CMS Physics Analysis Summary

Contact: cms-pag-conveners-higgs@cern.ch 2016/06/09

Search for lepton flavour violating decays of the Higgs boson in the μ-τ final state at 13 TeV

The CMS Collaboration

Abstract

A direct search for lepton flavour violating decays of the Higgs boson in the $H \to \mu\tau$ channel is described. In particular, the search examines the $H \to \mu\tau_e$ and the $H \to \mu\tau_h$ channels, where the τ leptons are reconstructed in the electronic and hadronic decay channels respectively. The data sample used in the search was collected in proton-proton collisions at $\sqrt{s} = 13$ TeV with the CMS experiment at the LHC and corresponds to an integrated luminosity of 2.3 fb^{-1}. No excess is observed, and a 95% CL upper limit of $\mathcal{B}(H \to \mu\tau) < 1.20\%$ (1.62 expected) is obtained.

1 Introduction

The discovery of the Higgs boson (H) [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 [4]. Such decays can occur naturally in models with more than one Higgs boson doublet [5]. They also arise in supersymmetric models [6–8], composite Higgs models [9, 10], models with flavour symmetries [11], Randall-Sundrum models [12–14], and others [15–22].

The CMS Collaboration has published a search in the $\mathrm{H} \to \mu\tau$ channel [23]. The data sample used in this search was collected in proton-proton collisions at a centre-of-mass energy of $\sqrt{s} = 8$ TeV and corresponds to an integrated luminosity of 19.7 fb^{-1}. 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 $\mathcal{B}(\mathrm{H} \to \mu\tau) = (0.84^{+0.39}_{-0.37})\%$. A small excess of data is observed with respect to the SM background-only hypothesis at $\mathrm{M_H} = 125\,\mathrm{GeV}$ with a significance of 2.4σ. A similar search in the $\mathrm{H} \to \mathrm{e}\tau$ and $\mathrm{H} \to \mathrm{e}\mu$ channels was performed by CMS also at a centre-of-mass energy of $\sqrt{s} = 8$ TeV [24] and found no evidence for lepton flavour violating decays in either final state. Upper limits on the branching fractions, $\mathcal{B}(\mathrm{H} \to \mathrm{e}\tau) < 0.69\%$ and $\mathcal{B}(\mathrm{H} \to \mathrm{e}\mu) < 0.035\%$, were set at the 95% CL. The ATLAS Collaboration finds a mild deviation of 1σ significance in the $\mathrm{H} \to \mu\tau$ channel and sets an upper limit of $\mathcal{B}(\mathrm{H} \to \mu\tau) < 1.43\%$ at 95% CL with a best fit branching fraction of $\mathcal{B}(\mathrm{H} \to \mu\tau) = (0.53 \pm 0.51)\%$ [25, 26]. The ATLAS collaboration also reports a search in the $\mathrm{H} \to \mathrm{e}\tau$ channel [26] in which no significant excess is observed and an upper limit of $\mathcal{B}(\mathrm{H} \to \mathrm{e}\tau) < 1.04\%$ at 95% CL is set. These two searches are based on the data sample of collisions collected by the ATLAS detector corresponding to an integrated luminosity of 20.3 fb^{-1} at a centre-of-mass energy of $\sqrt{s} = 8$ TeV. The presence of LFV Higgs boson couplings allows the transitions, $\mu \to \mathrm{e}$, $\tau \to \mu$, and $\tau \to \mathrm{e}$ to proceed via a virtual Higgs boson [27, 28]. The experimental limits on these decays have been translated into constraints on the branching fractions $\mathcal{B}(\mathrm{H} \to \mathrm{e}\mu)$, $\mathcal{B}(\mathrm{H} \to \mu\tau)$ and $\mathcal{B}(\mathrm{H} \to \mathrm{e}\tau)$ [4, 29]. The null search for $\mu \to \mathrm{e}\gamma$ [30] strongly constrains the $\mu \to \mathrm{e}$ transition $\mathcal{B}(\mathrm{H} \to \mathrm{e}\mu) < \mathcal{O}(10^{-8})$.

This analysis summary describes a search for LFV decays of the Higgs boson with $\mathrm{M_H} = 125\,\mathrm{GeV}$, based on proton-proton collision data recorded at $\sqrt{s} = 13$ TeV, corresponding to an integrated pp luminosity of 2.3 fb^{-1}. The search is performed in two channels, $\mathrm{H} \to \mu\tau_\mathrm{e}$, and $\mathrm{H} \to \mu\tau_\mathrm{h}$, where τ_e and τ_h are tau leptons reconstructed in the electronic and hadronic decay channels, respectively. The muonic decay of the τ lepton, τ_μ, is not considered due to the large background contribution from $\mathrm{Z} \to \mu\mu$ decays. The expected final state signatures are very similar to the SM $\mathrm{H} \to \tau_\mu\tau_\mathrm{h}$ and $\mathrm{H} \to \tau_\mathrm{e}\tau_\mu$ decays, which have been studied by CMS [31, 32] and ATLAS [33], but with some significant kinematic differences. The muon in the LFV $\mathrm{H} \to \mu\tau$ decay is produced promptly, and tends to have a larger momentum than in the SM $\mathrm{H} \to \tau_\mu\tau_\mathrm{h}$ decay.

