CMS Draft Analysis Note

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

2017/09/05
Head Id: 423044
Archive Id: 423865P
Archive Date: 2017/08/30
Archive Tag: trunk

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 flavour violating decays of the 125 GeV Higgs boson in the $\mu\tau$ and eτ decay modes is presented. The search is based on a data set corresponding to an integrated luminosity of 35.9 fb^{-1} of proton-proton collisions collected with the CMS detector in 2016, at a centre-of-mass energy of 13 TeV. No significant excess over the standard model expectation is observed. The observed (expected) upper limits on the lepton flavour violating branching fractions 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, which are found to be $\sqrt{|Y_{\mu\tau}|^2 + |Y_{\tau\mu}|^2} < 1.43 \times 10^{-3}$ and $\sqrt{|Y_{\mathrm{e}\tau}|^2 + |Y_{\tau\mathrm{e}}|^2} < 2.26 \times 10^{-3}$ at 95% confidence level. These limits on the lepton flavour violating branching fractions of the Higgs boson and on the associated Yukawa couplings are the best achieved to date.

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

PDFAuthor:	Cecile Caillol, Maria Cepeda, Sridhara Dasu, Nabarun Dev, Colin Jessop, Fanbo Meng, Silvia Taroni, Daniel Troendle
PDFTitle:	Search for lepton flavour violating decays of the Higgs boson to mu-tau and e-tau in proton-proton collisions at a centre-of-mass energy of 13 TeV
PDFSubject:	CMS
PDFKeywords:	CMS, physics, Higgs, taus

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

1 Introduction

The discovery of the Higgs boson at the CERN LHC [1–3] has stimulated further precision measurements of the properties of the new particle. A combined study of the 7 and 8 TeV data sets collected by the CMS and ATLAS Collaborations shows consistency between the measured couplings of the Higgs boson (H) and the standard model (SM) predictions [4]. However, the constraint on the branching fraction to non-SM decay modes derived from these measurements, $\mathcal{B}(\text{non-SM}) < 34\%$ at 95% confidence level (CL), still allows for the possibility of a significant contribution from exotic decays [4].

In this paper a search for lepton flavour violating (LFV) decays of the 125 GeV Higgs boson in the $\mu\tau$ and $e\tau$ channels is presented. These decays are forbidden in the SM but occur in many new physics scenarios. These include supersymmetric [5–13], composite Higgs [14, 15], or Randall–Sundrum models [16–18], SM extentions with more than one Higgs boson doublet [19, 20] or with flavour symmetries [21], and many other scenarios [22–36]. The presence of LFV Higgs boson couplings would allow $\tau \to \mu$ and $\tau \to e$ to proceed via a virtual Higgs boson [37, 38]. Consequently the experimental limits on rare τlepton decays, such as $\tau \to e\gamma$ and $\tau \to \mu\gamma$ [39], provide upper limits on $\mathcal{B}(H \to \mu\tau)$ and $\mathcal{B}(H \to e\tau)$ [40, 41] of $\mathcal{O}(10\%)$. Measurements of the electron and muon magnetic moment, and exclusion limits on the electric dipole moment of the electron also provide complementary constraints [42]. The LFV Higgs decay to μe is strongly constrained by the $\mu \to e\gamma$ limit, $\mathcal{B}(H \to e\mu) < \mathcal{O}(10^{-9})$ [43].

The CMS experiment published the first direct search for $H \to \mu\tau$ [44], followed by searches for $H \to e\tau$ and $H \to e\mu$ decays [45], using proton-proton (pp) collision data corresponding to an integrated luminosity of 19.7 fb^{-1} at a centre-of-mass energy of 8 TeV. A small excess of data with respect to the SM background-only hypothesis at $m_H = 125$ GeV was observed in the $H \to \mu\tau$ channel, with a significance of 2.4 standard deviations (σ), and the best fit for the branching fraction was found to be $\mathcal{B}(H \to \mu\tau) = (0.84^{+0.39}_{-0.37})\%$. A constraint was set on the observed (expected) branching fraction $\mathcal{B}(H \to \mu\tau) < 1.51\%$ (0.75%) at 95% CL. No excess of events over the estimated background was observed in the $H \to e\tau$ or $H \to e\mu$ channels, and observed (expected) upper limits on the branching fractions $\mathcal{B}(H \to e\tau) < 0.69\%$ (0.75%) and $\mathcal{B}(H \to e\mu) < 0.035\%$ (0.048%) at 95% CL were set. The ATLAS Collaboration reported searches for $H \to e\tau$ and $H \to \mu\tau$ using proton-proton (pp) collision data at a centre-of-mass energy of 8 TeV, finding no significant excess of events over the background expectation, and set observed (expected) limits of $\mathcal{B}(H \to \mu\tau) < 1.43\%$ (1.01%) and $\mathcal{B}(H \to e\tau) < 1.04\%$ (1.21%) at 95% CL [46, 47].

The search described in this paper is performed in four decay channels, $H \to \mu\tau_e$, $H \to \mu\tau_h$, $H \to e\tau_\mu$, $H \to e\tau_h$, where τ_e, τ_μ, and τ_h correspond to the electronic, muonic, and hadronic decay channels of τ leptons, respectively. The decay channels $H \to e\tau_e$ and $H \to \mu\tau_\mu$, are not considered because of the large background contribution from Z boson decays. The expected final state signatures are very similar to those for the SM $H \to \tau\tau$ decays, studied by CMS [48–50] and ATLAS [51], but with some significant kinematic differences. The electron (muon) in the LFV $H \to e(\mu)\tau$ decay is produced promptly, and tends to have a larger momentum than in the SM $H \to \tau_{e(\mu)}\tau_h$ decay. With respect to [44] and [45], the search described in this paper is performed using the output of a boosted decision tree (BDT) to distinguish signal from background events, and a separate analysis, similar in strategy to the 8 TeV CMS publications is performed as cross check. The results of both strategies are reported in this paper.

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 analyses (Section 3), the event reconstruction is described in Section 4. The backgrounds, which are common to all channels but with

different rates in each, are described in Section 5. The event selection and background validation are described separately for the two Higgs boson decay modes $H \to e\tau$ and $H \to \mu\tau$ in Section 6. The systematic uncertainties are described in Section 7 and the results are then presented in Section 8.

2 The CMS detector

The central feature of the CMS apparatus is a superconducting solenoid of 6 m internal diameter, providing a magnetic field of 3.8 T. Within the solenoid volume are a silicon pixel and strip tracker, a lead tungstate crystal electromagnetic calorimeter (ECAL), and a brass and scintillator hadron calorimeter (HCAL), each composed of a barrel and two endcap sections. Forward calorimeters extend the pseudorapidity coverage provided by the barrel and endcap detectors. Muons are detected in gas-ionization chambers embedded in the steel flux-return yoke outside the solenoid. 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 at a rate of around 100 kHz within a time interval of less than 4 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. 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. [52].

3 Collision data and simulated events

The analyses use samples of pp collisions collected in 2016 by the CMS experiment at the LHC at a centre-of-mass energy of $\sqrt{s} = 13$ TeV, corresponding to an integrated luminosity of 35.9 fb^{-1}. Isolated single muon triggers are used to collect the data samples in the $H \to \mu\tau$ search. Triggers requiring a single isolated electron, or a combination of an electron and a muon, are used in the $H \to e\tau_h$ and $H \to e\tau_\mu$ channels, respectively. Simulated samples of signal and background events are produced with several event generators. The Higgs bosons are produced in pp collisions predominantly by gluon fusion (ggH) [53], but also by vector boson fusion (VBF) [54], and in association with a W or Z boson [55]. The ggH and VBF Higgs samples are generated with POWHEG 2.0 [56–60] while the MINLO HVJ [61] extension of POWHEG 2.0 is used for the WH and ZH simulated samples. The MG5_aMC@NLO [62] generator is used for $Z + \text{jets}$ and $W + \text{jets}$ processes. They are simulated at leading order (LO) with the MLM jet matching and merging [63]. Diboson production is simulated at next-to-LO (NLO) using MG5_aMC@NLO generator with the FxFx jet matching and merging [64], whereas POWHEG 2.0 and 1.0 are used for $t\bar{t}$ and single top quark production, respectively. The POWHEG and MADGRAPH generators are interfaced with PYTHIA 8.212 [65] for parton showering, fragmentation, and decays. The PYTHIA parameters for the underlying event description are set to the CUETP8M1 tune [66]. Due to the high luminosities attained during data taking, many events have multiple pp interactions per bunch crossing (pileup). The effect is taken into account in simulated samples, by generating concurrent minimum bias events. All simulated samples are weighted to match the pileup distribution observed in data, that has an average of approximately 27 interactions per bunch crossing. The CMS detector response is modelled using GEANT4 [67].

4 Event reconstruction

The global event reconstruction, called particle-flow (PF) algorithm, consists of reconstructing and identifying each individual particle with an optimized combination of all subdetector information [68]. In this process, the identification of the particle type (photon, electron, muon, charged hadron, neutral hadron) plays an important role in the determination of the particle direction and energy. The charged particles are required to be consistent with the primary pp vertex, identified as the reconstructed vertex with the largest value of summed physics-object p_{T}^2, where p_{T} is the transverse momentum. The physics objects are returned by a jet finding algorithm [69, 70] applied to all charged tracks associated with the vertex, plus the corresponding associated missing transverse momentum. The missing transverse momentum, $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$, is defined as the negative of the vector $\vec{p}_{\mathrm{T}}$ sum of all identified PF objects in the event [71]. Its magnitude is referred to as $p_{\mathrm{T}}^{\mathrm{miss}}$.

A muon is identified as a track in the tracker consistent with either a track or several hits in the muon system, associated with an energy deficit in the calorimeters [72]. Identification is based on the number of spacial points measured in the tracker and in the muon system, the track quality and its consistency with the event vertex location. The energy is obtained from the corresponding track momentum.

An electron is identified as a primary charged particle track in combination with, potentially, one or more ECAL energy clusters corresponding to this track extrapolation to the ECAL and to possible bremsstrahlung photons emitted when interacting with the material of the tracker [73]. 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. They are identified using a multivariate (MVA) 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 as well as various shower shape observables in the calorimeters. Electrons from photon conversions are removed. The energy of electrons is determined from a combination of the track momentum at the primary vertex, the corresponding ECAL cluster energy, and the energy sum of all bremsstrahlung photons attached to the track.

