CMS Physics Analysis Summary

Contact: cms-pag-conveners-higgs@cern.ch 2017/05/15

Search for lepton flavour violating decays of the Higgs boson to $\mu\tau$ and eτ in proton-proton collisions at $\sqrt{s} = 13$ TeV

The CMS Collaboration

Abstract

A search for lepton flavor violating decays of the 125 GeV Higgs boson in the $\mu\tau$ and eτ decay modes is presented. The search is based on a dataset of 35.9 fb^{-1} of proton-proton collisions collected with the CMS detector in 2016, at a center-of-mass energy of 13 TeV. The tau leptons are reconstructed in the leptonic and hadronic decay modes. No significant excess over the standard model background expectation is observed. The observed (expected) upper limits on the branching fraction of the Higgs boson are found to be $\mathcal{B}(\mathrm{H} \to \mu\tau) < 0.25(0.25)\%$ and $\mathcal{B}(\mathrm{H} \to \mathrm{e}\tau) < 0.61(0.37)\%$ at 95% confidence level. These results are used to derive upper limits on the off-diagonal $\mu\tau$ and eτ Yukawa couplings, $\sqrt{|Y_{\mu\tau}|^2 + |Y_{\tau\mu}|^2} < 1.43 \times 10^{-3}$ and $\sqrt{|Y_{\mathrm{e}\tau}|^2 + |Y_{\tau\mathrm{e}}|^2} < 2.26 \times 10^{-3}$ at 95% CL.

1 Introduction

The discovery of the Higgs boson [1–3] has generated great interest in exploring its properties. In the standard model (SM), lepton flavour violating (LFV) decays of the Higgs boson are forbidden. Such decays can occur naturally in models with more than one Higgs boson doublet [4]. They also arise in supersymmetric models [5–11], composite Higgs models [12, 13], models with flavour symmetries [14], Randall–Sundrum models [15–17], and others [18–26]. The CMS Collaboration has published a search in the $\mathrm{H} \to \mu\tau$ channel [27] showing an excess of data with respect to the SM background-only hypothesis at $M_{\mathrm{H}} = 125\,\mathrm{GeV}$ with a significance of 2.4 standard deviations (σ). A constraint is set on the branching fraction $\mathcal{B}(\mathrm{H} \to \mu\tau) < 1.51\%$ at 95% confidence level (CL), while the best fit branching fraction is $\mathcal{B}(\mathrm{H} \to \mu\tau) = (0.84^{+0.39}_{-0.37})\%$. The CMS Collaboration has also reported searches for $\mathrm{H} \to \mathrm{e}\tau$ and $\mathrm{H} \to \mathrm{e}\mu$ [28] finding no evidence of a signal and setting limits, $\mathcal{B}(\mathrm{H} \to \mathrm{e}\tau) < 0.69\%$ and $\mathcal{B}(\mathrm{H} \to \mathrm{e}\mu) < 0.035\%$ at the 95% CL. The ATLAS Collaboration reported searches for $\mathrm{H} \to \mathrm{e}\tau$ and $\mathrm{H} \to \mu\tau$, finding no significant excess of events over the background expectation, and set limits of $\mathcal{B}(\mathrm{H} \to \mathrm{e}\tau) < 1.04\%$, $\mathcal{B}(\mathrm{H} \to \mu\tau) < 1.43\%$ at 95% CL [29, 30].

Indirect constraints arise because the presence of LFV Higgs boson couplings allows, $\mu \to \mathrm{e}$, $\tau \to \mu$, and $\tau \to \mathrm{e}$ to proceed via a virtual Higgs boson [31, 32]. The experimental limits on these decays have been translated into constraints on $\mathcal{B}(\mathrm{H} \to \mathrm{e}\mu)$, $\mathcal{B}(\mathrm{H} \to \mu\tau)$ and $\mathcal{B}(\mathrm{H} \to \mathrm{e}\tau)$ [33, 34]. The null result for $\mu \to \mathrm{e}\gamma$ [35] strongly constrains $\mathcal{B}(\mathrm{H} \to \mathrm{e}\tau) < \mathcal{O}(10^{-8})$. However, the constraints $\mathcal{B}(\mathrm{H} \to \mu\tau) < \mathcal{O}(10\%)$ and $\mathcal{B}(\mathrm{H} \to \mathrm{e}\tau) < \mathcal{O}(10\%)$ are much less stringent. These come from searches for rare tau lepton decays [36] such as $\tau \to \mathrm{e}\gamma$ and $\tau \to \mu\gamma$, and the measurements of the electron and muon magnetic moment. Exclusion limits on the electric dipole moment of the electron [37] also provide complementary constraints.

This paper describes a search for LFV decays of the Higgs boson with $M_{\mathrm{H}} = 125\,\mathrm{GeV}$. The search is performed in four decay channels, $\mathrm{H} \to \mu\tau_{\mathrm{e}}$, $\mathrm{H} \to \mu\tau_h$, $\mathrm{H} \to \mathrm{e}\tau_\mu$, $\mathrm{H} \to \mathrm{e}\tau_h$, where τ_{e},τ_μ and τ_h correspond to the electronic, muonic and hadronic decay channels of τ leptons, respectively. The decay channels, $\mathrm{H} \to \mathrm{e}\tau_{\mathrm{e}}$ and $\mathrm{H} \to \mu\tau_\mu$, are not considered due to the large background contribution from Z boson decays. The expected final state signatures are very similar to the SM $\mathrm{H} \to \tau\tau$ decays, studied by CMS [38, 39] and ATLAS [40], but with some significant kinematic differences. The electron(muon) in the LFV $\mathrm{H} \to \mathrm{e}(\mu)\tau$ decay is produced promptly, and tends to have a larger momentum than in the SM $\mathrm{H} \to \tau_{\mathrm{e}(\mu)}\tau_h$ decay.

This paper is organized as follows. After a description of the CMS detector (Section 2) and of the collision data and simulated samples used in the analysis (Section 3), the event reconstruction is described in Section 4. The backgrounds are common to all channels but with different rates in each. They are described in Section 5. The event selection and background validation are described separately for the two Higgs decay modes $\mathrm{H} \to \mathrm{e}\tau$ and $\mathrm{H} \to \mu\tau$ in Section 6. The results are then presented in Section 8.

2 The CMS detector

A detailed description of the CMS detector, together with a definition of the coordinate system used and the relevant kinematic variables, can be found in Ref. [41]. The momenta of charged particles are measured with a silicon pixel and strip tracker that covers the pseudorapidity range $|\eta| < 2.5$, in a 3.8 T axial magnetic field. A lead tungstate crystal electromagnetic calorimeter (ECAL) and a brass and scintillator hadron calorimeter, both consisting of a barrel section and two endcaps, cover the pseudorapidity range $|\eta| < 3.0$. A steel and quartz-fibre Cherenkov forward detector extends the calorimetric coverage to $|\eta| < 5.0$. The outermost component of

the CMS detector is the muon system, consisting of gas-ionization detectors placed in the steel flux-return yoke of the magnet to identify the muons traversing the detector. The two-level CMS trigger system selects events of interest for permanent storage. The first trigger level, composed of custom hardware processors, uses information from the calorimeters and muon detectors to select events in less than 3.2 μs. The software algorithms of the high-level trigger, executed on a farm of commercial processors, reduce the event rate to about 1 kHz using information from all detector subsystems.

3 Collision data and simulated events

The analysis uses samples of proton-proton collisions collected in 2016 by the CMS experiment at the LHC, at a center-of-mass energy of $\sqrt{s} = 13$ TeV, and corresponding to an integrated luminosity of 35.9 fb^{-1}. Isolated single muon and single electron triggers are used to collect the data samples in the $H \to \mu\tau$ and $H \to e\tau$ channels respectively. Simulated samples of signal and background events are produced with several event generators. The Higgs bosons are produced in proton-proton collisions predominantly by gluon fusion (GF) [42], but also by vector boson fusion (VBF) [43] and in association with a W or Z boson [44]. The $H \to \mu\tau$ and $H \to e\tau$ decay samples are generated with POWHEG 1.0 [45–49]. The MADGRAPH 5.1 [50] generator is used for $Z +$ jets and $W +$ jets processes, AMC@NLO [51] for diboson production, and POWHEG for $t\bar{t}$ and single-top-quark production. The POWHEG and MADGRAPH generators are interfaced with PYTHIA 8 for parton shower, fragmentation, as well as τ lepton decays. The PYTHIA parameters for the underlying event description are set to the CUETP8M1 tune [52]. Due to the high luminosities attained during data-taking, many events have multiple proton-proton interactions per bunch crossing (pileup). All simulated samples are reweighted to match the pileup distribution observed in data. The CMS detector response is modelled using GEANT4 [53].

4 Event reconstruction

The tracking system is able to separate collision vertices as close as 0.5 mm to each other along the beam direction [54]. The primary vertex, assumed to correspond to the hard-scattering process, is the vertex for which the sum of the squared transverse momentum p_T^2 of all associated tracks is largest. The pileup interactions affect the identification of most of the physics objects, such as jets and leptons, and variables such as lepton isolation.

A particle-flow (PF) algorithm [55–57] combines the information from all CMS subdetectors to identify and reconstruct the individual particles emerging from all interactions in the event: charged and neutral hadrons, photons, muons, and electrons. The charged particles are required to be consistent with the primary vertex. The individual particles are then used to reconstruct jets, hadronic τ decays, and to quantify the isolation of leptons and photons and reconstruct E_T^{miss}. The missing transverse energy vector, $\vec{E}_T^{\text{miss}}$, is defined as the negative of the vector p_T sum of all identified PF objects in the event [58]. Its magnitude is referred to as E_T^{miss}. The variable $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$, where ϕ is the azimuthal coordinate, is used to measure the separation between reconstructed objects in the detector.

Electron reconstruction requires the matching of an energy cluster in the ECAL with a track in the silicon tracker [59]. Electron candidates are accepted in the range $|\eta| < 2.5$, with the exception of the region $1.44 < |\eta| < 1.56$ where service infrastructure for the detector is located. Electron identification uses a multivariate discriminant that combines observables sensitive to

the amount of bremsstrahlung along the electron trajectory, the geometrical and momentum matching between the electron trajectory and associated clusters, isolation, and shower-shape observables. Additional requirements are imposed to remove electrons produced by photon conversions.

Muon candidates are obtained from combined fits of tracks in the tracker and muon detector seeded by track segments in the muon detector alone, including compatibility with small energy depositions in the calorimeters [60]. Identification is based on track quality and isolation. The muon momentum is measured with the combined fit.

Hadronically decaying τ leptons are reconstructed and identified using an algorithm [61, 62] that selects the decay modes with one charged hadron and up to two neutral pions, or three charged hadrons. A photon from a neutral-pion decay can convert in the tracker material into an electron-positron pair, which can then radiate photons. These particles give rise to several ECAL energy deposits at the same η value but separated in ϕ. They are reconstructed as several photons by the PF algorithm. To increase the acceptance for these converted photons, the neutral pions are identified by clustering the reconstructed photons and electrons in narrow strips along the ϕ direction. The charge of τ_h candidates is reconstructed by summing the charges of all particles included in the construction of the candidate, except for the electrons contained in strips. Identification and isolation conditions are enforced by requiring the τ_h lepton candidates to pass a working point of a multivariate (MVA) discriminator that takes as input various isolation variables and variables related to τ lepton lifetime information. The tight working point is used as it brings the best sensitivity to the analysis. Additionally, electrons and muons misidentified as τ_h candidates are rejected by using an MVA or a cut-based discriminator, respectively. The specific working points depend on the final state studied and on the background composition. leptons that decay to muons and electrons are reconstructed as prompt muons and electrons as described above.

