzzi, and G. Petrucciani.
https://indico.cern.ch/event/446448/contributions/1952877/attachments/1161380/1672225/lepawa

[52] Particle Data Group, "Review of particle physics", *Chin. Phys. C* **40** (2016) 100001, doi:10.1088/1674-1137/40/10/100001.

[53] C. Mueller et al., "Search for ttH in multilepton final states with 2016 data", *CMS Analysis Note* **2016/211** (2016).

[54] CMS Collaboration, "Reconstruction and identification of τ lepton decays to hadrons and ν_τ at CMS", *JINST* **11** (2016), no. 01, P01019, doi:10.1088/1748-0221/11/01/P01019, arXiv:1510.07488.

[55] CMS Collaboration, "Performance of reconstruction and identification of τ leptons in their decays to hadrons and ν_τ in LHC Run 2", CMS Physics Analysis Summary CMS-PAS-TAU-16-002, 2016.

[56] Perez, E. and Steggemann, J. and Takahashi, Y. and Veelken, C., "Improved τ identification algorithm for LHC Run 2", *CMS Analysis Note* **2015/083** (2015).

[57] Nayak, A. and Raspereza, A. and Steggemann, J. and Takahashi, Y. and Veelken, C., "Multivariate τ identification algorithms for 13 TeV data", *CMS Analysis Note* **2015/310** (2015).

[58] Tau POG.
https://twiki.cern.ch/twiki/bin/view/CMSPublic/SWGuidePFTauID#Rerunning_of_the_tau_ID_on_M

[59] M. Cacciari, G. P. Salam, and G. Soyez, "The anti-k_T jet clustering algorithm", *JHEP* **04** (2008) 063, doi:10.1088/1126-6708/2008/04/063, arXiv:0802.1189.

[60] JetMET POG. https://twiki.cern.ch/twiki/bin/view/CMS/JetID13TeVRun2017.

[61] CMS Collaboration, "Determination of jet energy calibration and transverse momentum resolution in CMS", *JINST* **6** (2011) P11002, doi:10.1088/1748-0221/6/11/P11002, arXiv:1107.4277.

[62] CMS Collaboration, "Identification of b quark jets with the CMS experiment", *JINST* **8** (2013) P04013, doi:10.1088/1748-0221/8/04/P04013, arXiv:1211.4462.

[63] CMS Collaboration, "Identification of heavy-flavour jets with the CMS detector in pp collisions at 13 TeV", *Submitted to JINST* (2017) arXiv:1712.07158.

[64] BTV POG Collaboration. https://twiki.cern.ch/twiki/bin/viewauth/CMS/BtagRecommendation94X.

[65] CMS Collaboration, "Performance of the CMS missing transverse momentum reconstruction in pp data at $\sqrt{s} = 8$ TeV", *JINST* **10** (2015), no. 02, P02006, doi:10.1088/1748-0221/10/02/P02006, arXiv:1411.0511.

[66] L. Bianchini, J. Conway, E. K. Friis, and C. Veelken, "Reconstruction of the Higgs mass in $H \to \tau\tau$ events by dynamical likelihood techniques", *J. Phys. Conf. Ser.* **513** (2014) 022035, doi:10.1088/1742-6596/513/2/022035.

[67] L. Bianchini et al., "Reconstruction of the Higgs mass in events with Higgs bosons decaying into a pair of τ leptons using matrix element techniques", *Nucl. Instrum. Meth.* **A 862** (2017) 54, doi:10.1016/j.nima.2017.05.001, arXiv:1603.05910.

[68] JetMET POG. https://twiki.cern.ch/twiki/bin/view/CMS/MissingETOptionalFiltersRun2#MET_Filter_Recommend

[69] A. Tiko. https://indico.cern.ch/event/606232/contributions/2444209/attachments/1398539/2133074/chargef

[70] CMS Collaboration, "Measurements of properties of the Higgs boson decaying into the four-lepton final state in pp collisions at $\sqrt{s} = 13$ TeV", *JHEP* **11** (2017) 047, doi:10.1007/JHEP11(2017)047, arXiv:1706.09936.

[71] M. Hildreth. https://twiki.cern.ch/twiki/bin/viewauth/CMS/PdmVPileUpDescription#Reweighting_Method.

[72] K. Ehatäht et al. https://gitlab.cern.ch/ttH_leptons/doc/blob/master/2017/datasets.md#52-data-samples.

[73] T. Boccali et al. https://indico.cern.ch/event/695872/contributions/2877123/attachments/1593469/2522749/pileup_

[74] J. Berryhill, V. Halyo, A. Hunt, and K. Mishra, "Generic Tag and Probe tool for measuring efficiency at CMS with early data", *CMS Analysis Note* **2009/111** (2009).

[75] T. Ruggles, H. Sert, and T. Strebler. https://indico.cern.ch/event/700042/contributions/2871830/attachments/1591232/2527113/180129_

[76] Bloch, D. and others, "Search for $t\bar{t}H$ in multilepton final states with the full 2016 dataset", *CMS Analysis Note* **2017/029** (2017).

[77] J. Bechtel et al., "Tau reconstruction and performance using the 2016 dataset", *CMS Analysis Note* **2017/094** (2017).

[78] Tau POG. https://indico.cern.ch/event/719250/contributions/2971854/attachments/1635435/2609013/tauid_re

[79] BTV POG Collaboration. https://twiki.cern.ch/twiki/bin/view/CMS/BTagSFMethods.

[80] M. Marionneau and C. Veelken, "Study of E_T^{miss} performance in events containing one Z boson decaying into two muons in 2012 data", *CMS Analysis Note* **2012/333** (2012).

[81] JetMET POG Collaboration. https://twiki.cern.ch/twiki/bin/view/CMSPublic/SWGuidePATTools#MET_Systematics_Tools.