Hadronically decaying τ leptons are reconstructed and identified using the hadrons-plus-strips (HPS) algorithm [74, 75]. The reconstruction starts from a jet and searches for the products of the main τ lepton decay modes: one charged hadron and up to two neutral pions, or three charged hadrons. To improve the reconstruction efficiency in case of conversion of the photons from neutral-pion decay, this algorithm considers the PF photons and electrons from a strip along the azimuthal direction ϕ. The charges of all the PF objects from tau lepton decay, except for the electrons from neutral pions, are summed to reconstruct the tau lepton charge. A MVA discriminator, based on the information of the reconstructed tau lepton and of the charged particles in a cone around it, is used to reduce the rate for quark- and gluon-initiated jets to be identified as τ candidates. The working point used in the analysis has an efficiency of about 60% for a genuine τ_h, with about 0.5% misidentification rate for quark and gluon jets [75]. Additionally, muons and electrons misidentified as tau leptons are rejected using a dedicated set of selection criteria based on the consistency between the measurements in tracker, calorimeters, and muon detectors. The specific identification criteria depend on the final state studied and on the background composition. The tau leptons that decay to muons and electrons are reconstructed as prompt muons and electrons as described above.

Charged hadrons are identified as charged particle tracks neither reconstructed as electrons nor as muons. Neutral hadrons are identified as HCAL energy clusters not assigned to any charged

hadron, or as ECAL and HCAL energy excesses with respect to the expected charged-hadron energy deposit. Both charged and neutral hadrons are clustered into hadronic jets using the infrared and collinear safe anti-k_T algorithm [69], implemented in the FASTJET package [76], with a distance parameter of 0.4. The jet momentum is determined as the vectorial sum of all particle momenta in this jet, and is found in the simulation to be on average within 10% of the true momentum over the whole p_T spectrum and detector acceptance. An offset correction is applied to jet energies to take into account the contribution from pileup [77]. Jet energy corrections are derived from the simulation, and are confirmed with in situ measurements of the energy balance of dijet, multijet, photon + jet, and Z+jet events [78]. The variable $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$is used to measure the separation between reconstructed objects in the detector. Any jet within $\Delta R = 0.4$ of the identified leptons is removed.

Jets misidentified as electrons, muons, or tau leptons are suppressed by imposing isolation requirements. The muon (electron) isolation is measured relative to its transverse momentum p_T^ℓ ($\ell = e, \mu$), by summing over the transverse momenta of PF particles in a cone with $\Delta R = 0.4$ (0.3) around the lepton:

$$I_{rel}^\ell = \left(\sum p_T^{charged} + \max\left[0, \sum p_T^{neutral} + \sum p_T^\gamma - p_T^{PU}(\ell)\right]\right)/p_T^\ell.$$

where $p_T^{charged}$, $p_T^{neutral}$, and p_T^γ indicate the p_T of a charged particle, a neutral particle, and a photon within the cone, respectively. The neutral contribution to isolation from pileup $p_T^{PU}(\ell)$ is estimated from the area of the jet and the average energy density of the event [79, 80] for the electron or through a correction, the modified $\Delta\beta$ correction described in Ref. [72], based on the reconstructed tracks and vertices for the muons. The charged contribution to isolation from pileup is rejected requiring the tracks to originate from the primary vertex.

The transverse mass $M_T(\ell)$ is a variable formed from the lepton momentum and the missing transverse momentum vectors: $M_T(\ell) = \sqrt{2|\vec{p}_T^\ell||\vec{p}_T^{miss}|(1 - \cos\Delta\phi_{\ell - p_T^{miss}})}$. It is used to discriminate the Higgs boson signal candidates from the W + jets background. 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 $m_H \gg m_\tau$, the τ lepton decay products are highly Lorentz boosted in the direction of the τ candidate [81]. The neutrino momenta can be approximated to have the same direction as the other visible decay products of the τ ($\vec{\tau}^{vis}$) and the component of the $\vec{p}_T^{miss}$ in the direction of the visible τ lepton decay products is used to estimate the transverse component of the neutrino momentum ($p_T^{\nu,\,est}$). The collinear mass can then be derived from the visible mass of the τ-μ or τ-e system (M_{vis}) as $M_{col} = M_{vis}/\sqrt{x_\tau^{vis}}$, where x_τ^{vis} is the fraction of energy carried by the visible decay products of the τ ($x_\tau^{vis} = p_T^{\vec{\tau}^{vis}}/(p_T^{\vec{\tau}^{vis}} + p_T^{\nu,\,est})$).

5 Background estimation

The signal contains a prompt isolated lepton, μ or e, 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 momentum. The main background processes are $Z \to \tau\tau$, which is estimated from simulation, and W+jets and QCD multijet productions where the jets are misidentified as leptons. Other backgrounds include $t\bar{t}$ pairs, SM Higgs boson production ($H \to \tau\tau, WW$), WW, WZ, ZZ, $W\gamma^{(*)}$ + jets processes, single top quark production, and $Z \to \ell\ell$ ($\ell = e, \mu, \tau$). All the background are estimated from simulation with the exception of the misidentified-lepton backgrounds that are estimated from data with either fully data-

driven or semi-data driven methods. The background estimation techniques are described in detail below. The background estimate is validated with with control regions that have enhanced contributions from the dominant backgrounds.

The $Z \to \ell\ell$ background is estimated from simulation. A reweighting is applied to correct the generator-level Z p_T and $m_{\ell\ell}$ distributions in LO MG5_aMC@NLO samples to reduce the shape discrepancy between collision data and simulation. The reweighting factors are extracted from a $Z \to \mu\mu$ control region and are applied to both $Z \to \mu\mu$ and $Z \to$ ee simulated samples in bins of Z p_T and $m_{\ell\ell}$. Additional corrections for $\mu \to \tau_h$ and e $\to \tau_h$ misidentification rates are applied when the reconstructed τ_h candidate is matched to a muon or an electron, respectively, at the generator level. These corrections are measured in $Z \to \ell\ell$ events and depend on the pseudorapidity of the lepton. The $t\bar{t}$ +jets background is particularly important in the $e\mu$ final state. A correction based on the generated p_T of the top quark and antiquark is applied to the simulation to match the p_T distribution observed in data. The SM Higgs boson production contributes a small but nonnegligible background. It arises predominantly from $H \to \tau\tau$ but also from $H \to WW$ decays and peaks at lower values of M_{col} than the signal, because of additional neutrinos in the decays. The event selection described below uses an MVA 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 MVA discriminant output.

Jets misidentified as leptons are a source of background arising from two sources, W + jets and QCD multijet events. In W + jets background events one lepton candidate is a real lepton from the W boson decay and the other is a jet misidentified as a lepton. In QCD multijet events, both lepton candidates are misidentified jets. In each of the four channels for this analysis ($\mu\tau_h$, eτ_h, $\mu\tau_e$, eτ_μ), the misidentified-lepton background has been estimated using data driven methods. In the $\mu\tau_e$ and eτ_μ channels it is also estimated using a technique, called semi-data driven in following lines, partially based on control samples in data and partially on simulation used in the SM $H \to \tau\tau$ analysis [50], in which the misidentified W + jets background is estimated from simulation and the QCD background with data. The two techniques give consistent results; the semi-data driven technique is chosen for the leptonically decaying tau channels as the fully data driven technique is statistically limited.

Fully data driven technique

The probabilities with which jets are misidentified as e (f_e), μ (f_μ), or τ_h (f_τ), are estimated in Z + jets events in collision data. This event sample is obtained by requiring a Z boson candidate plus one jet that can be misidentified as a lepton. The Z boson candidate is formed from two muons with $p_T > 26$ GeV, $|\eta| < 2.4$, and $I^\ell_{rel} < 0.15$. In the case of the jet $\to$ e misidentification rate, the muon relative isolation is loosened to $I^\ell_{rel} < 0.25$ to increase the number of selected Z bosons preserving the purity. The muons are required to have an opposite sign charges and the invariant mass ($M_{\mu\mu}$) to satisfy the requirement $70 < M_{\mu\mu} < 110$ GeV. The contribution from diboson events, where the third lepton candidate corresponds to a genuine lepton, is subtracted using simulation.
The requirements for the third lepton vary depending on the lepton flavour but are the same in both $Z + \ell$ and LFV Higgs search datasets. These requirements define two regions: the signal-like region, where the third lepton satisfies the criteria to select $H \to e\tau$ or $H \to \mu\tau$ events in LFV Higgs search datasets, and the background-enriched region.

- **Jets $\to \mu$ estimation**: the background-enriched region is defined requiring muon candidates, defined as not-isolated, having $0.15 < I^\mu_{rel} < 0.25$, $p^\mu_T > 10$ GeV, and $|\eta^\mu| < 2.4$. In the signal-like region, muons are required to have $I^\mu_{rel} < 0.15$. The

misidentification rate is computed as a function of lepton p_T.

- **Jets $\rightarrow$ e estimation**: the background-enriched region for electron candidates is defined requiring electron having $0.1 < I^e_{rel} < 0.5$, $p^e_T > 10$ GeV, $|\eta^e| < 2.4$. In the signal-like region the electrons are required to have $I^e_{rel} < 0.1$. The misidentification rate is calculated as a function of lepton p_T as it shows a dependence on this variable and is flat in $|\eta|$.
- **Jets $\rightarrow \tau_h$ estimation**: hadronically decaying τ candidates in the background-enriched region, having $p^\tau_T > 30$ GeV and $|\eta^\tau| < 2.3$, are identified using the HPS working point having an efficiency of about 75% for genuine τ_h and are not identified with the working point used for the signal-like region that has an efficiency of about 60% for genuine τ_h [75]. The misidentification rates show a p_T dependence that varies with the τ decay mode and $|\eta|$. The misidentification rates are thus obtained as a function of p_T for the different decay modes and $|\eta|$ regions ($|\eta| < 1.5$ or $|\eta| > 1.5$).

From $Z + \ell$ events, the misidentification rate (f_e, f_μ or f_τ) is computed as

$$f_i = \frac{N_i(\text{signal-like region})}{N_i(\text{background-enriched region}) + N_i(\text{signal-like region})},$$

where N_i(signal-like region) are the events with a third lepton that pass the signal-like region selection, N_i(backgroun enriched region) is the number of events in the background-enriched region and $i = e$, μ or τ.

These misidentification rates are applied to the events in the background-enriched region of the LFV Higgs search datasets, to estimate the number of misidentified background events N_i(misidentified) in the signal region:

$$N_i(\text{misidentified}) = \frac{f_i}{1 - f_i} N_i(\text{background-enriched region}).$$

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

$$N_{\tau\ell}(\text{misidentified}) = \frac{f_\tau \cdot f_\ell}{(1 - f_\tau) \cdot (1 - f_\ell)} N_{\tau\ell}(\text{background-enriched region}),$$

where $\ell = e$ or μ.

The background estimation is validated in a like-sign region applying the misidentification rate f_i to events selected inverting the change requirement of the lepton pair in both the background-enriched and the signal-like regions in the LFV Higgs search dataset. The background estimation can also be validated in a W boson enriched control region. This data sample is obtained by applying the signal-like region requirements and $M_T(\ell) > 60$ GeV ($\ell = e$ or μ) and $M_T(\tau_h) > 80$ GeV. The background validation in these control regions shows a good agreement of estimated and observed background: the results of these validations are given in the Section 6.