Jets misidentified as electrons or muons or are suppressed by imposing isolation requirements, summing the neutral and charged particle contributions in cones of $\Delta R = 0.4$ around the lepton. The isolation of muons (electrons) is measured relative to their transverse momentum p_{T}^{ℓ}, by summing over the transverse momenta of PF particles in a cone with $\Delta R < 0.4$ (0.3) around the lepton:

$$R_{\mathrm{Iso}}^{\ell} = \left(\sum p_{\mathrm{T}}^{\mathrm{charged}} + \max\left[0, \sum p_{\mathrm{T}}^{\mathrm{neutral}} + \sum p_{\mathrm{T}}^{\gamma} - p_{\mathrm{T}}^{\mathrm{PU}}\left(\ell\right)\right] \right) / p_{\mathrm{T}}^{\ell}.$$

The energy deposited within the isolation cone is contaminated by energy from pileup and the underlying event. The effect of pileup is reduced by requiring the tracks considered in the isolation sum to be compatible with originating from the production vertex of the lepton. The contribution to the isolation from pileup and the underlying event is subtracted on an event-by-event basis. In the case of electrons, this contribution is estimated from the product of the measured energy density ρ for the event, determined using the ρ median estimator implemented in FASTJET [63], and an effective area corresponding to the isolation cone. In the case of muons and hadronically decaying τ leptons, it is estimated on a statistical basis through the modified $\Delta\beta$ correction described in Ref. [60].

Jets are reconstructed from all the particles using the anti-k_{t} jet clustering algorithm [64] implemented in FASTJET, with a distance parameter of $\Delta R = 0.4$. The jet energies are corrected subtracting the contribution of particles created in pileup interactions and in the underlying event [65]. Any jet within $\Delta R = 0.4$ of the identified leptons is removed.

The transverse mass $M_{\mathrm{T}}(\ell)$ is a variable formed from the lepton (ℓ) momentum and the missing

transverse energy vector: $M_{\mathrm{T}\ell} = \sqrt{2p_{\mathrm{T}}^{\ell} E_{\mathrm{T}}^{\mathrm{miss}}(1-\cos\Delta\phi_{\vec{p}_{\mathrm{T}}^{\ell}-\vec{E}_{\mathrm{T}}^{\mathrm{miss}}})}$. It is used to discriminate the Higgs boson signal candidates from the W + jets backgrounds. The collinear mass, M_{col}, provides an estimate of M_{H} using the observed decay products of the Higgs boson candidate. It is constructed using the collinear approximation based on the observation that, since $M_{\mathrm{H}} \gg M_{\tau}$, the τ decay products are highly Lorentz boosted in the direction of the τ [66]. The neutrino momenta can be approximated to have the same direction as the other visible decay products of the τ (τ^{vis}) and the component of the $\vec{E}_{\mathrm{T}}^{\mathrm{miss}}$ in the direction of the visible τ lepton decay products is used to estimate the transverse component of the neutrino momentum ($p_{\mathrm{T}}^{\nu,\,\mathrm{est}}$). The collinear mass can then be derived from the visible mass of the τ-μ or τ-e system (M_{vis}) as $M_{\mathrm{col}} = M_{\mathrm{vis}}/\sqrt{x_{\tau}^{\mathrm{vis}}}$, where x_{τ}^{vis} is the fraction of energy carried by the visible decay products of the τ ($x_{\tau}^{\mathrm{vis}} = p_{\mathrm{T}}^{\tau^{\mathrm{vis}}}/(p_{\mathrm{T}}^{\tau^{\mathrm{vis}}} + p_{\mathrm{T}}^{\nu,\,\mathrm{est}})$).

5 Background Estimation

The signal is a prompt isolated lepton, e or μ, along with an oppositely charged isolated lepton of different flavour (τ_{μ}, τ_{e} or τ_h). The neutrino in the τ lepton decay leads to the presence of significant missing energy. The largest backgrounds to this signature are from $\mathrm{Z} \to \tau\tau$ decays and from W+jets and QCD multijet production. In the latter, PF objects (predominantly jets), are misidentified as leptons. The misidentified lepton backgrounds are estimated with collision data. Smaller backgrounds are estimated with simulation. These include $\mathrm{t\bar{t}}$ pairs, SM Higgs boson production ($\mathrm{H} \to \tau\tau$), WW, WZ, ZZ + jets, $\mathrm{W}\gamma^{(*)}$ + jets processes, and single top quark production. The background estimation techniques are described in detail below. The background estimation is validated with a loose selection for each channel (described in following sections) by comparing the distributions of the kinematic variables with data. It is also validated with control regions that are enhanced with the dominant backgrounds.

The $Z \to \tau\tau$ background is estimated from simulation. The $m_{\ell\ell}$ distribution in the Z + jets samples is found to be different in data and simulation. This is corrected by reweighting the generator-level $p_{\mathrm{T}}^{\ell\ell}$ and $m_{\ell\ell}$ distributions in simulation so that the reconstructed distribution matches the data. The $Z \to \ell\ell$ background is also estimated from simulation. Corrections for $e \to \tau_h$ and $\mu \to \tau_h$ misidentification rates are applied when the reconstructed τ_h candidate is matched to an electron or a muon, respectively, at the generator level. They are measured in Drell-Yan events. The correction factors depend on the pseudorapidity of the lepton. The $t\bar{t}$+jets background is particularly important in the $e\mu$ final state. It is estimated with simulation. A correction based on the generated p_{T} of the top and antitop quarks is applied to match the p_{T} distribution observed in data. The diboson and single top quark backgrounds, as well as the background from SM Higgs boson decays are estimated with simulation. The SM Higgs boson production contributes a small but not negligible background. It arises predominantly from $\mathrm{H} \to \tau\tau$ but also $\mathrm{H} \to \mathrm{WW}$ decays and peaks at lower values of M_{col} than the signal due to additional neutrinos in the decays. The event selection described below uses a multivariate discriminant that combines M_{col} with a set of other kinematic variables. The SM Higgs background also peaks below the signal in the distribution of the mutivariate discriminant output.

Jets misidentified as leptons are a source of background arising from two sources, W + jets and QCD multijet events. In W + jets events one lepton candidate is a real lepton from the W decay and the other is a jet misidentified as a lepton. In QCD multijet events both lepton candidates are misidentified jets. In the Run 1 analysis, a fully data driven technique was used to estimate the misidentified lepton background. It estimates the misidentified lepton background from the two sources combined. In each of the four channels for this analysis ($\mu\tau_h$,eτ_h,$\mu\tau_{\mathrm{e}}$,eτ_{μ}) the misidentified lepton background has been estimated using the same method. In the $\mu\tau_{\mathrm{e}}$ and

eτ_μ channels it is also estimated using a semi-data driven technique used previously in the SM $H \to \tau\tau$ analysis in which the misidentified $W+$jets background is estimated using simulation and the QCD background with data. In this case we have checked that the two techniques are consistent with the loose selection sample. However we use the semi-data driven technique with the full selection in these channels as the fully data driven technique suffers from lack of statistics.

Fully data driven technique

The misidentified lepton background is estimated from collision data by defining a sample with the same selection as the signal sample, but inverting the isolation requirements on one of the leptons, to enrich the contribution from W+jets and QCD multijets. The probability for PF objects to be misidentified as leptons is measured using an independent collision data set, defined below, and this probability is applied to the background enriched sample to compute the misidentified lepton background in the signal sample. The technique is shown schematically in Table 1 in which four regions are defined including the signal (I) and background (III) enriched regions and two control Regions (II & IV), defined with the same selections as Regions I & III respectively, except the leptons are required to have the same charge.

Table 1: Definition of the samples used to estimate the misidentified lepton (ℓ) background. They are defined by the charge of the two leptons and by the isolation requirements on each. The definition of not-isolated differs in each channel.

Region I	**Region II**
$\ell_1^{\pm}$(isolated)	$\ell_1^{\pm}$(isolated)
$\ell_2^{\mp}$(isolated)	$\ell_2^{\pm}$(isolated)
Region III	**Region IV**
$\ell_1^{\pm}$(isolated)	$\ell_1^{\pm}$(isolated)
$\ell_2^{\mp}$(not-isolated)	$\ell_2^{\pm}$(not-isolated)

The rates with which jets are misidentified as electrons (f_e), muons (f_μ), or τ_h (f_τ), are estimated in $Z+$jets events in data. This event sample is obtained by requiring a good Z boson candidate plus one jet. The Z boson candidate is formed from two muons with $p_\mathrm{T} > 26$ GeV, $|\eta| < 2.4$, passing a medium identification requirement, and have $R^{\ell}_{\mathrm{Iso}} < 0.15$. In the case of jet $\to$ e misidentification rate, the muon relative isolation is loosened to $R^{\ell}_{\mathrm{Iso}} < 0.25$ to increase the statistics while keeping a high purity Z sample. The muons are required to have an opposite sign charge and $70 < \mathrm{M}_{\mu\mu} < 110$ GeV. The misidentification rates are measured both in data and in $Z+$jets simulation. The contribution from diboson events, where the jet candidate corresponds to a real lepton, is subtracted using simulation.

The misidentification rates are then defined as

$$f_i = \frac{N_i(\text{region I})}{N_i(\text{region III}) + N_i(\text{region I})}$$

where N_i is the number of events passing the requirements for the regions as defined below and $i = \mathrm{e}, \mu, \tau_h$.

Jets $\to \tau_h$ Estimation

τ_h candidates are defined as not-isolated if they pass a very loose isolation requirement and not pass the tight isolation requirement used for the signal region (region I). The misidentification

rates show a p_T dependence that varies with the τ decay mode and $|\eta|$. The misidentification rates are applied as a function of p_T for the different decay modes and $|\eta|$ regions ($|\eta| < 1.5$ or $|\eta| > 1.5$).

Jets $\rightarrow$ e Estimation

Electron candidates are defined as not-isolated if they pass the tight identification working point and have $0.1 < R^{\ell}_{\text{Iso}} < 0.5$. In the signal region the electrons are required to pass the tight identification working point and have $R^{\ell}_{\text{Iso}} < 0.1$. The misidentification rate shows a dependence on p_T and is flat in $|\eta|$.

Jets $\rightarrow \mu$ Estimation

Muon candidates are defined as not-isolated if they pass a medium muon identification requirement and have $0.15 < R^{\ell}_{\text{Iso}} < 0.25$. In the signal region, muons are required to have $R^{\ell}_{\text{Iso}} < 0.15$. The misidentification rates are applied as a function of p_T.

The number of misidentified background events in the signal region $N_i(regionI)$ is then given by

$$N_i(\text{misidentified}) = \frac{f_i}{1 - f_i} N_i(\text{region III})$$

Double-counted events with two misidentified leptons are subtracted. For example events with a misidentified muon (electron) and a misidentified τ_h are subtracted in the $H \rightarrow \mu\tau_h$ and $H \rightarrow e\tau_h$ channels. This contribution is estimated as

$$N_j(\text{misidentified}) = \frac{f_\tau \cdot f_j}{(1 - f_\tau) \cdot (1 - f_j)} N_j(\text{region III})$$

where $j = e$ or μ. The contribution of diboson events is estimated by applying the method to the corresponding simulation samples, and subtracted. The background estimation is validated in like sign events by following the same procedure but with regions II and IV instead of I and III. The background estimation can also be validated in a W-enriched control region. The region is obtained by selecting opposite sign events with $M_T(\mu/e) > 60$ GeV and $M_T(\tau_h) > 80$ GeV. Selecting opposite sign events better reproduces the jet flavour composition in the signal region. The results of the background validation are given in the next section after describing the event selection.