[82] CMS Collaboration, "Performance of missing energy reconstruction in 13 TeV pp collision data using the CMS detector", CMS Physics Analysis Summary CMS-PAS-JME-16-004, 2016.

[83] A. Brinkerhoff et al., "Search for the standard model Higgs boson produced in association with top quarks and decaying to leptons ", *CMS Analysis Note* **2013/159** (2013).

[84] B. Stieger et al., "Search for ttH in multilepton final states at 13 TeV", *CMS Analysis Note* **2015/321** (2016).

[85] Conway, J. and Friis, E. K. and Veelken, C., "Estimation of background contributions to τ analyses via fake-rate technique", *CMS Analysis Note* **2010/074** (2010).

[86] C. E. Cid, S. S. Cruz, and P. Vischia. https://twiki.cern.ch/twiki/pub/CMS/HIG18019Reviews/document.pdf.

[87] Tau POG. https://twiki.cern.ch/twiki/bin/viewauth/CMS/TauIDRecommendation13TeV.

[88] JetMET POG. https://twiki.cern.ch/twiki/bin/view/CMSPublic/WorkBookJetEnergyCorrections.

[89] BTV POG Collaboration. https://twiki.cern.ch/twiki/bin/viewauth/CMS/BtagRecommendation80X.

[90] LHC Higgs Cross Section Working Group Collaboration, "Handbook of LHC Higgs Cross Sections: 3. Higgs Properties: Report of the LHC Higgs Cross Section Working Group", Technical Report arXiv:1307.1347. CERN-2013-004, Geneva, 2013. Comments: 404 pages, 139 figures, to be submitted to CERN Report. Working Group web page: https://twiki.cern.ch/twiki/bin/view/LHCPhysics/CrossSections.

[91] CMS Collaboration, "CMS luminosity measurement for the 2017 data-taking period at $\sqrt{s} = 13$ TeV", CMS Physics Analysis Summary CMS-PAS-LUM-17-004, 2017.

[92] ATLAS and CMS Collaborations and LHC Higgs Combination Group, "Procedure for the LHC Higgs boson search combination in summer 2011", Technical Report ATL-PHYS-PUB-2011-011, CMS-NOTE-2011-005, 2011.

[93] CMS Collaboration, "Combined results of searches for the standard model Higgs boson in pp collisions at $\sqrt{s} = 7$ TeV", *Phys. Lett. B* **710** (2012) 26, doi:10.1016/j.physletb.2012.02.064, arXiv:1202.1488.

[94] J. S. Conway, "Incorporating nuisance parameters in likelihoods for multisource spectra", arXiv:1103.0354.

[95] R. Barlow and C. Beeston, "Fitting using finite Monte Carlo samples", *Comput. Phys. Commun.* **77** (1993) 219.

[96] J. S. Conway, "Incorporating nuisance parameters in likelihoods for multisource spectra", (2011). arXiv:1103.0354.

[97] A. L. Read, "Presentation of search results: The CL(s) technique", *J. Phys.* **G 28** (2002) 2693, doi:10.1088/0954-3899/28/10/313.

[98] A. Korytov, G. Petrucciani, M. Chen et al., "Procedure for the lhc higgs boson search combination in summer 2011", *CMS Note* **2011/298** (2011).

[99] Higgs PAG. https://twiki.cern.ch/twiki/bin/viewauth/CMS/SWGuideHiggsAnalysisCombinedLimit.

[100] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system", *CoRR* **abs/1603.02754** (2016) arXiv:1603.02754.

[101] Carvalho, A. https://indico.cern.ch/event/718591/contributions/3022720/attachments/1658083/2655364/BDT_20

[102] Pata, J. https://github.com/jpata/mlglue.

[103] Beaudette, F. and others, "Search for the associated production of Higgs boson with top-quark pair in final states with a τ lepton at $\sqrt{s} = 13$ TeV", *CMS Analysis Note* **2016/372** (2016).

[104] CMS Collaboration, "The fast simulation of the CMS detector at LHC", *J. Phys. Conf. Ser.* **331** (2011) 032049, doi:10.1088/1742-6596/331/3/032049.

[105] Tao, Z. and others. https://indico.cern.ch/event/718596/contributions/3015984/attachments/1654622/2648165/180523_

[106] CMS Collaboration, "Measurement of the cross section for top quark pair production in association with a W or Z boson in proton-proton collisions at $\sqrt{s} = 13$ TeV", arXiv:1711.02547.

[107] Tau POG. https://hypernews.cern.ch/HyperNews/CMS/get/tauid/822/1/1.html.

[108] I. Volobouev, "Matrix element method in HEP: transfer functions, efficiencies, and likelihood normalization. ", arXiv:1101.2259.

[109] Bianchini, L. and others, "Search for the associated $t\bar{t}H$ ($\rightarrow b\bar{b}$) production using the matrix element method", *CMS Analysis Note* **2013/313** (2013).

[110] Aldaya, M. and others, "Search for ttH, H $\rightarrow$ bb decays using the full 2016 data sample", *CMS Analysis Note* **2017/063** (2017).

[111] CMS Collaboration, "Jet algorithms performance in 13 TeV data", CMS Physics Analysis Summary CMS-PAS-JME-16-003, 2016.

[112] R. K. Dewanjee. https://indico.cern.ch/event/718591/contributions/3022719/attachments/1657989/2654969/ttH_me

[113] Wikipedia Collaboration. https://en.wikipedia.org/wiki/Linear_least_squares_(mathematics)#Weighted_linear_least_squares.

[114] A. Tiko. https://indico.cern.ch/event/606232/contributions/2444209/attachments/1398539/2133074/chargef