Semi data-driven technique

The W + jets background contribution to the misidentified-lepton background is estimated with simulated samples generated with different jet multiplicities and combined to improve

the statistical precision. The QCD multijet contribution is estimated with like-sign collision data events that pass the signa-like region requirement. The expected yield from non-QCD processes is subtracted using simulation. The resulting sample is then rescaled to account for the 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 inverted isolation requirements as illustrated in Ref. [50].

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 boson production mechanisms. The jets are required to have $p_{\mathrm{T}} > 30$ GeV and $|\eta| < 4.7$. The 0-jet category enhances the ggH contribution, while the 1-jet category enhances ggH production with initial-state radiation. The 2-jet ggH category has a further requirement that the invariant mass of the two jets $\mathrm{M}_{jj} < 550$ GeV while the 2-jet VBF category with the requirement $\mathrm{M}_{jj} > 550$ 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 the same in all the channels except that in the $\mathrm{H} \to \mathrm{e}\tau$ channels the M_{jj} threshold is 500 GeV. This threshold is set accordingly to the result of the optimization based on the best expected limits for this channel.

After the loose selection, a binned likelihood is used to fit the distribution of a 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 a fit is performed to the the M_{col} distribution. This is referred to as the M_{col} fit analysis. Requirements on additional kinematic variables such as $M_{\mathrm{T}}(\ell)$ are chosen to obtain the most stringent expected limits. The lepton p_{T} has been excluded from this optimization to avoid the backgrounds being led to peak under the signal in the M_{col} distribution.

6.1 $\mathrm{H} \to \mu\tau_h$

The loose selection begins requiring an isolated μ and an isolated τ_h of opposite charges and separated by $\Delta R > 0.5$. Events with additional electrons, muons, or tau lepton candidates, or at least one jet identified by the combined secondary vertex (CSVv2) b-tagging algorithm [82] as arising from a b quark, are vetoed. The candidates are required to have $p_{\mathrm{T}}^{\mu} > 26$ GeV, $|\eta^{\mu}| < 2.4$, $p_{\mathrm{T}}^{\tau_h} > 30$ GeV, and $|\eta^{\tau_h}| < 2.3$. The tighter selection used for the M_{col} fit analysis further requires $M_{\mathrm{T}}(\tau_h) < 105$ GeV in the 0, 1 and 2-jet ggH categories, and $M_{\mathrm{T}}(\tau_h) < 85$ GeV in the 2-jet VBF category. The selections are summarised in Table 1.

A BDT is trained after the loose selection and combining all categories. The signal training sample used is a mixture of simulated ggH and VBF events, weighted according to their respective SM production cross sections. The background training sample is a collision data set of misidentified-lepton events, as this is the dominant background in this channel. The leptons are required to be like-sign and not-isolated in order to select an orthogonal data set to the signal region, and yet have the same kinematic properties. The input variables to the BDT are: p_{T}^{μ}, $p_{\mathrm{T}}^{\tau_h}$, M_{col}, $p_{\mathrm{T}}^{\mathrm{miss}}$, $M_{\mathrm{T}}(\tau_h)$, $\Delta\eta(\mu, \tau_h)$, $\Delta\phi(\mu, \tau_h)$, and $\Delta\phi(\tau_h, \vec{p}_{\mathrm{T}}^{\,\mathrm{miss}})$. They are shown in Fig. 1

The control region tests of the misidentified-lepton backgrounds described in Section 5 are performed after the loose selection. Figure 3 (left) shows the data compared to the background estimation in the like-sign control region. The background estimate is cross-checked using a

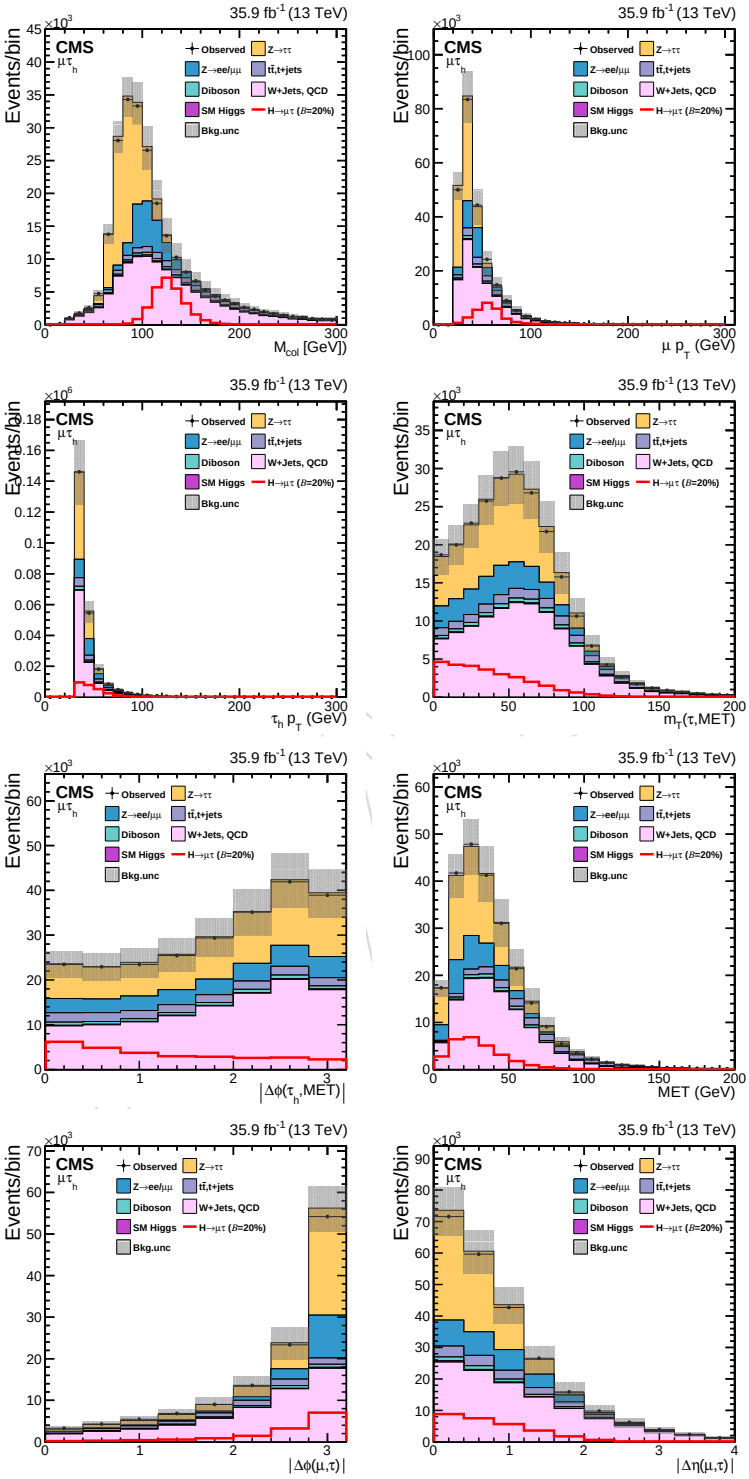

Figure 1: Distributions of the input variables to the BDT for the H $\rightarrow \mu\tau_h$ channel.

W + jets enriched control region and the agreement is comparable (Fig. 3 central). To increase the number of events, the distributions are shown after the loose selection. 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 charges and separated by $\Delta R > 0.3$ are requested. Events with additional electrons, muons, or tau lepton candidates are rejected. Events with at least one b-tagged 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 $I_{rel}^\mu < 0.15$. The electron is required to have $p_T^e > 10$ GeV, $|\eta^e| < 2.3$, and $I_{rel}^e < 0.1$.

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 ggH 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 tau lepton decay are approximately collinear to the electron direction so an additional requirement is made on the azimuthal angle between the electron and the $\vec{p}_T^{miss}$: $\Delta\phi(e, \vec{p}_T^{miss})$ <0.7, 0.7, 0.5, 0.3 for the 0-jet, 1-jet, 2-jets ggH, and 2-jets VBF categories, respectively. In the 0-jet, 1-jet category it is further required that $\Delta\phi(e, \mu) > 2.5, 1.0$, respectively. The selections are summarised in Table 1.

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 the 1-jet category. 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 to the BDT are: p_T^μ, p_T^e, M_{col}, $M_T(\mu)$, $M_T(e)$, $\Delta\phi(e, \mu)$, $\Delta\phi(e, \vec{p}_T^{miss})$, and $\Delta\phi(\mu, \vec{p}_T^{miss})$. The distributions of these variables are shown in Fig. 2. 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 3 (right) shows the data compared to the background estimation in the $t\bar{t}$ control region. This control region has the loose 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 $H \to e\tau_\mu$ channeland show similar agreement.

6.3 $H \to e\tau_h$

The electron candidate is required to have $p_T^e > 26$ GeV, $|\eta^e| < 2.1$, and $I_{rel}^e < 0.1$. The τ_h candidate is required to have $p_T^{\tau_h} > 30$ GeV and $|\eta^{\tau_h}| < 2.3$. In each event, the electron and the tau candidate are required to have opposite sign charges, and to be separated by $\Delta R > 0.5$. Events with additional electrons, muons, or tau lepton candidates are rejected. No requirement is made on the number of b-tagged jets as the $t\bar{t}$ contribution is small. The additional selection used for the M_{col} fit analysis further requires that $M_T(\tau_h) < 60$ GeV. The selections are summarised in Table 2. A BDT is trained after the loose selection. The same training samples as for the $H \to \mu\tau_h$ channel are used, except with an electron at the place of the muon . The input variables to the BDT are also the same except for the addition of the visible mass, M_{vis}, and the removal of p_T^{miss}.