Semi Data Driven Technique

The W + jets contribution to the misidentified lepton background is estimated with simulation. Simulated samples with different generated jet multiplicities are combined to improve the statistical precision. The QCD multijet contribution is estimated with like sign collision data events that pass the signal region requirement. The expected yield from non-QCD processes is subtracted using simulation. The resulting sample is then rescaled to account for differences between the composition in the like-sign and opposite-sign regions. The scaling factors are extracted from QCD multijet enriched control regions, composed of events with the lepton candidates satisfying relaxed isolation requirements.

6 Event Selection

In each decay mode a loose selection is defined first. The events are then divided into categories within each sample according to the number of jets in the event. This is designed to

enhance the contribution of different Higgs production mechanisms. The jets are required to have $p_T > 30$ GeV and $|\eta| < 4.7$. The 0-jet category enhances the GF contribution, while the 1-jet category enhances GF production with initial state radiation. The 2-jet GF category has a further requirement that the invariant mass of the two jets $M_{jj} < 550$ GeV while the 2-jet VBF category with the requirement $M_{jj} > 550$ GeV enhances the VBF contribution. The threshold on M_{jj} has been optimized to give the best expected exclusion limits. The definition of the categories is that same in both channels except that in the eμ channel the M_{jj} threshold is 500 GeV. The events selected are used to provide a high statistics validation of the background estimation.

After the loose selection a set of eight kinematic variables is combined into a boosted decision tree (BDT). A binned likelihood is used to fit the distributions of the BDT output for the signal and the background contributions. This is referred to as the BDT-fit analysis. As a cross-check an analysis using a tighter set of selection criteria is also presented. In this case selection requirements are placed on the kinematic variables and the M_{col} distribution is fit. This is referred to as the M_{col}-fit analysis. In this case the lepton p_T requirements are constrained to be very loose to avoid the backgrounds being biased to peak under the signal in the M_{col} distribution. This degrades sensitivity compared to the BDT-fit analysis but requirements on additional kinematic variables such as $M_T(\ell)$ are chosen to get the most stringent expected limits with this constraint. In both analysis the modified-frequentist CL_s method [67, 68] is used to set upper bounds on the signal strength μ, or determine a signal significance.

6.1 $H \to \mu\tau_h$

In the loose selection an isolated μ and an isolated τ_h of opposite charge are required to lie in the fiducial region and be separated by $\Delta R > 0.5$. Events with additional electrons, muons, or τ_h candidates or at least one b-jet are vetoed. The lepton candidates are required to have $p_T^{\mu} > 26$ GeV, $|\eta^{\mu}| < 2.4$, $p_T^{\tau_h} > 30$ GeV and $|\eta^{\tau_h}| < 2.3$. The tighter selection used for the M_{col}-fit analysis further requires $M_T(\tau_h) < 105$ GeV in the 0, 1 and 2-jet GF categories and $M_T(\tau_h) < 85$ GeV in the 2-jet VBF category.

A BDT is trained after the loose selection and combining all categories. The signal used is a mixture of simulated GF and VBF events, weighted according to their respective production cross sections. The same signal sample is used in all categories. The background is a collision data sample of misidentified lepton events. This is the dominant background in this channel. The leptons are required to be like sign and be not-isolated in order to be orthogonal to the signal region, yet have the same kinematic properties. The input variables to the BDT are; p_T^{μ}, $p_T^{\tau_h}$, M_{col}, E_T^{miss}, $M_T(\tau_h)$, $\Delta\eta(\mu, \tau_h)$, $\Delta\phi(p_T^{\mu}, p_T^{\tau_h})$ and $\Delta\phi(p_T^{\tau_h}, \vec{E}_T^{miss})$.

The control region tests of the misidentified lepton backgrounds described in section 5 are performed after the loose selection. The τ_h and μ candidate isolation requirements are relaxed to enhance statistics in a like-sign control region. Figure 1 shows the data compared to the background estimation in the control region. A 30% uncertainty is assigned to the misidentified lepton background estimation. This uncertainty accounts for uncertainties in the misidentification rate, and for differences of background composition or jet flavour. The background estimate is cross-checked using another W + jets enriched control region and the agreement is comparable. The same samples and control regions are used in the $H \to e\tau_h$ channel with similar agreement.

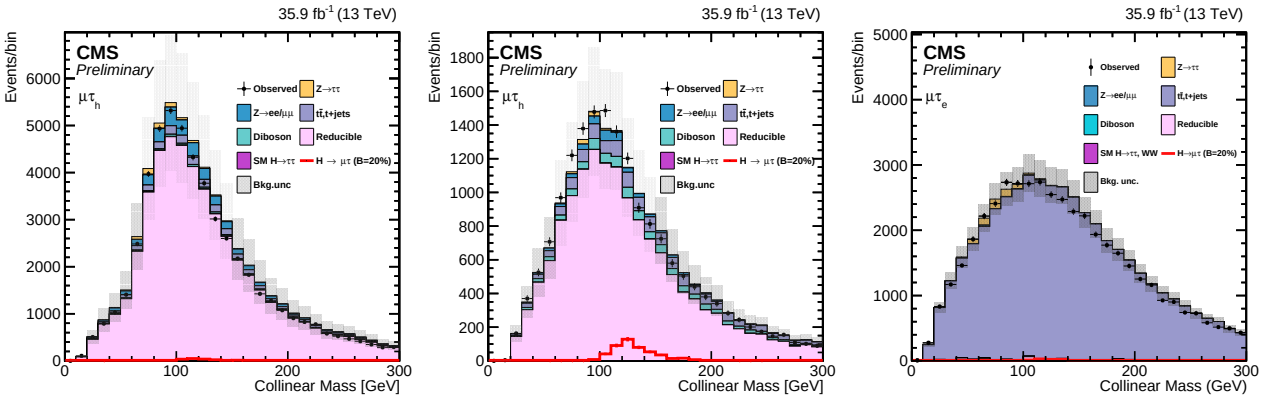
6.2 $H \to \mu\tau_e$

In the loose selection an isolated μ and an isolated e of opposite charge are required to lie in the fiducial region and be separated by $\Delta R > 0.3$. Events with additional electrons, muons, or τ_h candidates are rejected. Events with at least one b-jet are also rejected to reduce the $t\bar{t}$ background contribution. The muon is then required to have $p_T^\mu > 26$ GeV, $|\eta^\mu| < 2.4$, and $R_{Iso}^\ell < 0.15$. The electron is required to have $p_T^e > 10$ GeV, $|\eta^e| < 2.3$, and $R_{Iso}^\ell < 0.1$.

The tighter selection used in the M_{col}-fit analysis requires $p_T^\mu > 30$ GeV for the 0-jet category and $p_T^\mu > 26$ GeV in the other categories. In the 0-jet, 1-jet , 2-jets GF and 2-jets VBF categories the transverse mass $M_T(\mu)$ is required to be greater than 60, 40, 15 and 15 GeV respectively. The neutrinos from the τ lepton decay are approximately collinear to the electron direction so an additional requirement is made on the azimuthal angle between the electron and the $\vec{E}_T^{miss}$: $\Delta\phi(p_T^e, \vec{E}_T^{miss}) < 0.7, 0.7, 0.5, 0.3$ for the 0-jet, 1-jet, 2-jets GF and 2-jets VBF categories respectively. In the 0-jet and 1-jet category it is further required that $\Delta\phi(p_T^e, p_T^\mu) > 2.5, 1.0$.

A BDT is trained after the loose selection, combining all categories. The background is a mixed sample of $t\bar{t}$ and Drell-Yan events weighted by their cross-sections. The $t\bar{t}$ background is the dominant background in this channel for the 2-jets category and also very significant in 1-jet. It has many kinematic characteristics in common with the other backgrounds such as diboson and single top. The Drell-Yan background is the dominant background in 0-jet and 1-jet category. The input variables of the BDT are; p_T^μ, p_T^e, M_{col}, E_T^{miss}, $M_T(\mu)$, $M_T(e)$, $\Delta\phi(p_T^e, p_T^\mu)$, $\Delta\phi(p_T^e, \vec{E}_T^{miss})$ and $\Delta\phi(p_T^\mu, \vec{E}_T^{miss})$.

The control region tests of the misidentified lepton backgrounds described in section 5 are performed after the loose selection. The misidentified lepton background in this channel is much less important than in the hadronic final states. Figure 1 shows the data compared to the background estimation in the $t\bar{t}$ control region. The control region is the 2-jet selection but with the additional requirement that at least one of the jets is b-tagged in order to enhance the $t\bar{t}$ contribution The same sample and control region are used in the $H \to e\tau_\mu$ channel, except with $\mu \to e$, and give similar agreement.



(a) Like-sign lepton region (b) W + jets enriched region (c) $t\bar{t}$ enriched region

Figure 1: M_{col} in different control regions defined in the text. The distributions are pre-fit and include both statistical and systematic uncertainties.

6.3 $H \to e\tau_h$

The electron candidates are required to have $p_T^e > 26$ GeV, $|\eta^e| < 2.1$ and $R_{Iso}^{\ell} < 0.1$. The τ_h candidates are required to have $p_T^{\tau_h} > 30$ GeV and $|\eta^{\tau_h}| < 2.3$ and to pass a tight isolation requirement. In each event, the electron and the τ_h candidate are required to have opposite sign charge, and to be separated by $\Delta R > 0.5$. Events with additional electrons, muons, or τ_h candidates are rejected. No requirement is made on the number of b-tagged jets as the $t\bar{t}$ contribution is small. The tighter selection used for the M_{col}-fit analysis further requires that $M_T(\tau_h) < 60$ GeV. A BDT is trained after the loose selection. The same training samples as the $H \to \mu\tau_h$ channel are used, except with $\mu \to e$. The input variables to the BDT are also the same except for the addition of the visible mass, M_{vis}.

6.4 $H \to e\tau_\mu$

In the loose selection the electron candidates are required to have $p_T^e > 26$ GeV, $|\eta^e| < 2.1$ and $R_{Iso}^{\ell} < 0.15$. Muon candidates are required to have $p_T^{\mu} > 10$ GeV, $|\eta^{\mu}| < 2.4$ and $R_{Iso}^{\ell} < 0.1$. The electron and the muon candidates are required to have an opposite sign charge, and to be separated by $\Delta R > 0.5$. Events with additional electrons, muons, or τ_h candidates are rejected. No requirement is made on the number of b-tagged jets.