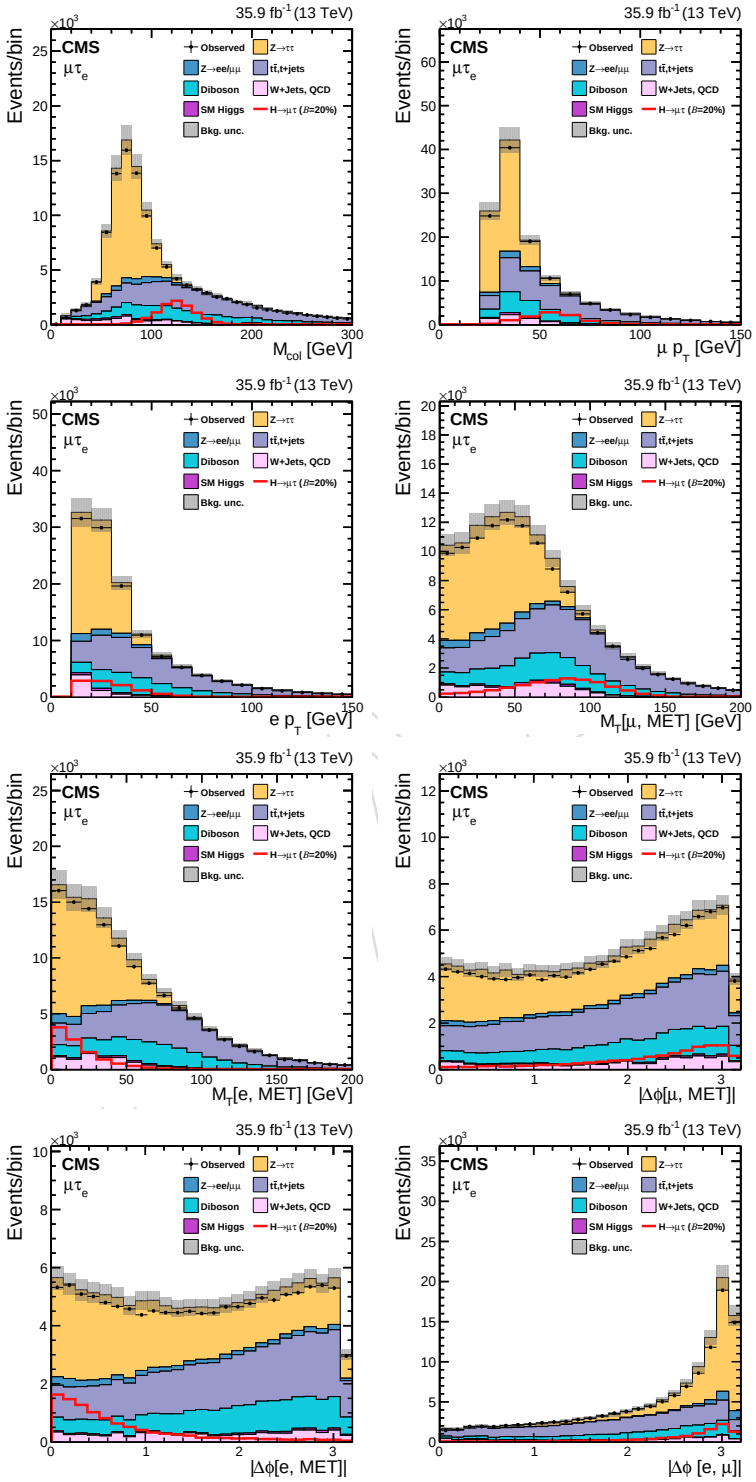

Figure 2: Distributions of the input variables to the BDT for the H $\to \mu\tau_e$ channel.

Table 1: Event selection criteria for the kinematic variables for the $H \to \mu\tau$ channels.

Variable		$H \to \mu\tau_e$				$H \to \mu\tau_h$			
		0 jet	1 jet	2 jet		0 jet	1 jet	2 jet	
				ggH	VBF			ggH	VBF
M_{jj}	[GeV]	—	—	<550	>550	—	—	<550	>550
p_T^e	[GeV]	>10				—			
p_T^μ	[GeV]	>26				>26			
$p_T^{\tau_h}$	[GeV]	—				>30			
$\lvert\eta^e\rvert$		<2.4				—			
$\lvert\eta^\mu\rvert$		<2.4				<2.4			
$\lvert\eta^{\tau_h}\rvert$		—				<2.3			
I_{rel}^e		<0.1				—			
I_{rel}^μ		<0.15				<0.15			
		M_{col} fit selection							
p_T^μ	[GeV]	<30	—	—	—	—			
$M_T(\mu)$	[GeV]	<60	<40	<15	<15	—			
$M_T(\tau_h)$	[GeV]	—				<105	<105	<105	<85
$\Delta\phi(e, \vec{p}_T^{\,miss})$	[radians]	<0.7	<0.7	<0.5	<0.3	—			
$\Delta\phi(e, \mu)$	[radians]	>2.5	<1.0	—	—	—			

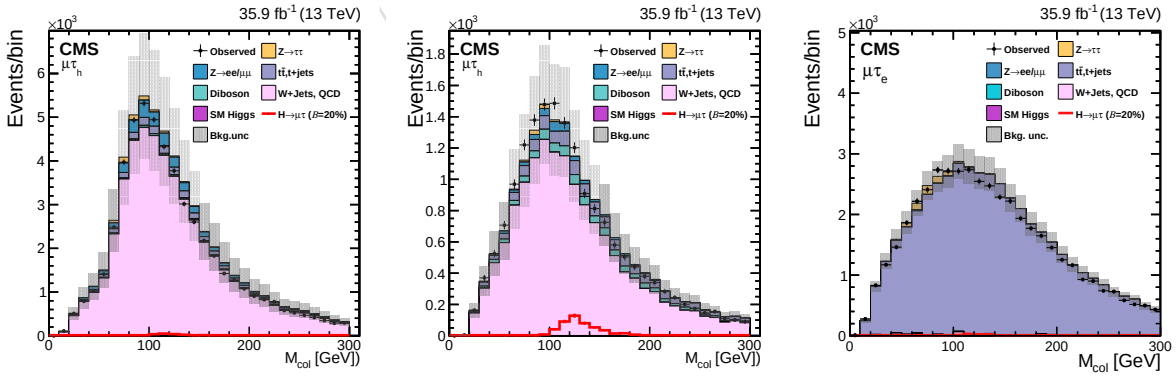

Figure 3: M_{col} distribution in like-sign lepton (left), W + jets enriched (central), and $t\bar{t}$ enriched (right) control regions defined in the text. The distributions are pre-fit and include both statistical and systematic uncertainties.

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 $I_{rel}^e < 0.15$. Muon candidates are required to have $p_T^\mu > 10$ GeV, $|\eta^\mu| < 2.4$, and $I_{rel}^\mu < 0.1$. The electron and the muon candidates are required to have opposite sign charges, and to be separated by $\Delta R > 0.5$. Events with additional electrons, muons, or tau candidates are rejected. Events with a b-tagged jet are rejected.

The tighter selection used in the M_{col} fit analysis further requires $\Delta\phi(e, \vec{p}_T^{miss}) < 1.0$ and $M_T(\mu) > 60$ GeV. The large $t\bar{t}$ background is reduced by vetoing events with b-tagged jets, and requiring $p_\zeta - 0.85 \cdot p_\zeta^{vis} > -60$ GeV. This topological selection is based on the projections

$$p_\zeta = (\vec{p}_T^{\,e} + \vec{p}_T^{\,\mu} + \vec{p}_T^{miss}) \cdot \frac{\vec{\zeta}}{|\vec{\zeta}|} \qquad \text{and} \qquad p_\zeta^{vis} = (\vec{p}_T^{\,e} + \vec{p}_T^{\,\mu}) \cdot \frac{\vec{\zeta}}{|\vec{\zeta}|}$$

on the axis $\vec{\zeta}$ bisecting the directions of the electron, $\vec{p}_T^{\,e}$, and of the muon, $\vec{p}_T^{\,\mu}$. This selction criterion is higly efficient in rejecting background as the $\vec{p}_T^{miss}$ is oriented in the direction of the visible τdecay products in signal events. The selections are summarised in Table 2.
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}, and the removal of $M_T(e)$. The background used for the training 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 is very significant in the 1-jet category.

Table 2: Event selection criteria for the kinematic variables for the $H \to e\tau$ channels.

Variable		$H \to e\tau_\mu$				$H \to e\tau_h$			
		0 jet	1 jet	2 jet		0 jet	1 jet	2 jet	
				ggH	VBF			ggH	VBF
M_{jj}	[GeV]	—	—	<500	>500	—	—	<500	>500
p_T^e	[GeV]	>26				>26			
p_T^μ	[GeV]	>10				—			
$p_T^{\tau_h}$	[GeV]	—				>30			
$\lvert\eta^e\rvert$		<2.1				<2.1			
$\lvert\eta^\mu\rvert$		<2.4				—			
$\lvert\eta^{\tau_h}\rvert$		—				<2.3			
I_{rel}^e		<0.1				<0.15			
I_{rel}^μ		<0.1				—			
		M_{col} fit selection							
$M_T(\tau_h)$	[GeV]	<60				<60			
$\Delta\phi(e, \vec{p}_T^{miss})$	[radians]	<1.0				—			
$p_\zeta - 0.85 \cdot p_\zeta^{vis}$	[GeV]	>-60				—			

7 Systematic uncertainties

The systematic uncertainties affect the normalization and the shape of the distributions of the different processes, and arise from either experimental or theoretical sources. They are summarized in Table 3. 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 data sets of Z bosons decaying to ee, $\mu\mu$, $\tau_\mu\tau_h$ [72–75, 84]. The b tagging efficiency in the simulation is adjusted to match the efficiency measured in data. The uncertainty in this measurement

Table 3: The systematic uncertainties in the expected event yields. All uncertainties are treated as correlated between the categories, except those that have two values separated by the $\oplus$ sign. 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. Theoretical uncertainties on SM VBF Higgs production [83] are also applied to SM VH production.

Systematic uncertainty	$H \to \mu\tau_e$	$H \to \mu\tau_h$	$H \to e\tau_\mu$	$H \to e\tau_h$
Muon trigger/identification/isolation	2%	2%	2%	-
Electron trigger/identification/isolation	2%	-	2%	2%
Hadronic tau lepton 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%$\oplus$5%	-	10%$\oplus$5%	-
$Z \to \tau\tau$ + jets background	10%$\oplus$5%	10%$\oplus$5%	10%$\oplus$5%	10%$\oplus$5%
W + jets background	10%	-	10%	-
QCD multijet background	30%	-	30%	-
WW, ZZ background	5%$\oplus$5%	5%$\oplus$5%	5%$\oplus$5%	5%$\oplus$5%
$t\bar{t}$ background	10%$\oplus$5%	10%$\oplus$5%	10%$\oplus$5%	10%$\oplus$5%
$W\gamma$ background	10%$\oplus$5%	-	10%$\oplus$5%	-
Single top quark background	5%$\oplus$5%	5%$\oplus$5%	5%$\oplus$5%	5%$\oplus$5%
$\mu \to \tau_h$ background	-	25%	-	-
$e \to \tau_h$ background	-	-	-	12%
jet $\to \tau_h, \mu, e$ background	-	30%$\oplus$10%	-	30%$\oplus$10%
jet energy scale	3–20%	3–20%	3-20%	3–20%
τ_h energy scale	-	1.2%	-	1.2%
$\mu, e \to \tau_h$ energy scale	-	1.5%	-	3%
e energy scale	0.1 – 0.5%	-	0.1 – 0.5%	0.1 – 0.5%
μ energy scale	0.2%	0.2%	0.2%	-
unclustered energy scale	$\pm 1\sigma$	$\pm 1\sigma$	$\pm 1\sigma$	$\pm 1\sigma$
renorm./fact. acceptance (ggH)	–(2.0 – 3.0)%			
renorm./fact. acceptance (VBF and VH)	–(0.3 – 1.0)%			
renorm./fact. scales (ggH) [83]	3.9%			
renorm./fact. scales (VBF and VH) [83]	0.4%			
PDF + α_s acceptance (ggH)	–(0.5 – 1.5)%			
PED + α_s acceptance (VBF and VH)	–(1.0 – 1.5)%			
PDF + α_s (ggH) [83]	3.2%			
PDF + α_s (VBF and VH) [83]	2.1%			
Integrated luminosity	2.5%			

is taken as the systematic uncertainty. The uncertainties on the $Z \to ee, Z \to \mu\mu, Z \to \tau\tau$, WW, ZZ, $W\gamma$, $t\bar{t}$, and single top production background contributions arise predominantly from the uncertainties in the measured cross sections of these processes. The uncertainties in the estimate of the misidentified-lepton backgrounds ($\mu \to \tau_h$, $e \to \tau_h$, jet $\to \tau_h, \mu, e$) 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 the effect of altering each source of jet energy scale uncertainty by one standard deviation to the fit templates of each process. This takes into account differences in yield and shape. The uncertainties on the e, μ, τ_h energy scale are propagated to the M_{col} and BDT distributions. For τ_h, the energy scale uncertainty is treated independently for each reconstructed hadronic decay mode of the τ lepton. The systematic uncertainties in the energy resolutions of lepton candidates have negligible effect. The energy scale of electrons (muons) misidentified as hadronically decaying tau candidates ($\mu, e \to \tau_h$ energy scale) is considered independently from true hadronic tau leptons. There is also an uncertainty in the unclustered energy scale. The unclustered energy comes from jets having $p_{\text{T}} < 10$ GeV and PF candidates not within jets. It is propagated to $p_{\text{T}}^{\text{miss}}$. 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 varying the energy of each particle by its uncertainty leads to changes in both shape of the distribution and yield. The four different systematic uncertainties are uncorrelated.

The uncertainties in the Higgs boson production cross sections due to the factorization and the renormalization scales as well as the parton distribution functions (PDF) and the strong coupling constant (α_s) result in changes in normalization and they are taken from Ref. [83]. They also affect the acceptance and lead to migration of events between the categories. They are listed as acceptance uncertainties in Table 3 and depend on the production process, Higgs boson decay channel, and category. Theoretical uncertainties in SM VBF Higgs production are also applied to SM VH production. They are 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 of 2% on the integrated luminosity [85] affects all processes with normalization taken directly from simulation. Shape uncertainties related to the pileup have been considered by varying the weights applied to simulation. The weight variation is obtained by a 5% change of the total inelastic cross section used to estimate the number of 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 produce alternate 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 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. A profile likelihood ratio is used as test statistic. The upper limits on the signal branching fraction are calculated with the asymptotic formula, using the CL_s criterion [86–88].

The BDT discriminator distributions of signal and background for each category are shown in Fig. 4 and 7 in the $H \to \mu\tau$ and $H \to e\tau$ channels respectively. Figures 5 and 8 show the corresponding M_{col} distributions used as cross-check. All the distributions are shown after

they have been adjusted by the fit. No excess over the background expectation is observed. The observed and median expected 95% CL upper limits, and best fit branching fractions, for $\mathcal{B}(H \to \mu\tau)$ and $\mathcal{B}(H \to e\tau)$, assuming m_H=125 GeV, are given for each category in Tables 5, 4, 7, and 6. The limits are also summarized graphically in Figs. 6 and 9.

Table 4: Expected and observed upper limits at 95% CL, and best fit branching fractions in percent for each individual jet category, and combined, in the $H \to \mu\tau$ 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				

Table 5: Expected and observed upper limits at 95% CL, and best fit branching fractions in percent for the individual jet categories for the $H \to \mu\tau$ 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				

No evidence is found for either the $H \to \mu\tau$ or $H \to e\tau$ processes in this search. The observed exclusion limits are a significant improvement over the 8 TeV results. The new results exclude the branching fraction that corresponded to the best fit for the 2.4 σ excess observed in the 8 TeV $H \to \mu\tau$ channel results at 95% CL, in both the M_{col} fit and BDT fit analysis. Table 8

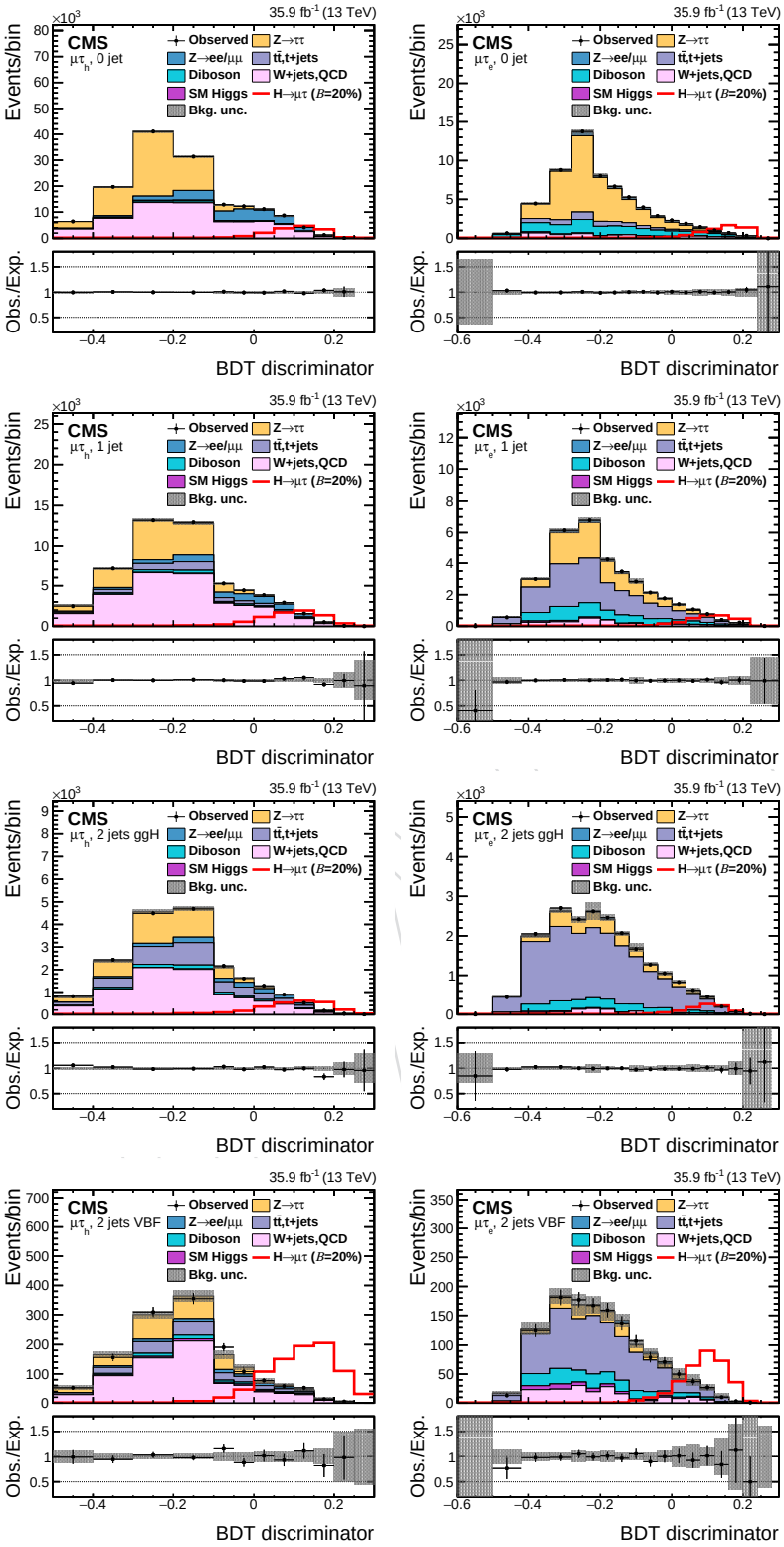

Figure 4: Distribution of the BDT discriminator for the $H \to \mu\tau$ process in the BDT fit analysis, in the individual channels and categories compared to the signal and background estimation. The background is normalized to the best fit values from the signal plus background fit while the simulated 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.

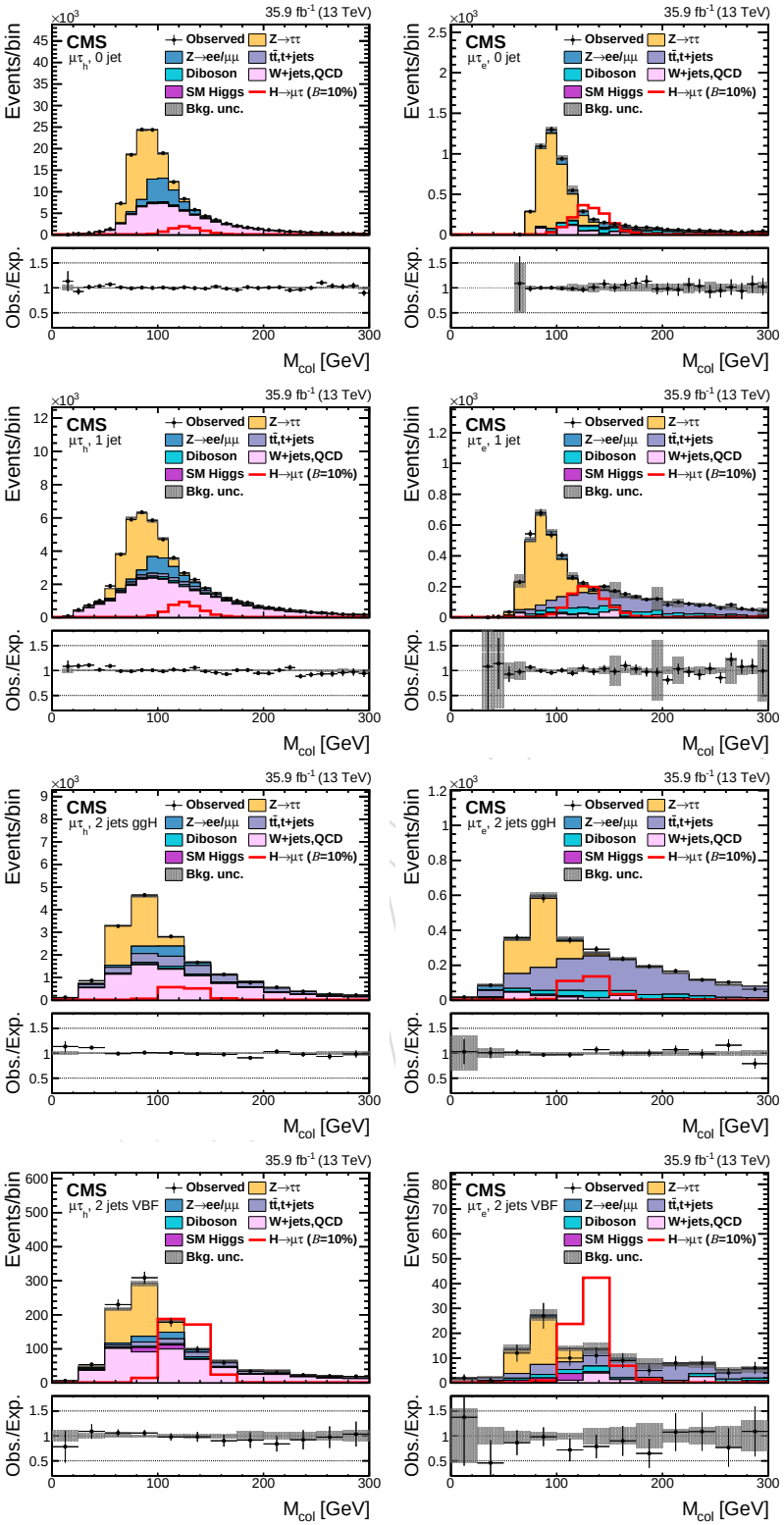

Figure 5: Distribution of the collinear mass M_{col} for the $\text{H} \to \mu\tau$ process in M_{col} fit analysis, in different channels and categories compared to the signal and background estimation. The background is normalized to the best fit values from the signal plus background fit while the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