The tighter selection used in the M_{col}-fit analysis further requires $\Delta\phi(p_T^e, \vec{E}_T^{miss}) < 1.0$ and $M_T(\mu) > 60$ GeV. The large $t\bar{t}$ background is reduced by vetoing events with b-tagged jets, and requiring $P_\zeta > -60$ GeV. P_ζ is calculated as the difference of the projections of visible transverse momentum of τ decay products plus missing transverse momentum, $\vec{P}_{T,1}^{vis} + \vec{P}_{T,2}^{vis} + \vec{P}_T^{mis}$, and the visible transverse momentum of τ decay products on the axis ζ. Here, ζ is the axis bisecting the directions $\vec{P}_{T,1}^{vis}$ and $\vec{P}_{T,2}^{vis}$ of the electron and the muon in the transverse plane. The definition is the following:

$$P_\zeta = P_\zeta^{mis} - 1.85 P_\zeta^{vis}$$
$$\text{with } P_\zeta^{mis} = (\vec{P}_{T,1}^{vis} + \vec{P}_{T,2}^{vis} + \vec{P}_T^{mis}) \frac{\vec{\zeta}}{|\vec{\zeta}|}$$
$$\text{and } P_\zeta^{vis} = (\vec{P}_{T,1}^{vis} + \vec{P}_{T,2}^{vis}) \frac{\vec{\zeta}}{|\vec{\zeta}|}$$

A BDT is trained after the loose selection. It uses the same input variables as for the $H \to \mu\tau_e$ channel with the addition of the visible mass, M_{vis}. The background is a sample of simulated $t\bar{t}$ events. The $t\bar{t}$ background is the dominant background in this channel for the 2-jet category and very significant in the 1-jet category. It has many kinematic characteristics in common with other backgrounds such as diboson and single top.

7 Systematic uncertainties

The systematic uncertainties affect the normalization or the shape of the distribution of the different processes, and arise from experimental or theoretical sources. They are summarized in Table 2. The uncertainties in the lepton (e, μ, τ_h) selection including the trigger, identification, and isolation efficiencies are estimated using tag and probe measurements in collision datasets of Z bosons decaying to $ee, \mu\mu, \tau_\mu\tau_h$ [59–62]. The b-tagging efficiency is measured in collision data and the simulation is corrected accordingly with the measurement uncertainty as the systematic error. The uncertainties on the $Z \to ee, \mu\mu, \tau\tau$, WW, ZZ, W$\gamma$,$t\bar{t}$ and single top

Table 2: The systematic uncertainties in the expected event yield. All uncertainties are treated as correlated between the categories, except those which have two values. In this case the first value is the correlated uncertainty and the second value is the uncorrelated uncertainty for each individual category. Anticorrelations arise due to migration of events between the categories and are expressed as negative numbers.

Systematic uncertainty	$H \to \mu\tau_e$	$H \to \mu\tau_h$	$H \to e\tau_\mu$	$H \to e\tau_h$
Muon trigger/ID/isolation	2%	2%	2%	-
Electron trigger/ID/isolation	2%	-	2%	2%
Hadronic τ efficiency	-	5%	-	5%
b-tagging veto	2.0–4.5%	2.0–4.5%	2.0–4.5%	-
$Z \to \mu\mu/ee$ +jets background	10%⊕5%	-	10%⊕5%	-
$Z \to \tau\tau$ +jets background	10%⊕5%	10%⊕5%	10%⊕5%	10%⊕5%
W + jets background	10%	-	10%	-
QCD multijet background	30%	-	30%	-
WW, ZZ background	5%⊕5%	5%⊕5%	5%⊕5%	5%⊕5%
$t\bar{t}$ background	10%⊕5%	10%⊕5%	10%⊕5%	10%⊕5%
$W + \gamma$ background	10%⊕5%	-	10%⊕5%	-
Single top production background	5%⊕5%	5%⊕5%	5%⊕5%	5%⊕5%
$\mu \to \tau_h$ background	-	25%	-	-
$e \to \tau_h$ background	-	-	-	12%
jet $\to \tau_h, \mu, e$ background	-	30%⊕10%	-	30%⊕10%
Jet energy scale	3–20%	3–20%	3-20%	3–20%
Hadronic τ energy scale	-	1.2%	-	1.2%
$e \to \tau_h$ energy scale	-	1.5%	-	3%
Electron energy scale	$\pm\sigma$	-	$\pm\sigma$	$\pm\sigma$
Muon energy scale	0.2%	0.2%	-	$\pm\sigma$
Unclustered energy scale	$\pm\sigma$	$\pm\sigma$	$\pm\sigma$	$\pm\sigma$
acceptance scale (GF H)	-3.0 – 2.0%	-3.0 – 2.0%	-3.0 – 2.0%	-3.0 – 2.0%
acceptance scale (VBF H)	-0.3 – 1.0%	-0.3 – 1.0%	-0.3 – 1.0%	-0.3 – 1.0%
QCD scale YR4 (GF H)	3.2%	3.2%	3.2%	3.2%
QCD scale YR4 (VBF H)	2.1%	2.1%	2.1%	2.1%
acceptance PDF (GF H)	-1.5 – 0.5%	-1.5 – 0.5%	-1.5 – 0.5%	-1.5 – 0.5%
acceptance PDF (VBF H)	-1.5 – 1.0%	-1.5 – 1.0%	-1.5 – 1.0%	-1.5 – 1.0%
PDF YR4 (GF H)	3.9%	3.9%	3.9%	3.9%
PDF YR4 (VBF H)	0.4%	0.4%	0.4%	0.4%
Bin-by-bin	Shape	Shape	Shape	Shape
Luminosity	2.5%	2.5%	2.5%	2.5%
Pile-up	Shape	Shape	Shape	Shape

production background contributions arise predominantly from the measured cross sections of these processes. The uncertainties in the estimation of the misidentified lepton backgrounds are from the validation tests in control regions, described in section 6.

Shape and normalization uncertainties arising from the uncertainty in the jet energy scale are computed by propagating to the fit templates of each process the effect of altering each source of jet energy scale uncertainty by $\pm 1\sigma$. This takes into account differences in yield and shape. There are 27 sources of jet energy scale uncertainty considered independently, fully correlated between categories and τ decay channels. The uncertainty on the τ_h energy scale is treated independently for each reconstructed decay mode. It is propagated to the collinear mass and BDT distributions. The uncertainty in the energy scale of electrons and muons misidentified as taus is propagated to the M_{col} and BDT distributions. Systematic uncertainties on the electron energy scale and resolution include the effects of electron selection efficiency, pseudorapidity dependence, and categorization summed in quadrature. The resolution systematics have negligible effect. There is also an uncertainty in the unclustered energy scale. The unclustered energy comes from jets below 10 GeV and PF candidates not within jets. It is also propagated to $E_{\mathrm{T}}^{\mathrm{miss}}$. The unclustered energy scale is considered independently for charged particles, photons, neutral hadrons, and very forward particles which are not contained in jets. The effect of shifting the energy of each particle by its uncertainty leads to both changes in shape of the distribution and in the yield. The four different systematic uncertainties are uncorrelated.

The uncertainties on the Higgs production cross sections cause changes in acceptance when the factorization and the renormalization scale are shifted. They affect the normalization and result in migration between categories. Normalization uncertainties taken from Ref. [69] are from the renormalization and factorization scale as well as the PDF and α_s variations. They depend on the production process and are correlated between categories. It is estimated from simulation, and depends on the production process, Higgs boson decay channel and category. It is correlated between categories.

The bin-by-bin uncertainties account for the statistical uncertainties in every bin of the template distributions of every process. They are uncorrelated between bins, processes, and categories. The uncertainty on the integrated luminosity affects all processes with normalization taken directly from simulation. Shape uncertainties related to the pile up have been considered varying by 5% the minimum bias cross section in the computation of the pileup events in data. The new values are then used to compute the weights for the simulation samples and these are applied, event by event, to produced alternative collinear mass and BDT distributions used as shape uncertainties in the fit.

8 Results

After applying the selection criteria, a maximum likelihood fit is performed in a final discrimination variable to derive the expected and observed limits. Each systematic uncertainty is used as a nuisance parameter in the fit. The fits are performed simultaneously in all channels and categories. The modified-frequentist CL_s method [67, 68] is used to set upper bounds on the signal strength, or determine a signal significance.

The M_{col}-fit distributions for each category compared to the signal and background contributions are shown in Fig. 2 and Fig. 5 in the $\mathrm{H} \to \mu\tau$ and $\mathrm{H} \to \mathrm{e}\tau$ channels respectively. Figures 3 and 6 show the corresponding BDT-fit distributions. No excess over the background expectation is observed. The observed and median expected 95% CL upper limits, and best fit branching fractions, for $\mathcal{B}(\mathrm{H} \to \mu\tau)$ and $\mathcal{B}(\mathrm{H} \to \mu\tau)$, assuming M_{H}=125 GeV, are given for

each category in Tables 3, 4, 5 and 6. The limits are also summarized graphically in Figures 4 and 7.

Table 3: The expected and observed upper limits at 95% CL, and best fit branching fractions in percent for the different jet categories for the H $\rightarrow \mu\tau$ process obtained with the M_{col}-fit analysis.

Expected limits (%)					
	0-jet	1-jet	2-jets	VBF	Combined
$\mu\tau_e$	< 1.01	< 1.47	< 3.23	< 1.73	< 0.75
$\mu\tau_h$	< 1.14	< 1.26	< 2.12	< 1.41	< 0.71
$\mu\tau$	< 0.49				
Observed limits (%)					
	0-jet	1-jet	2-jets	VBF	Combined
$\mu\tau_e$	< 1.08	< 1.35	< 3.33	< 1.40	< 0.71
$\mu\tau_h$	< 1.04	< 1.74	< 1.65	< 1.30	< 0.66
$\mu\tau$	< 0.51				
Best fit branching fractions (%)					
	0-jet	1-jet	2-jets	VBF	Combined
$\mu\tau_e$	0.13 ± 0.43	-0.22 ± 0.75	0.22 ± 1.39	-1.73 ± 1.05	-0.04 ± 0.33
$\mu\tau_h$	-0.30 ± 0.45	0.68 ± 0.56	-1.23 ± 1.04	-0.23 ± 0.66	-0.08 ± 0.34
$\mu\tau$	0.02 ± 0.20				

Table 4: The expected and observed upper limits at 95% CL, and the best fit branching fractions in percent for each individual jet category, and combined, in the H $\rightarrow \mu\tau$ process obtained with the BDT-fit analysis.

Expected limits (%)					
	0-jet	1-jet	2-jets	VBF	Combined
$\mu\tau_e$	< 0.83	< 1.19	< 1.98	< 1.62	< 0.59
$\mu\tau_h$	< 0.43	< 0.56	< 0.94	< 0.58	< 0.29
$\mu\tau$	< 0.25				
Observed limits (%)					
	0-jet	1-jet	2-jets	VBF	Combined
$\mu\tau_e$	< 1.30	< 1.34	< 2.27	< 1.79	< 0.86
$\mu\tau_h$	< 0.51	< 0.53	< 0.56	< 0.51	< 0.27
$\mu\tau$	< 0.25				
Best fit branching fractions (%)					
	0-jet	1-jet	2-jets	VBF	Combined
$\mu\tau_e$	0.61 ± 0.36	0.22 ± 0.46	0.39 ± 0.83	0.10 ± 1.37	0.35 ± 0.26
$\mu\tau_h$	0.12 ± 0.20	-0.05 ± 0.25	-0.72 ± 0.43	-0.22 ± 0.31	-0.04 ± 0.14
$\mu\tau$	0.00 ± 0.12				

No evidence is found for either the H $\rightarrow \mu\tau$ or H $\rightarrow e\tau$ processes in the the 35.9 fb^{-1} dataset at 13 TeV. The observed exclusion limits are a significant improvement over 8 TeV results [27, 28] and the early 13 TeV 2015 result [70]. The new results exclude the branching fraction that corresponded to the best fit for the 2.4 σ excess observed in the 8 TeV H $\rightarrow \mu\tau$ channel results at 95% CL, in both the M_{col}-fit and BDT-fit analysis. Table 7 shows a summary of the new 95%

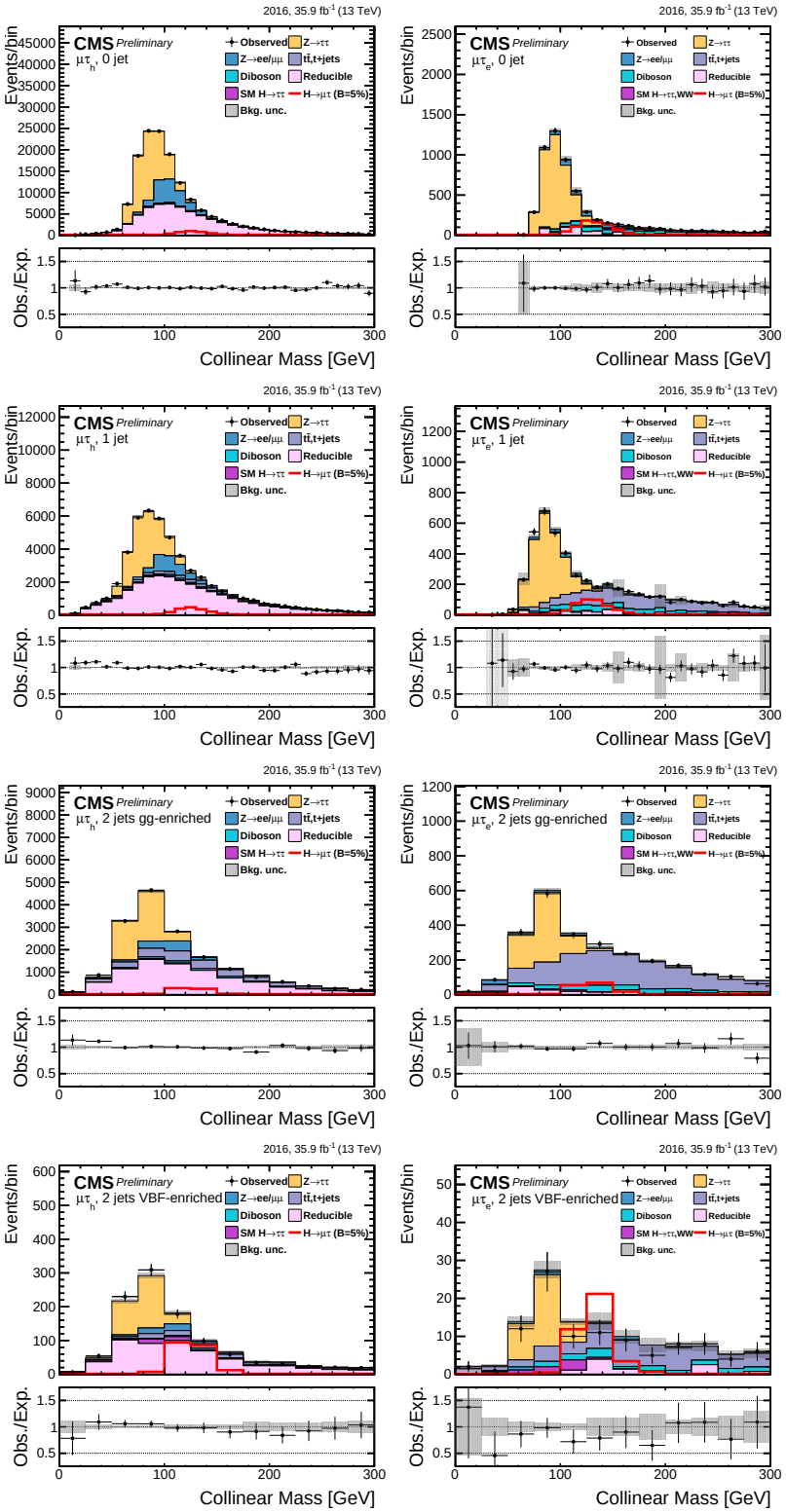


Figure 2: Distribution of the collinear mass M_{col} for the H $\rightarrow \mu\tau$ process in M_{col}-fit analysis, in the different channels and categories compared to the signal and background estimation. The background is normalized to the best-fit values from the signal plus background fit while the overlayed simulated signal corresponds to $\mathcal{B}(\mathrm{H} \rightarrow \mu\tau) = 5\%$. The bottom panel in each plot shows the fractional difference between the observed data and the fitted background. The left column of plots corresponds to the H $\rightarrow \mu\tau_h$ categories, from 0-jets (first row) to 2-jets VBF (fourth row). The right one to their H $\rightarrow \mu\tau_e$ counterparts.

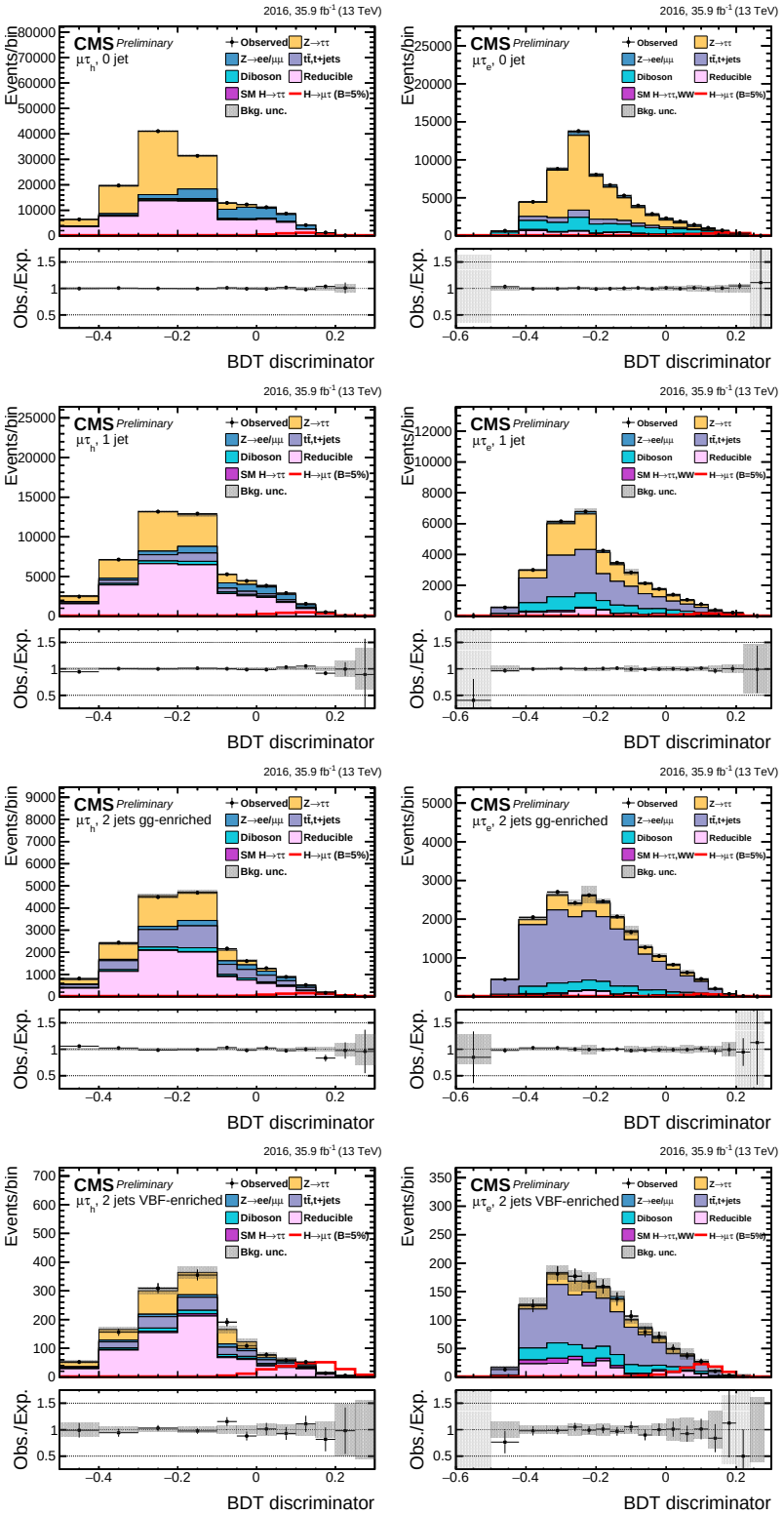


Figure 3: Distribution of the BDT output for the $H \to \mu\tau$ process in the BDT-fit analysis, in the different channels and categories compared to the signal and background estimation. The background is normalized to the best-fit values from the signal plus background fit while the simulated signal corresponds to $\mathcal{B}(H \to \mu\tau) = 5\%$. The bottom panel in each plot shows the fractional difference between the observed data and the fitted background. The left column of plots corresponds to the $H \to \mu\tau_h$ categories, from 0-jets (first row) to 2-jets VBF (fourth row). The right one to their $H \to \mu\tau_e$ counterparts.

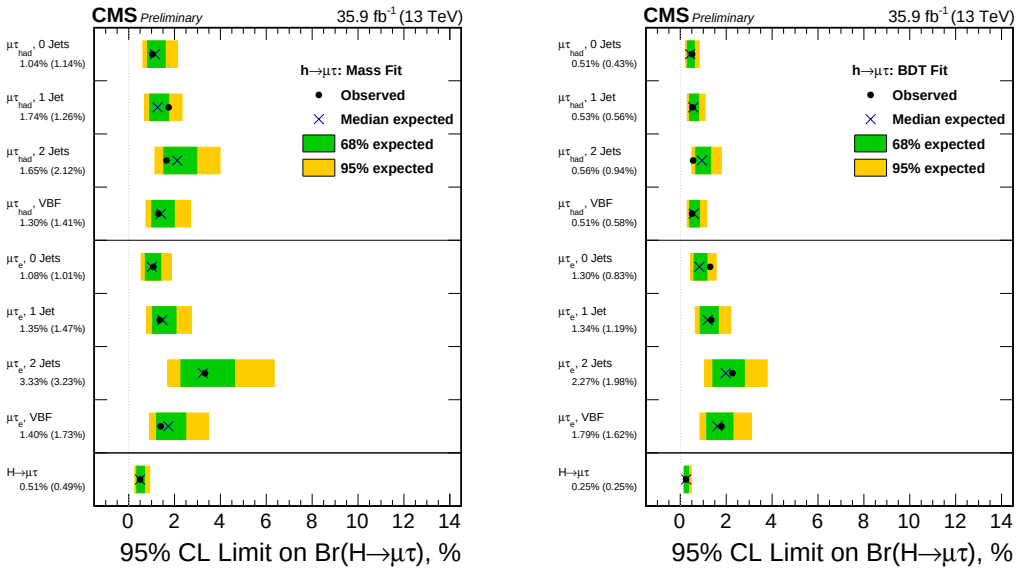


Figure 4: Observed and expected 95% CL upper limits on the $\mathcal{B}(\mathrm{H} \to \mu\tau)$ for each individual category and combined. Left: M_{col}-fit analysis. Right: BDT-fit analysis.

Table 5: The expected and observed upper limits at 95% CL and best fit branching fractions in percent for each individual jet category, and combined, in the $\mathrm{H} \to \mu\tau$ process obtained with the BDT-fit analysis.