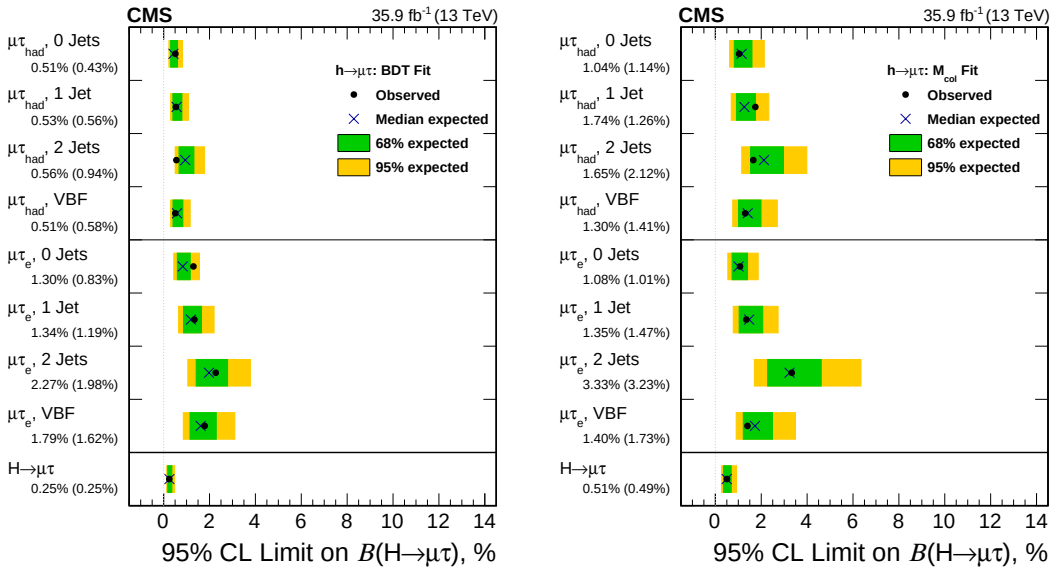

Figure 6: Observed and expected 95% CL upper limits on the $\mathcal{B}(\mathrm{H} \to \mu\tau)$ for each individual category and combined. Left: BDT fit analysis. Right: M_{col} fit analysis.

Table 6: Expected and observed upper limits at 95% CL and best fit branching fractions in percent for the individual jet categories in the $\mathrm{H} \to \mathrm{e}\tau$ process obtained with the BDT fit analysis.

Expected limits (%)					
	0-jet	1-jet	2-jets	VBF	Combined
$\mathrm{e}\tau_\mu$	< 0.90	< 1.59	< 2.54	< 1.84	< 0.64
$\mathrm{e}\tau_h$	< 0.79	< 1.13	< 1.59	< 0.74	< 0.49
$\mathrm{e}\tau$	< 0.37				
Observed limits (%)					
	0-jet	1-jet	2-jets	VBF	Combined
$\mathrm{e}\tau_\mu$	< 1.22	< 1.66	< 2.25	< 1.10	< 0.78
$\mathrm{e}\tau_h$	< 0.73	< 0.81	< 1.94	< 1.49	< 0.72
$\mathrm{e}\tau$	< 0.61				
Best fit branching fractions (%)					
	0-jet	1-jet	2-jets	VBF	Combined
$\mathrm{e}\tau_\mu$	0.47 ± 0.42	0.17 ± 0.79	-0.42 ± 1.01	-1.54 ± 0.44	0.18 ± 0.32
$\mathrm{e}\tau_h$	-0.13 ± 0.39	-0.63 ± 0.40	0.54 ± 0.53	0.70 ± 0.38	0.33 ± 0.24
$\mathrm{e}\tau$	0.30 ± 0.18				

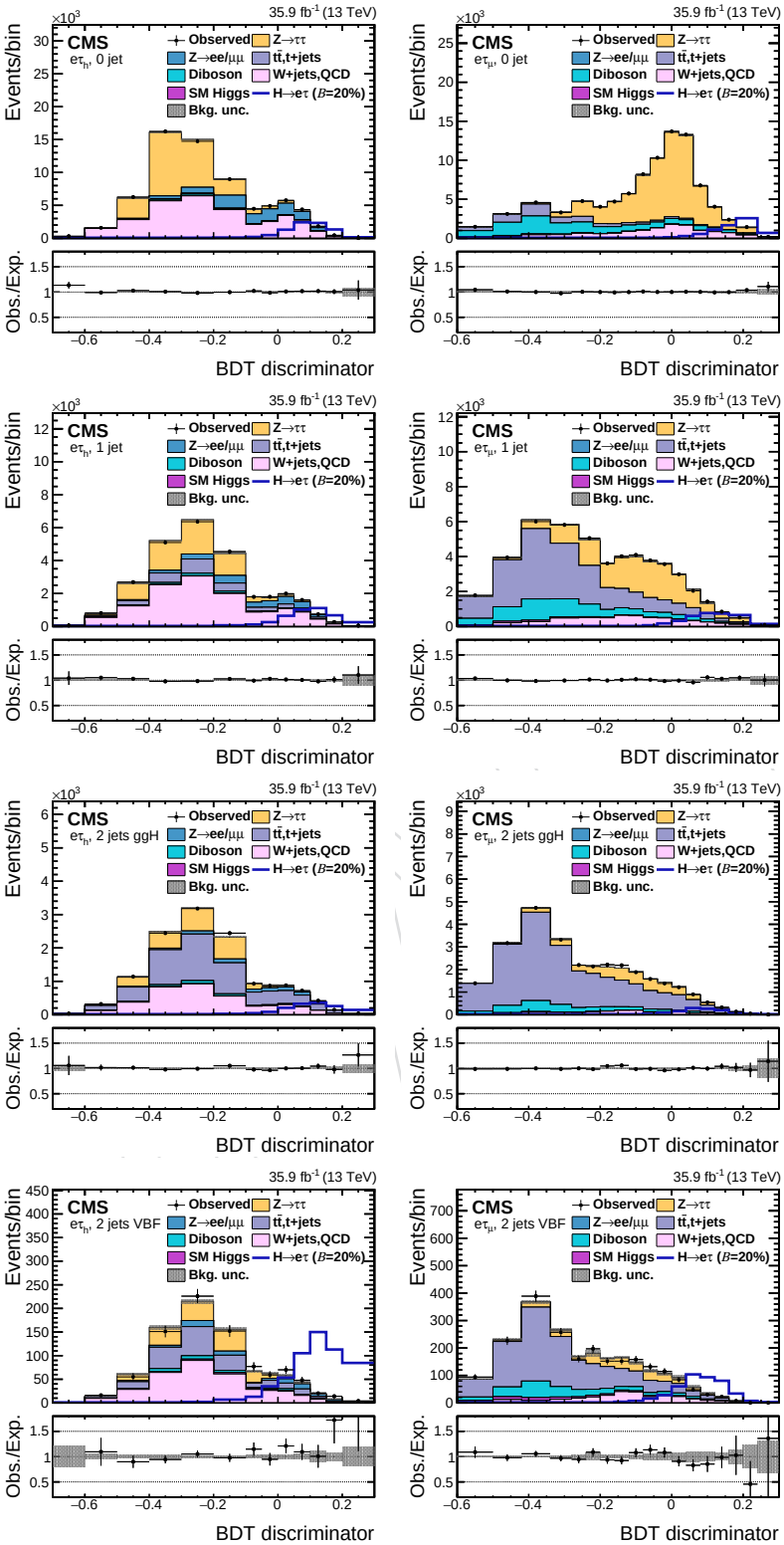

Figure 7: Distribution of the BDT discriminator for the H $\rightarrow$ eτ process for the BDT fit analysis, in different channels and categories compared to the signal and background estimation. The background is normalized to the best fit values from the signal plus background fit while the simulated signal corresponds to $\mathcal{B}(\mathrm{H} \rightarrow \mathrm{e}\tau) = 5\%$. The bottom panel in each plot shows the ratio between the observed data and the fitted background. The left column of plots corresponds to the H $\rightarrow$ eτ_h categories, from 0-jets (first row) to 2-jets VBF (fourth row). The right one to their H $\rightarrow$ eτ_μ counterparts.