Expected limits (%)					
	0-jet	1-jet	2-jets	VBF	Combined
$e\tau_\mu$	< 0.94	< 1.21	< 3.73	< 2.76	< 0.71
$e\tau_h$	< 1.52	< 1.93	< 3.55	< 1.76	< 0.97
$e\tau$	< 0.56				

Observed limits (%)					
	0-jet	1-jet	2-jets	VBF	Combined
$e\tau_\mu$	< 1.27	< 1.26	< 3.90	< 1.78	< 0.85
$e\tau_h$	< 1.53	< 2.07	< 3.65	< 3.39	< 1.31
$e\tau$	< 0.72				

Best fit branching fractions (%)					
	0-jet	1-jet	2-jets	VBF	Combined
$e\tau_\mu$	0.46 ± 0.43	0.07 ± 0.39	0.13 ± 1.13	-1.38 ± 1.03	0.21 ± 0.36
$e\tau_h$	0.18 ± 0.35	0.45 ± 0.60	0.29 ± 1.13	2.03 ± 0.47	0.51 ± 0.41
$e\tau$	0.23 ± 0.24				

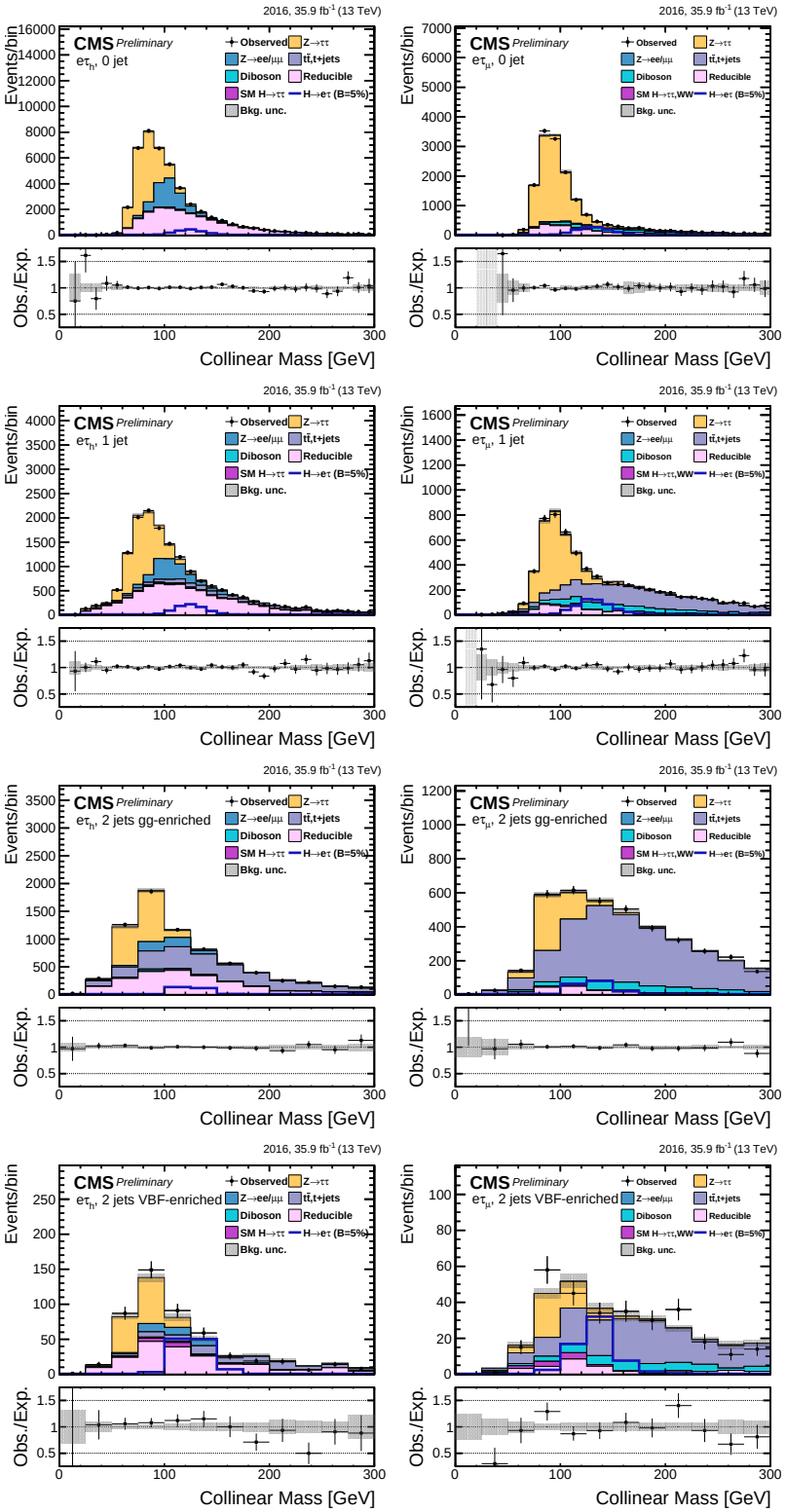


Figure 5: Distribution of the collinear mass M_{col} for the $H \to e\tau$ process in the M_{col}-fit analysis, in the different channels and categories compared to the signal and background estimation. The background is normalized to the best-fit values from the signal plus background fit while the simulated signal corresponds to $\mathcal{B}(H \to e\tau) = 5\%$. The lower panel in each plot shows the fractional difference between the observed data and the fitted background. The left column of plots correspond to the $H \to e\tau_h$ categories, from 0-jets (first row) to 2 jets VBF (fourth row). The right one to their $H \to e\tau_\mu$ counterparts.

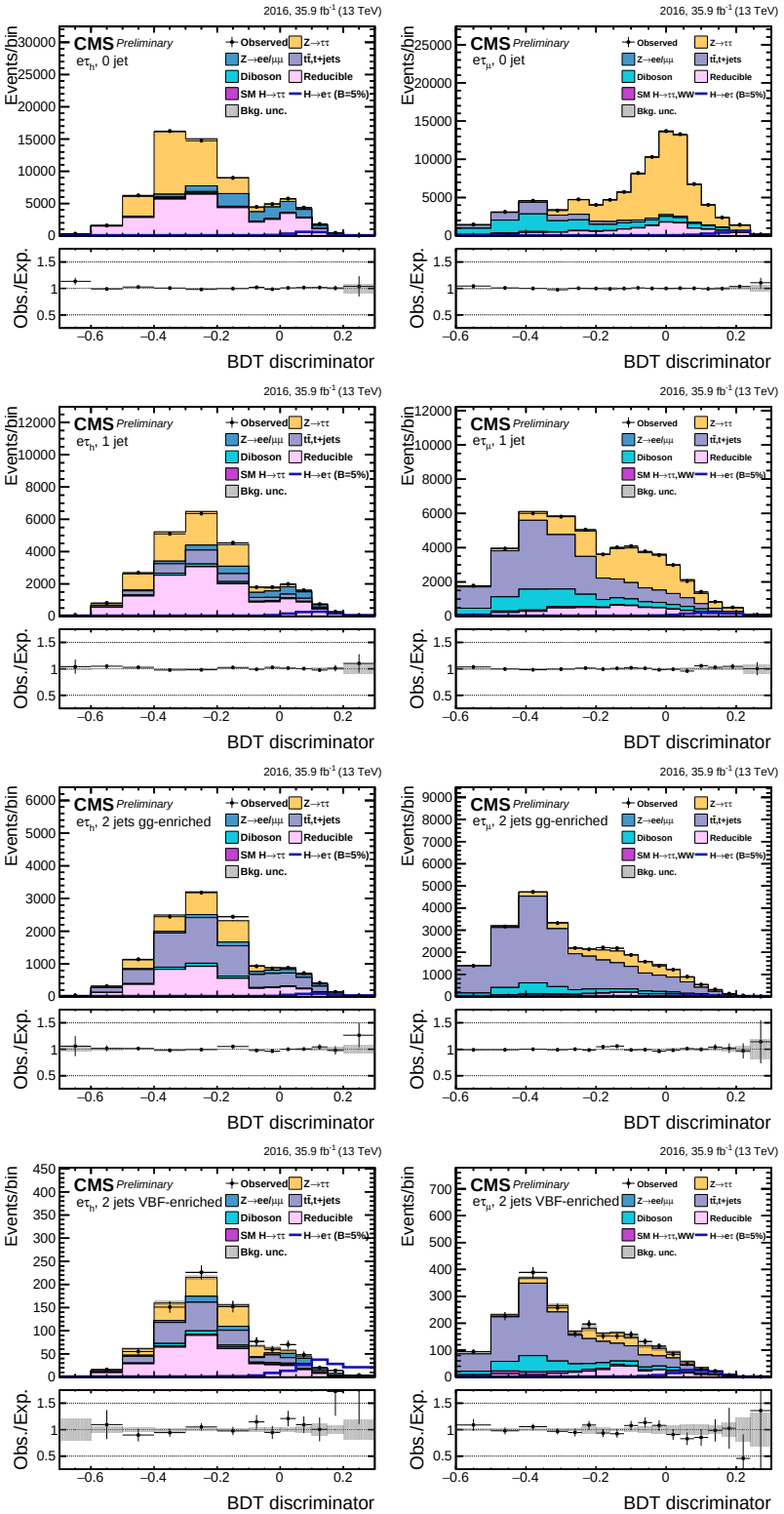


Figure 6: Distribution of the BDT output for the $H \to e\tau$ process for the BDT-fit analysis, in the different channels and categories compared to the signal and background estimation. The background is normalized to the best-fit values from the signal plus background fit while the simulated signal corresponds to $\mathcal{B}(H \to e\tau) = 5\%$. The bottom panel in each plot shows the fractional difference between the observed data and the fitted background. The left column of plots corresponds to the $H \to e\tau_h$ categories, from 0-jets (first row) to 2-jets VBF (fourth row). The right one to their $H \to e\tau_\mu$ counterparts.

Table 6: The observed and expected upper limits at 95% CL and the best fit branching fractions in percent for the different jet categories in the H $\rightarrow$ eτ process obtained with the BDT-fit analysis.

Expected limits (%)					
	0-jet	1-jet	2-jets	VBF	Combined
eτ_μ	< 0.90	< 1.59	< 2.54	< 1.84	< 0.64
eτ_h	< 0.79	< 1.13	< 1.59	< 0.74	< 0.49
eτ	< 0.37				

Observed limits (%)					
	0-jet	1-jet	2-jets	VBF	Combined
eτ_μ	< 1.22	< 1.66	< 2.25	< 1.10	< 0.78
eτ_h	< 0.73	< 0.81	< 1.94	< 1.49	< 0.72
eτ	< 0.61				

Best fit branching fractions (%)					
	0-jet	1-jet	2-jets	VBF	Combined
eτ_μ	0.47 ± 0.42	0.17 ± 0.79	-0.42 ± 1.01	-1.54 ± 0.44	0.18 ± 0.32
eτ_h	-0.13 ± 0.39	-0.63 ± 0.40	0.54 ± 0.53	0.70 ± 0.38	0.33 ± 0.24
eτ	0.30 ± 0.18				

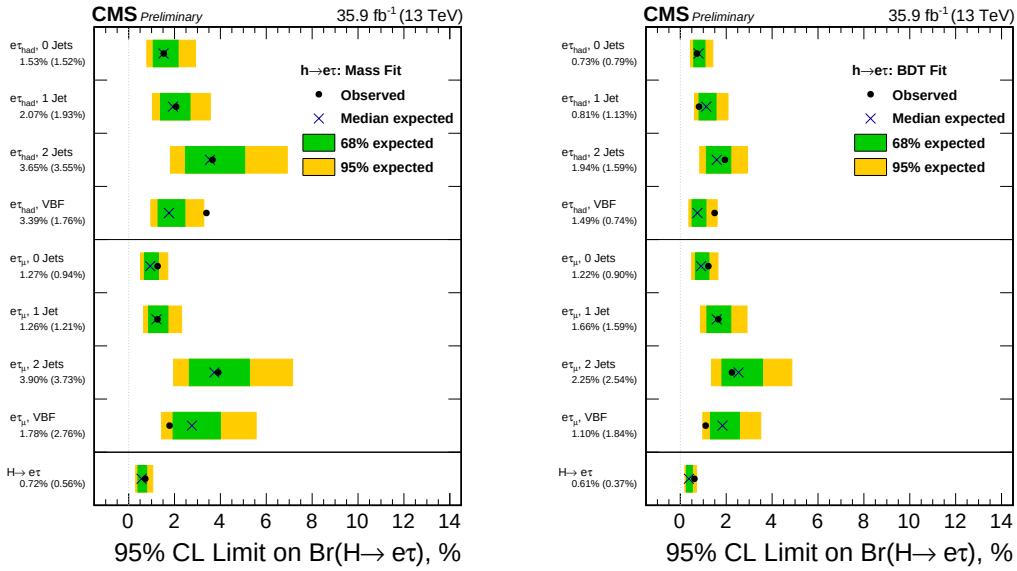


Figure 7: Observed and expected 95% CL upper limits on the $\mathcal{B}$(H $\rightarrow$ eτ) for each individual category and combined. Left: M_{col}-fit analysis. Right: BDT-fit analysis.

CL upper limits. The BDT-fit analysis is more sensitive than the M_{col}-fit analysis, with limits reduced by about a factor two.

Table 7: The observed and expected upper limits at the 95% CL and the best fit branching fractions in percent for the $H \to \mu\tau$ and $H \to e\tau$ processes, with the different selections.

	Observed(Expected) limits (%)		Best fit branching fraction (%)	
	M_{col}-fit	BDT-fit	M_{col}-fit	BDT-fit
$H \to \mu\tau$	<0.51 (0.49) %	<0.25 (0.25)%	0.02 ± 0.20%	0.00 ± 0.12 %
$H \to e\tau$	<0.72 (0.56) %	<0.61 (0.37) %	0.23 ± 0.24 %	0.30 ± 0.18 %

The constraints on $\mathcal{B}(H \to \mu\tau)$ and $\mathcal{B}(H \to e\tau)$ can be interpreted in terms of LFV Yukawa couplings [34]. The LFV decays $e\tau$ and $\mu\tau$ arise at tree level from the assumed flavour violating Yukawa interactions, $Y_{\ell^\alpha \ell^\beta}$ where ℓ^α, ℓ^β denote the leptons, $\ell^\alpha, \ell^\beta = e, \mu, \tau$ and $\ell^\alpha \neq \ell^\beta$. The decay width $\Gamma(H \to \ell^\alpha \ell^\beta)$ in terms of the Yukawa couplings is given by:

$$\Gamma(H \to \ell^\alpha \ell^\beta) = \frac{m_H}{8\pi}\left(|Y_{\ell^\beta \ell^\alpha}|^2 + |Y_{\ell^\alpha \ell^\beta}|^2\right),$$

and the branching fraction by:

$$B(H \to \ell^\alpha \ell^\beta) = \frac{\Gamma(H \to \ell^\alpha \ell^\beta)}{\Gamma(H \to \ell^\alpha \ell^\beta) + \Gamma_{SM}}.$$

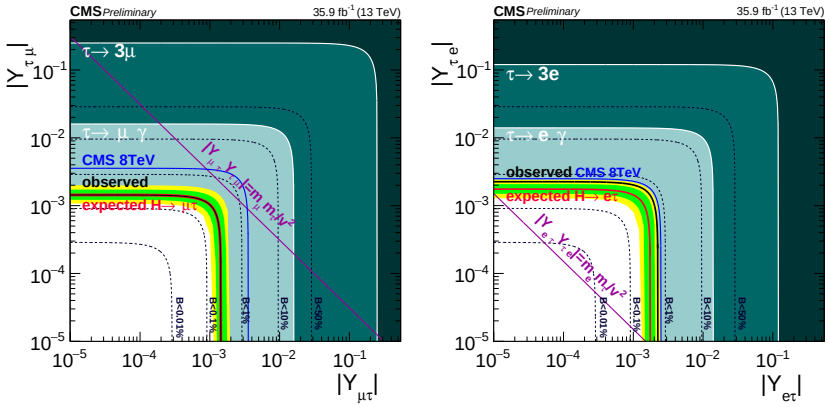
The SM H decay width is assumed to be $\Gamma_{SM} = 4.1\,\text{MeV}$ [71] for $M_H = 125\,\text{GeV}$. The 95% CL upper limit on the Yukawa couplings derived from the expression for the branching fraction above is shown in Table 8. The limits on the Yukawa couplings derived from the BDT-fit analysis results are shown in Figure 8.

Table 8: 95% CL upper limit on the Yukawa couplings

	M_{col}-fit	BDT-fit
$\sqrt{\lvert Y_{\mu\tau}\rvert^2 + \lvert Y_{\tau\mu}\rvert^2}$	$< 2.05 \times 10^{-3}$	$< 1.43 \times 10^{-3}$
$\sqrt{\lvert Y_{e\tau}\rvert^2 + \lvert Y_{\tau e}\rvert^2}$	$< 2.45 \times 10^{-3}$	$< 2.26 \times 10^{-3}$

9 Summary

This article presents the search for LFV decays of the Higgs boson in the $\mu\tau$ and $e\tau$ final states, with the 2016 data collected by the CMS detector. The dataset analyzed corresponds to an integrated luminosity of $35.9\,\text{fb}^{-1}$ of proton-proton collision data recorded at $\sqrt{s} = 13\,\text{TeV}$. The results are extracted by a fit to the output of a BDT trained to discriminate the signal from backgrounds. The results are cross-checked with alternate analysis that fits the M_{col} distribution after applying selection criteria on kinematic variables. No evidence is found for LFV Higgs boson decays. The observed (expected) limits on the branching fraction of the Higgs boson to $\mu\tau$ and to $e\tau$ are found to be less than 0.25(0.25)% and 0.61(0.37)%, respectively, at 95% confidence level, and constitute a significant improvement with respect to the previously obtained limits by CMS and ATLAS using $20\,\text{fb}^{-1}$ of 8 TeV proton-proton collision data. Upper limits on the off-diagonal $\mu\tau$ and $e\tau$ Yukawa couplings are derived from these constraints on the branching ratios, and found to be $\sqrt{|Y_{\mu\tau}|^2 + |Y_{\tau\mu}|^2} < 1.43 \times 10^{-3}$ and $\sqrt{|Y_{e\tau}|^2 + |Y_{\tau e}|^2} < 2.26 \times 10^{-3}$ at 95% CL.



(a) Constraints on $|Y_{\mu\tau}|, |Y_{\tau\mu}|$ (b) Constraints on $|Y_{e\tau}|, |Y_{\tau e}|$

Figure 8: Constraints on the flavour violating Yukawa couplings, $|Y_{\mu\tau}|, |Y_{\tau\mu}|$ and $|Y_{e\tau}|, |Y_{\tau e}|$, from the BDT result. The expected (red solid line) and observed (black solid line) limits are derived from the limit on $B(\mathrm{H} \to \mu\tau)$ and $B(\mathrm{H} \to \mathrm{e}\tau)$ from the present analysis. The flavour diagonal Yukawa couplings are approximated by their SM values. The green (yellow) band indicates the range that is expected to contain 68% (95%) of all observed limit excursions from the expected limit. The shaded regions are derived constraints from null searches for $\tau \to 3\mu$ or $\tau \to 3\mathrm{e}$ (dark green) and $\tau \to \mu\gamma$ or $\tau \to \mathrm{e}\gamma$ (lighter green).The purple diagonal line is the theoretical naturalness limit $Y_{ij}Y_{ji} \leq m_i m_j / v^2$.

References

[1] 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.

[2] 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.

[3] CMS Collaboration, "Observation of a new boson with mass near 125 GeV in pp collisions at $\sqrt{s}$ = 7 and 8 TeV", *JHEP* **06** (2013) 081, doi:10.1007/JHEP06(2013)081, arXiv:1303.4571.

[4] J. D. Bjorken and S. Weinberg, "Mechanism for Nonconservation of Muon Number", *Phys. Rev. Lett.* **38** (1977) 622, doi:10.1103/PhysRevLett.38.622.

[5] J. L. Diaz-Cruz and J. J. Toscano, "Lepton flavor violating decays of Higgs bosons beyond the standard model", *Phys. Rev. D* **62** (2000) 116005, doi:10.1103/PhysRevD.62.116005, arXiv:hep-ph/9910233.

[6] T. Han and D. Marfatia, "$h \to \mu\tau$ at Hadron Colliders", *Phys. Rev. Lett.* **86** (2001) 1442, doi:10.1103/PhysRevLett.86.1442, arXiv:hep-ph/0008141.

[7] E. Arganda, A. M. Curiel, M. J. Herrero, and D. Temes, "Lepton flavor violating Higgs boson decays from massive seesaw neutrinos", *Phys. Rev. D* **71** (2005) 035011, doi:10.1103/PhysRevD.71.035011, arXiv:hep-ph/0407302.

[8] A. Arhrib, Y. Cheng, and O. C. W. Kong, "Comprehensive analysis on lepton flavor violating Higgs boson to $\mu\bar{\tau} + \tau\bar{\mu}$ decay in supersymmetry without R parity", *Phys. Rev. D* **87** (2013) 015025, doi:10.1103/PhysRevD.87.015025, arXiv:1210.8241.

[9] M. Arana-Catania, E. Arganda, and M. J. Herrero, "Non-decoupling SUSY in LFV Higgs decays: a window to new physics at the LHC", *JHEP* **09** (2013) 160, doi:10.1007/JHEP09(2013)160, arXiv:1304.3371. [Erratum: doi:10.1007/JHEP10(2015)192].

[10] E. Arganda, M. J. Herrero, R. Morales, and A. Szynkman, "Analysis of the $h, H, A \to \tau\mu$ decays induced from SUSY loops within the Mass Insertion Approximation", *JHEP* **03** (2016) 055, doi:10.1007/JHEP03(2016)055, arXiv:1510.04685.

[11] E. Arganda, M. J. Herrero, X. Marcano, and C. Weiland, "Enhancement of the lepton flavor violating Higgs boson decay rates from SUSY loops in the inverse seesaw model", *Phys. Rev. D* **93** (2016), no. 5, 055010, doi:10.1103/PhysRevD.93.055010, arXiv:1508.04623.

[12] K. Agashe and R. Contino, "Composite Higgs-mediated flavor-changing neutral current", *Phys. Rev. D* **80** (2009) 075016, doi:10.1103/PhysRevD.80.075016, arXiv:0906.1542.

[13] A. Azatov, M. Toharia, and L. Zhu, "Higgs mediated flavor changing neutral currents in warped extra dimensions", *Phys. Rev. D* **80** (2009) 035016, doi:10.1103/PhysRevD.80.035016, arXiv:0906.1990.