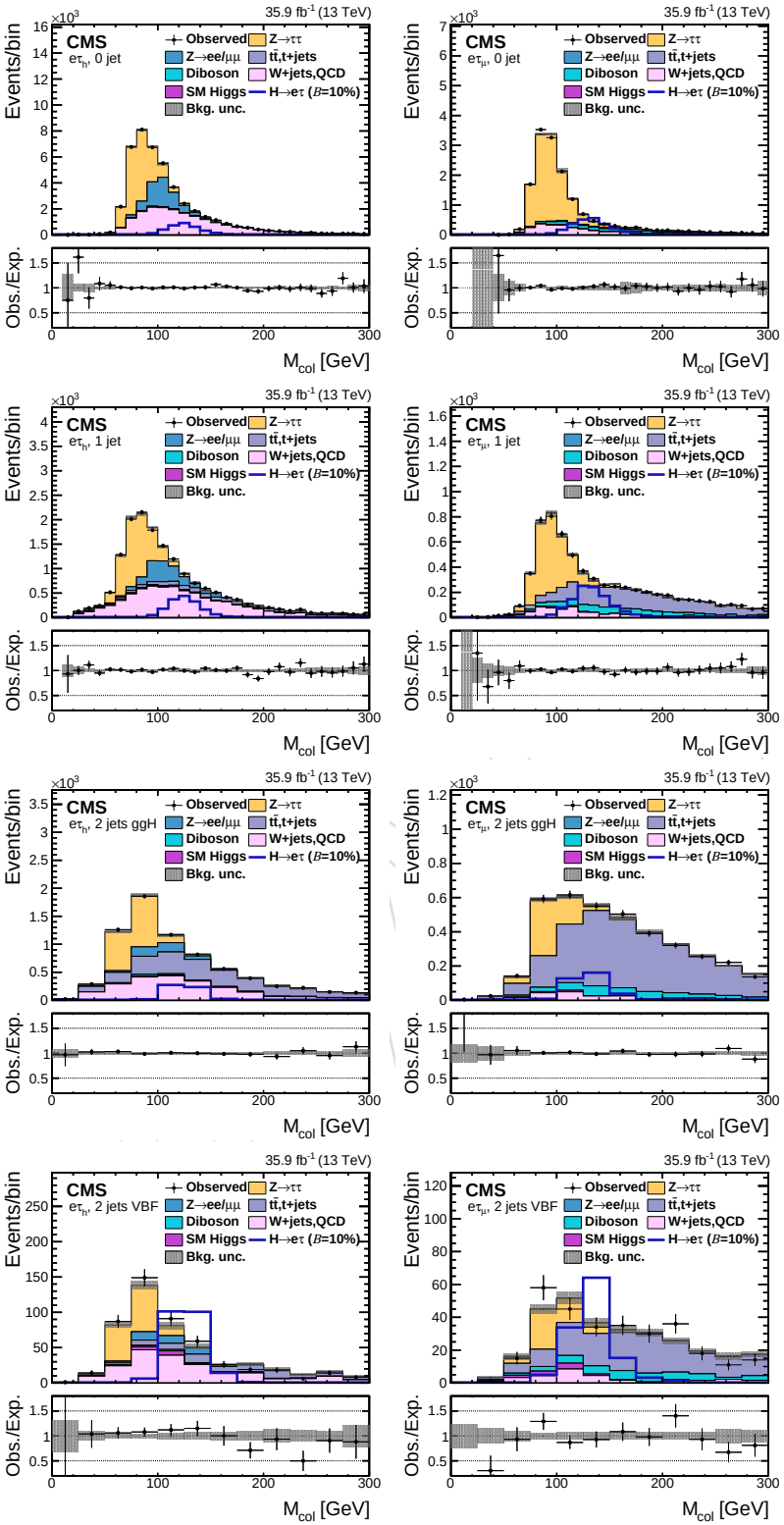

Figure 8: Distribution of the collinear mass M_{col} for the $H \to e\tau$ process in the M_{col} fit analysis, in different channels and categories compared to the signal and background estimation. The background is normalized to the best fit values from the signal plus background fit while the simulated signal corresponds to $\mathcal{B}(H \to e\tau) = 5\%$. The lower panel in each plot shows the ratio 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.

Table 7: Expected and observed upper limits at 95% CL and best fit branching fractions in percent for each individual jet category, and combined, in the $H \to e\tau$ process obtained with the M_{col} 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				

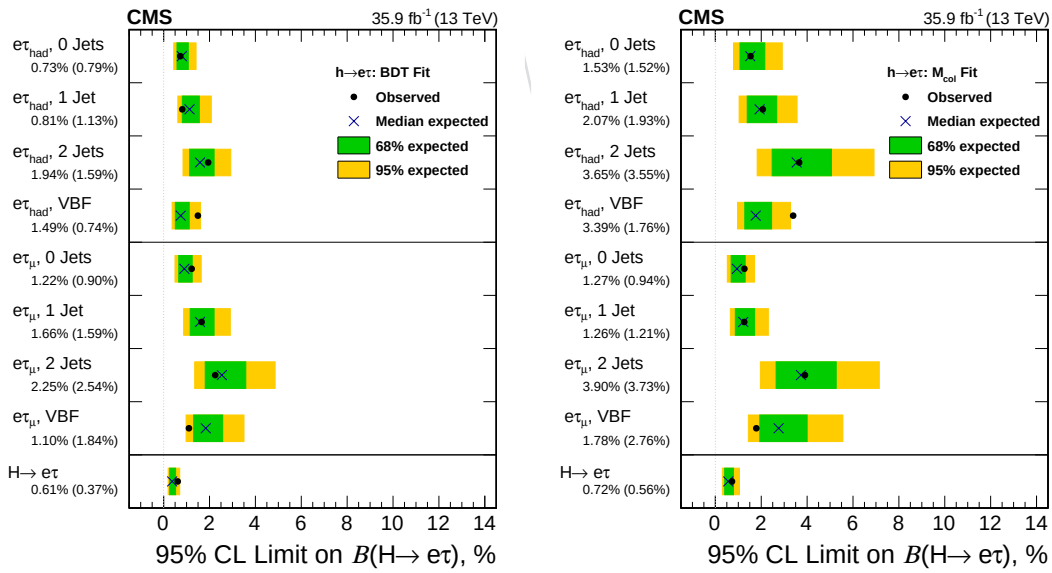

Figure 9: Observed and expected 95% CL upper limits on the $\mathcal{B}(H \to e\tau)$ for each individual category and combined. Left: BDT fit analysis. Right: M_{col} fit analysis.

shows a summary of the new 95% CL upper limits. The BDT fit analysis is more sensitive than the M_{col} fit analysis, with expected limits reduced by about a factor of two.

Table 8: Summary of the observed and expected upper limits at the 95% CL and the best fit branching fractions in percent for the $\text{H} \to \mu\tau$ and $\text{H} \to \text{e}\tau$ processes, for the main analysis (BDT fit) and the cross check (M_{col} fit) method.

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

The constraints on $\mathcal{B}(\text{H} \to \mu\tau)$ and $\mathcal{B}(\text{H} \to \text{e}\tau)$ can be interpreted in terms of LFV Yukawa couplings [41]. The LFV decays $\text{e}\tau$ 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, $\ell^\alpha, \ell^\beta = \text{e}, \mu, \tau$ and $\ell^\alpha \neq \ell^\beta$. The decay width $\Gamma(\text{H} \to \ell^\alpha \ell^\beta)$ in terms of the Yukawa couplings is given by:

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

and the branching fraction by:

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

The SM H decay width is assumed to be $\Gamma_{\text{SM}} = 4.1$ MeV [89] for $m_{\text{H}} = 125$ GeV. The 95% CL upper limit on the Yukawa couplings derived from the expression for the branching fraction above is shown in Table 9. The limits on the Yukawa couplings derived from the BDT fit analysis results are shown in Fig. 10.

Table 9: 95% CL observed upper limit on the Yukawa couplings, for the main analysis (BDT fit) and the cross check (M_{col} fit) method.

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

9 Summary

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

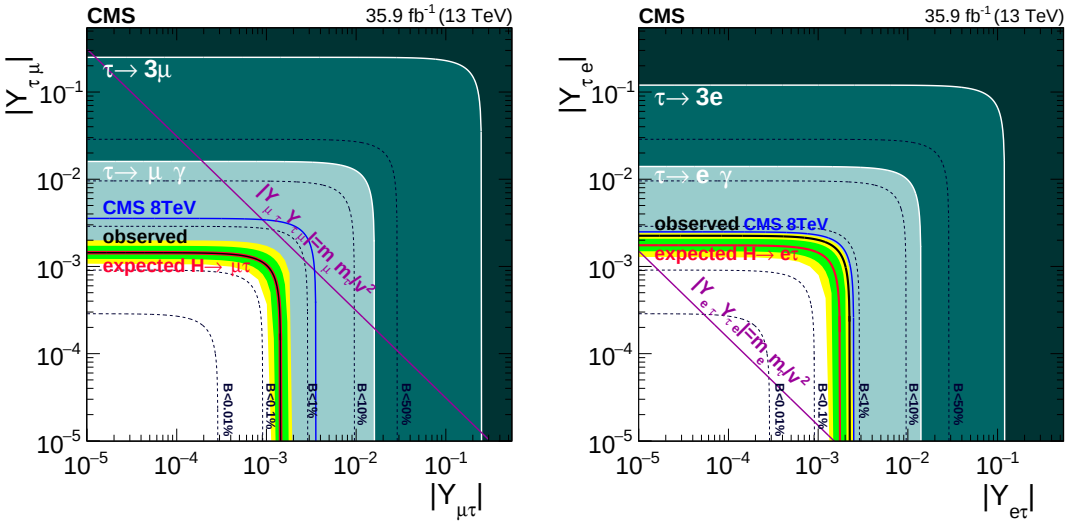

Figure 10: Constraints on the flavour violating Yukawa couplings, $|Y_{\mu\tau}|, |Y_{\tau\mu}|$ (left) and $|Y_{e\tau}|, |Y_{\tau e}|$ (right), from the BDT result. The expected (red solid line) and observed (black solid line) limits are derived from the limit on $\mathcal{B}(H \to \mu\tau)$ and $\mathcal{B}(H \to e\tau)$ 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 $\tau \to 3e$ (dark green) [41, 90, 91] and $\tau \to \mu\gamma$ or $\tau \to e\gamma$ (lighter green) [41, 91]. The blue solid lines are the CMS limits from [44] (left) and [45](right). The purple diagonal line is the theoretical naturalness limit $Y_{ij}Y_{ji} \leq m_i m_j / v^2$ [41].

Acknowledgments

We congratulate our colleagues in the CERN accelerator departments for the excellent performance of the LHC and thank the technical and administrative staffs at CERN and at other CMS institutes for their contributions to the success of the CMS effort. In addition, we gratefully acknowledge the computing centres and personnel of the Worldwide LHC Computing Grid for delivering so effectively the computing infrastructure essential to our analyses. Finally, we acknowledge the enduring support for the construction and operation of the LHC and the CMS detector provided by the following funding agencies: the Austrian Federal Ministry of Science, Research and Economy and the Austrian Science Fund; the Belgian Fonds de la Recherche Scientifique, and Fonds voor Wetenschappelijk Onderzoek; the Brazilian Funding Agencies (CNPq, CAPES, FAPERJ, and FAPESP); the Bulgarian Ministry of Education and Science; CERN; the Chinese Academy of Sciences, Ministry of Science and Technology, and National Natural Science Foundation of China; the Colombian Funding Agency (COLCIENCIAS); the Croatian Ministry of Science, Education and Sport, and the Croatian Science Foundation; the Research Promotion Foundation, Cyprus; the Secretariat for Higher Education, Science, Technology and Innovation, Ecuador; the Ministry of Education and Research, Estonian Research Council via IUT23-4 and IUT23-6 and European Regional Development Fund, Estonia; the Academy of Finland, Finnish Ministry of Education and Culture, and Helsinki Institute of Physics; the Institut National de Physique Nucléaire et de Physique des Particules / CNRS, and Commissariat à l'Énergie Atomique et aux Énergies Alternatives / CEA, France; the Bundesministerium für Bildung und Forschung, Deutsche Forschungsgemeinschaft, and Helmholtz-Gemeinschaft Deutscher Forschungszentren, Germany; the General Secretariat for Research and Technology, Greece; the National Scientific Research Foundation, and National Innova-