[14] H. Ishimori et al., "Non-Abelian Discrete Symmetries in Particle Physics", *Prog. Theor. Phys. Suppl.* **183** (2010) 1, doi:10.1143/PTPS.183.1, arXiv:1003.3552.

[15] G. Perez and L. Randall, "Natural neutrino masses and mixings from warped geometry", *JHEP* **01** (2009) 077, doi:10.1088/1126-6708/2009/01/077, arXiv:0805.4652.

[16] S. Casagrande et al., "Flavor physics in the Randall-Sundrum model I. Theoretical setup and electroweak precision tests", *JHEP* **10** (2008) 094, doi:10.1088/1126-6708/2008/10/094, arXiv:0807.4937.

[17] A. J. Buras, B. Duling, and S. Gori, "The impact of Kaluza-Klein fermions on Standard Model fermion couplings in a RS model with custodial protection", *JHEP* **09** (2009) 076, doi:10.1088/1126-6708/2009/09/076, arXiv:0905.2318.

[18] M. Blanke et al., "$\Delta F = 2$ observables and fine-tuning in a warped extra dimension with custodial protection ", *JHEP* **03** (2009) 001, doi:10.1088/1126-6708/2009/03/001, arXiv:0809.1073.

[19] G. F. Giudice and O. Lebedev, "Higgs-dependent Yukawa couplings", *Phys. Lett. B* **665** (2008) 79, doi:10.1016/j.physletb.2008.05.062, arXiv:0804.1753.

[20] J. A. Aguilar-Saavedra, "A minimal set of top-Higgs anomalous couplings", *Nucl. Phys. B* **821** (2009) 215, doi:10.1016/j.nuclphysb.2009.06.022, arXiv:0904.2387.

[21] M. E. Albrecht et al., "Electroweak and flavour structure of a warped extra dimension with custodial protection", *JHEP* **09** (2009) 064, doi:10.1088/1126-6708/2009/09/064, arXiv:0903.2415.

[22] A. Goudelis, O. Lebedev, and J. H. Park, "Higgs-induced lepton flavor violation", *Phys. Lett. B* **707** (2012) 369, doi:10.1016/j.physletb.2011.12.059, arXiv:1111.1715.

[23] D. McKeen, M. Pospelov, and A. Ritz, "Modified Higgs branching ratios versus CP and lepton flavor violation", *Phys. Rev. D* **86** (2012) 113004, doi:10.1103/PhysRevD.86.113004, arXiv:1208.4597.

[24] A. Pilaftsis, "Lepton flavour nonconservation in H^0 decays", *Phys. Lett. B* **285** (1992) 68, doi:10.1016/0370-2693(92)91301-O.

[25] J. G. Körner, A. Pilaftsis, and K. Schilcher, "Leptonic CP asymmetries in flavor-changing H^0 decays", *Phys. Rev. D* **47** (1993) 1080, doi:10.1103/PhysRevD.47.1080.

[26] E. Arganda, M. J. Herrero, X. Marcano, and C. Weiland, "Imprints of massive inverse seesaw model neutrinos in lepton flavor violating Higgs boson decays", *Phys. Rev. D* **91** (2015), no. 1, 015001, doi:10.1103/PhysRevD.91.015001, arXiv:1405.4300.

[27] CMS Collaboration, "Search for lepton-flavour-violating decays of the Higgs boson", *Phys. Lett. B* **749** (2015) 337, doi:10.1016/j.physletb.2015.07.053, arXiv:1502.07400.

[28] CMS Collaboration, "Search for lepton flavour violating decays of the Higgs boson to $e\tau$ and $e\mu$ in proton-proton collisions at $\sqrt{s}$=8 TeV", *Physics Letters B* **763** (2016) 472, doi:http://dx.doi.org/10.1016/j.physletb.2016.09.062.

[29] ATLAS Collaboration, "Search for lepton-flavour-violating decays of the Higgs and Z bosons with the ATLAS detector", *Eur. Phys. J.* **C77** (2017), no. 2, 70, doi:10.1140/epjc/s10052-017-4624-0, arXiv:1604.07730.

[30] ATLAS Collaboration, "Search for lepton-flavour-violating $H \to \mu\tau$ decays of the Higgs boson with the ATLAS detector", *JHEP* **11** (2015) 211, doi:10.1007/JHEP11(2015)211, arXiv:1508.03372.

[31] B. McWilliams and L.-F. Li, "Virtual effects of Higgs particles", *Nucl. Phys. B* **179** (1981) 62, doi:10.1016/0550-3213(81)90249-2.

[32] O. U. Shanker, "Flavour violation, scalar particles and leptoquarks", *Nucl. Phys. B* **206** (1982) 253, doi:10.1016/0550-3213(82)90534-X.

[33] G. Blankenburg, J. Ellis, and G. Isidori, "Flavour-changing decays of a 125 GeV Higgs-like particle", *Phys. Lett. B* **712** (2012) 386, doi:10.1016/j.physletb.2012.05.007, arXiv:1202.5704.

[34] R. Harnik, J. Kopp, and J. Zupan, "Flavor violating Higgs decays", *JHEP* **03** (2013) 26, doi:10.1007/JHEP03(2013)026, arXiv:1209.1397.

[35] Particle Data Group, J. Beringer et al., "Review of Particle Physics", *Phys. Rev. D* **86** (2012) 010001, doi:10.1103/PhysRevD.86.010001.

[36] A. Celis, V. Cirigliano, and E. Passemar, "Lepton flavor violation in the Higgs sector and the role of hadronic tau-lepton decays", *Phys. Rev. D* **89** (2014) 013008, doi:10.1103/PhysRevD.89.013008, arXiv:1309.3564.

[37] S. M. Barr and A. Zee, "Electric Dipole Moment of the Electron and of the Neutron", *Phys. Rev. Lett.* **65** (1990) 21, doi:10.1103/PhysRevLett.65.21.

[38] CMS Collaboration, "Evidence for the direct decay of the 125 GeV Higgs boson to fermions", *Nature Phys.* **10** (2014) 557, doi:10.1038/nphys3005, arXiv:1401.6527.

[39] CMS Collaboration, "Evidence for the 125 GeV Higgs boson decaying to a pair of τ leptons", *JHEP* **05** (2014) 104, doi:10.1007/JHEP05(2014)104, arXiv:1401.5041.

[40] ATLAS Collaboration, "Evidence for the Higgs-boson Yukawa coupling to tau leptons with the ATLAS detector", *JHEP* **04** (2015) 117, doi:10.1007/JHEP04(2015)117, arXiv:1501.04943.

[41] CMS Collaboration, "The CMS experiment at the CERN LHC", *JINST* **3** (2008) S08004, doi:10.1088/1748-0221/3/08/S08004.

[42] H. M. Georgi, S. L. Glashow, M. E. Machacek, and D. V. Nanopoulos, "Higgs Bosons from Two Gluon Annihilation in Proton Proton Collisions", *Phys. Rev. Lett.* **40** (1978) 692, doi:10.1103/PhysRevLett.40.692.

[43] R. N. Cahn, S. D. Ellis, R. Kleiss, and W. J. Stirling, "Transverse-momentum signatures for heavy Higgs bosons", *Phys. Rev. D* **35** (1987) 1626, doi:10.1103/PhysRevD.35.1626.

[44] S. L. Glashow, D. V. Nanopoulos, and A. Yildiz, "Associated production of Higgs bosons and Z particles", *Phys. Rev. D* **18** (1978) 1724, doi:10.1103/PhysRevD.18.1724.

[45] 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.

[46] 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.

[47] 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.

[48] S. Alioli et al., "Jet pair production in POWHEG", *JHEP* **04** (2011) 081, doi:10.1007/JHEP04(2011)081, arXiv:1012.3380.

[49] S. Alioli, P. Nason, C. Oleari, and E. Re, "NLO Higgs boson production via gluon fusion matched with shower in POWHEG", *JHEP* **04** (2009) 002, doi:10.1088/1126-6708/2009/04/002, arXiv:0812.0578.

[50] J. Alwall et al., "MadGraph 5: going beyond", *JHEP* **06** (2011) 128, doi:10.1007/JHEP06(2011)128, arXiv:1106.0522.

[51] 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", *Journal of High Energy Physics* **2014** (2014), no. 7, 79, doi:10.1007/JHEP07(2014)079.

[52] CMS Collaboration, "Event generator tunes obtained from underlying event and multiparton scattering measurements", *Eur. Phys. J.* **C76** (2016), no. 3, 155, doi:10.1140/epjc/s10052-016-3988-x, arXiv:1512.00815.

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

[54] CMS Collaboration, "Description and performance of track and primary-vertex reconstruction with the CMS tracker", *JINST* **9** (2014) P10009, doi:10.1088/1748-0221/9/10/P10009, arXiv:1405.6569.

[55] CMS Collaboration, "Particle–Flow Event Reconstruction in CMS and Performance for Jets, Taus, and $E_{\mathrm{T}}^{\mathrm{miss}}$", CMS Physics Analysis Summary CMS-PAS-PFT-09-001, 2009.

[56] CMS Collaboration, "Particle flow reconstruction of 0.9 TeV and 2.36 TeV collision events in CMS", CMS Physics Analysis Note CMS-PAS-PFT-10-001, 2010.

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

[58] CMS Collaboration, "Performance of the CMS missing transverse momentum reconstruction in pp data at $\sqrt{s}$ = 8 TeV", *JINST* **10** (2015) P02006, doi:10.1088/1748-0221/10/02/P02006, arXiv:1411.0511.

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

[60] 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.

[61] CMS Collaboration, "Reconstruction and identification of tau lepton decays to hadrons and ν_τ at CMS", *JINST* **11** (2016) P01019, doi:10.1088/1748-0221/11/01/P01019, arXiv:1510.07488.

[62] CMS Collaboration, "Performance of reconstruction and identification of tau leptons in their decays to hadrons and tau neutrino in LHC Run-2",.

[63] M. Cacciari, G. P. Salam, and G. Soyez, "FastJet user manual", *Eur. Phys. J. C* **72** (2012) 1896, doi:10.1140/epjc/s10052-012-1896-2, arXiv:1111.6097.

[64] 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.

[65] CMS Collaboration, "Determination of jet energy calibration and transverse momentum resolution in CMS", *JINST* **6** (2011) 11002, doi:10.1088/1748-0221/6/11/P11002, arXiv:1107.4277.

[66] R. K. Ellis, I. Hinchliffe, M. Soldate, and J. J. van der Bij, "Higgs Decay to $\tau^+\tau^-$: A possible signature of intermediate mass Higgs bosons at high energy hadron colliders", *Nucl. Phys. B* **297** (1988) 221, doi:10.1016/0550-3213(88)90019-3.

[67] T. Junk, "Confidence level computation for combining searches with small statistics", *Nucl. Instrum. Meth. A* **434** (1999) 435, doi:10.1016/S0168-9002(99)00498-2, arXiv:hep-ex/9902006.

[68] 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.

[69] LHC Higgs Cross Section Working Group Collaboration, "Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector", arXiv:1610.07922.

[70] CMS Collaboration, "Search for Lepton Flavour Violating Decays of the Higgs Boson in the mu-tau final state at 13 TeV", CMS Physics Analysis Summary CMS-PAS-HIG-16-005, 2016.

[71] A. Denner et al., "Standard model Higgs-boson branching ratios with uncertainties", *Eur. Phys. J. C* **71** (2011) 1753, doi:10.1140/epjc/s10052-011-1753-8, arXiv:1107.5909.