tion Office, Hungary; the Department of Atomic Energy and the Department of Science and Technology, India; the Institute for Studies in Theoretical Physics and Mathematics, Iran; the Science Foundation, Ireland; the Istituto Nazionale di Fisica Nucleare, Italy; the Ministry of Science, ICT and Future Planning, and National Research Foundation (NRF), Republic of Korea; the Lithuanian Academy of Sciences; the Ministry of Education, and University of Malaya (Malaysia); the Mexican Funding Agencies (BUAP, CINVESTAV, CONACYT, LNS, SEP, and UASLP-FAI); the Ministry of Business, Innovation and Employment, New Zealand; the Pakistan Atomic Energy Commission; the Ministry of Science and Higher Education and the National Science Centre, Poland; the Fundação para a Ciência e a Tecnologia, Portugal; JINR, Dubna; the Ministry of Education and Science of the Russian Federation, the Federal Agency of Atomic Energy of the Russian Federation, Russian Academy of Sciences, the Russian Foundation for Basic Research and the Russian Competitiveness Program of NRNU "MEPhI"; the Ministry of Education, Science and Technological Development of Serbia; the Secretaría de Estado de Investigación, Desarrollo e Innovación, Programa Consolider-Ingenio 2010, Plan de Ciencia, Tecnología e Innovación 2013-2017 del Principado de Asturias and Fondo Europeo de Desarrollo Regional, Spain; the Swiss Funding Agencies (ETH Board, ETH Zurich, PSI, SNF, UniZH, Canton Zurich, and SER); the Ministry of Science and Technology, Taipei; the Thailand Center of Excellence in Physics, the Institute for the Promotion of Teaching Science and Technology of Thailand, Special Task Force for Activating Research and the National Science and Technology Development Agency of Thailand; the Scientific and Technical Research Council of Turkey, and Turkish Atomic Energy Authority; the National Academy of Sciences of Ukraine, and State Fund for Fundamental Researches, Ukraine; the Science and Technology Facilities Council, UK; the US Department of Energy, and the US National Science Foundation.

Individuals have received support from the Marie-Curie programme and the European Research Council and Horizon 2020 Grant, contract No. 675440 (European Union); the Leventis Foundation; the A. P. Sloan Foundation; the Alexander von Humboldt Foundation; the Belgian Federal Science Policy Office; the Fonds pour la Formation à la Recherche dans l'Industrie et dans l'Agriculture (FRIA-Belgium); the Agentschap voor Innovatie door Wetenschap en Technologie (IWT-Belgium); the Ministry of Education, Youth and Sports (MEYS) of the Czech Republic; the Council of Scientific and Industrial Research, India; the HOMING PLUS programme of the Foundation for Polish Science, cofinanced from European Union, Regional Development Fund, the Mobility Plus programme of the Ministry of Science and Higher Education, the National Science Center (Poland), contracts Harmonia 2014/14/M/ST2/00428, Opus 2014/13/B/ST2/02543, 2014/15/B/ST2/03998, and 2015/19/B/ST2/02861, Sonata-bis 2012/07/E/ST2/01406; the National Priorities Research Program by Qatar National Research Fund; the Programa Clarín-COFUND del Principado de Asturias; the Thalis and Aristeia programmes cofinanced by EU-ESF and the Greek NSRF; the Rachadapisek Sompot Fund for Postdoctoral Fellowship, Chulalongkorn University and the Chulalongkorn Academic into Its 2nd Century Project Advancement Project (Thailand); and the Welch Foundation, contract C-1845.

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] ATLAS, CMS Collaboration, "Measurements of the Higgs boson production and decay rates and constraints on its couplings from a combined ATLAS and CMS analysis of the LHC pp collision data at $\sqrt{s} = 7$ and 8 TeV", *JHEP* **08** (2016) 045, doi:10.1007/JHEP08(2016)045, arXiv:1606.02266.

[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) 055010, doi:10.1103/PhysRevD.93.055010, arXiv:1508.04623.

[12] M. E. Gomez, S. Heinemeyer, and M. Rehman, "Lepton flavor violating Higgs Boson Decays in Supersymmetric High Scale Seesaw Models", arXiv:1703.02229.

[13] H.-B. Zhang et al., "125 GeV Higgs decay with lepton flavor violation in the $\mu\nu$SSM", *Chin. Phys. C* **41** (2017) 043106, doi:10.1088/1674-1137/41/4/043106, arXiv:1511.08979.

[14] 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.

[15] 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.

[16] 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.

[17] 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.

[18] 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.

[19] J. D. Bjorken and S. Weinberg, "Mechanism for nonconservation of muon number", *Phys. Rev. Lett.* **38** (1977) 622, doi:10.1103/PhysRevLett.38.622.

[20] J.-P. Lee and K. Y. Lee, "$B_s \to \mu\tau$ and $h \to \mu\tau$ decays in the general two Higgs doublet model", arXiv:1612.04057.

[21] 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.

[22] 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.

[23] 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.

[24] 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.

[25] 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.

[26] 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.

[27] 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.

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

[29] 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.

[30] 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) 015001, doi:10.1103/PhysRevD.91.015001, arXiv:1405.4300.

[31] J. Herrero-Garcia, N. Rius, and A. Santamaria, "Higgs lepton flavour violation: UV completions and connection to neutrino masses", *JHEP* **11** (2016) 084, doi:10.1007/JHEP11(2016)084, arXiv:1605.06091.

[32] F. del Aguila et al., "Lepton Flavor Changing Higgs decays in the Littlest Higgs Model with T-parity", arXiv:1705.08827.

[33] N. H. Thao, L. T. Hue, H. T. Hung, and N. T. Xuan, "Lepton flavor violating Higgs boson decays in seesaw models: new discussions", *Nucl. Phys. B* **921** (2017) 159, doi:10.1016/j.nuclphysb.2017.05.014, arXiv:1703.00896.

[34] I. Galon and J. Zupan, "Dark sectors and enhanced $h \to \tau\mu$ transitions", *JHEP* **05** (2017) 083, doi:10.1007/JHEP05(2017)083, arXiv:1701.08767.

[35] E. Arganda et al., "Effective lepton flavor violating $H\ell_i\ell_j$ vertex from right-handed neutrinos within the mass insertion approximation", *Phys. Rev. D* **95** (2017) 095029, doi:10.1103/PhysRevD.95.095029, arXiv:1612.09290.

[36] D. Choudhury, A. Kundu, S. Nandi, and S. K. Patra, "Unified resolution of the $R(D)$ and $R(D^*)$ anomalies and the lepton flavor violating decay $h \to \mu\tau$", *Phys. Rev. D* **95** (2017) 035021, doi:10.1103/PhysRevD.95.035021, arXiv:1612.03517.

[37] 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.

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

[39] 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.

[40] 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.

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

[42] 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.

[43] MEG Collaboration, "Search for the lepton flavour violating decay $\mu^+ \to e^+\gamma$ with the full dataset of the MEG experiment", *Eur. Phys. J. C* **76** (2016) 434, doi:10.1140/epjc/s10052-016-4271-x, arXiv:1605.05081.

[44] 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.

[45] 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", *Phys. Lett. B* **763** (2016) 472, doi:10.1016/j.physletb.2016.09.062, arXiv:1607.03561.

[46] ATLAS Collaboration, "Search for lepton-flavour-violating decays of the Higgs and Z bosons with the ATLAS detector", *Eur. Phys. J. C* **77** (2017) 70, doi:10.1140/epjc/s10052-017-4624-0, arXiv:1604.07730.

[47] 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.

[48] 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.

[49] 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.

[50] CMS Collaboration, "Observation of the Higgs boson decay to a pair of τ leptons", arXiv:1708.00373. Submitted to *Phys. Lett. B*.

[51] 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.

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

[53] 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.

[54] 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.

[55] 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.

[56] 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.

[57] 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.

[58] 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.

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

[60] 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.

[61] G. Luisoni, P. Nason, C. Oleari, and F. Tramontano, "$HW^{\pm}/HZ + 0$ and 1 jet at NLO with the POWHEG BOX interfaced to GoSam and their merging within MiNLO", *JHEP* **10** (2013) 083, doi:10.1007/JHEP10(2013)083, arXiv:1306.2542.

[62] 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.

[63] J. Alwall et al., "Comparative study of various algorithms for the merging of parton showers and matrix elements in hadronic collisions", *Eur. Phys. J. C* **53** (2008) 473, doi:10.1140/epjc/s10052-007-0490-5, arXiv:0706.2569.

[64] R. Frederix and S. Frixione, "Merging meets matching in MC@NLO", *JHEP* **12** (2012) 061, doi:10.1007/JHEP12(2012)061, arXiv:1209.6215.

[65] T. Sjostrand et al., "An introduction to PYTHIA 8.2", *Comput. Phys. Commun.* **191** (2015) 159, doi:10.1016/j.cpc.2015.01.024, arXiv:1410.3012.

[66] CMS Collaboration, "Event generator tunes obtained from underlying event and multiparton scattering measurements", *Eur. Phys. J. C* **76** (2016) 155, doi:10.1140/epjc/s10052-016-3988-x, arXiv:1512.00815.

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

[68] CMS Collaboration, "Particle-flow reconstruction and global event description with the CMS detector", arXiv:1706.04965. Submitted to *JINST*.

[69] 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.

[70] 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.

[71] 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.

[72] 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.

[73] 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.

[74] 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.

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

[76] M. Cacciari, G. P. Salam, "Dispelling the N^3 myth for the k_t jet-finder", *Phys. Lett. B* **641** (2006) 57, doi:10.1016/j.physletb.2006.08.037, arXiv:hep-ph/0512210.

[77] 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.

[78] CMS Collaboration, "Jet energy scale and resolution in the CMS experiment in pp collisions at 8 TeV", *JINST* **12** (2017) P02014, doi:10.1088/1748-0221/12/02/P02014, arXiv:1607.03663.

[79] 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.

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

[81] 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.

[82] CMS Collaboration Collaboration, "Identification of b quark jets at the CMS Experiment in the LHC Run 2",.

[83] LHC Higgs Cross Section Working Group Collaboration, "Handbook of LHC Higgs cross sections: 4. Deciphering the nature of the Higgs sector", arXiv:1610.07922.

[84] CMS Collaboration, "Measurements of inclusive w and z cross sections in pp collisions at $\sqrt{s} = 7$ tev", *JHEP* **01** (2011) 80, doi:10.1007/JHEP01(2011)080, arXiv:1012.2466.

[85] CMS Collaboration Collaboration, "CMS Luminosity Measurements for the 2016 Data Taking Period", Technical Report CMS-PAS-LUM-17-001, CERN, Geneva, 2017.

[86] 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.

[87] 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.

[88] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, "Asymptotic formulae for likelihood-based tests of new physics", *Eur. Phys. J. C* **71** (2011) 1554, doi:10.1140/epjc/s10052-011-1554-0, arXiv:1007.1727. [Erratum: doi:10.1140/epjc/s10052-013-2501-z].

[89] 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.

[90] K. Hayasaka et al., "Search for lepton flavor violating τ decays into three leptons with 719 million produced $\tau^+\tau^-$ pairs", *Phys. Lett. B* **687** (2010) 139, doi:10.1016/j.physletb.2010.03.037, arXiv:1001.3221.

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