This analysis summary is organized as follows. A brief description of the CMS detector is given in Section 2; the collision data and simulated samples used in the analysis are listed in Section 3; the event reconstruction and selection are described in Section 4; the estimation of the background and its components are studied in Section 5; the systematic uncertainties are reviewed in Section 6; and the results are presented in Section 7.

2 The CMS detector

A detailed description of the CMS detector, together with the definition of the coordinate system used and the relevant kinematic variables, can be found in Ref [34]. The momenta of charged particles are measured with a silicon pixel and strip tracker that covers the pseudora-

pidity 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-fiber 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 measure the momenta of 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 at a maximum rate of 100 kHz. The software algorithms of the high-level trigger, executed on a farm of commercial processors, reduce the event rate to approximately 1 kHz using information from all detector subsystems.

3 Collision data and simulated events

The trigger used to select the data sample analyzed requires a single muon with transverse momentum (p_T) of at least 20 GeV in the pseudorapidity range $|\eta| < 2.4$. Simulated samples of signal and background events are produced with several event generators. The CMS detector response is modelled using GEANT4 [35]. The Higgs bosons are produced in proton-proton collisions predominantly by gluon-gluon fusion (GF). Higgs bosons are also produced by vector boson fusion (VBF), in association with a W or Z boson, or in association with two top quarks. The SM Higgs boson samples considered in this analysis are GF and VBF $H \to \tau\tau$ decays. The $H \to \mu\tau$ and SM Higgs boson samples are generated using POWHEG 1.0 [36–40], with CT10 parton distribution functions, interfaced to PYTHIA 8.212 [41]. The MADGRAPH 5.1.3.30 [42] generator is used for Z+jets, W+jets, $t\bar{t}$, and diboson production, and POWHEG for single top quark production. The POWHEG and MADGRAPH generators are interfaced to PYTHIA for parton shower and fragmentation. The PYTHIA parameters for the underlying event description are set to the CUETP8M1 tune [43]. Due to the high luminosities attained during data-taking, many events have multiple proton-proton interactions per bunch crossing (pileup). The presence of pileup interactions is incorporated by simulating additional protonproton collisions with PYTHIA. An event weight based on the number of simulated pileup events and the instantaneous luminosity per bunch-crossing is applied to match the pileup distribution observed in data.

4 Event reconstruction and selection

The tracking system is able to separate collision vertices as close as 0.5 mm to each other along the beam direction [44]. 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 also affect the identification of most of the physics objects, such as jets, and variables such as lepton isolation.

A particle-flow (PF) algorithm [45–47] 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 then required to be consistent with the primary vertex and used to reconstruct jets, hadronic τ decays, quantify the isolation of leptons and photons and reconstruct E_T^{miss}. The missing transverse energy vector, $\vec{E}_T^{miss}$, is defined as the negative of the vector p_T sum of all identified PF objects in the event [48]. 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 [49]. 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, 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 [50]. 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 [51], 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 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. Dedicated discriminators veto against electrons and muons.

Jets misidentified as electrons, muon or tau leptons are suppressed by imposing isolation requirements, summing the neutral and charged particle contributions in cones of $\Delta R = 0.4$ around the lepton. 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 [52], 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. [51].

Jets are reconstructed from all the particles using the anti-k_t jet clustering algorithm [53] 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 [54]. Any jet within $\Delta R = 0.4$ of the identified leptons is removed.

The event selection consists of three steps, based on the electron, muon, hadronically decaying τ lepton and jet objects just described. First, a loose selection exploiting the basic signature is applied. The sample is then divided into categories, according to the number of jets in the event. Finally, requirements are placed on a set of kinematic variables designed to suppress the backgrounds.

The loose selection for the $H \to \mu\tau_e$ channel requires an isolated μ ($p_T > 25\,\text{GeV}$, $|\eta| < 2.1$) [55] and an isolated e ($p_T > 10\,\text{GeV}$, $|\eta| < 2.3$) [49] of opposite charge. The $H \to \mu\tau_h$ channel requires an isolated μ ($p_T > 30\,\text{GeV}$, $|\eta| < 2.1$) and an isolated hadronically decaying τ ($p_T >$

30 GeV, $|\eta| < 2.3$) [51] of opposite charge.

Events in the $H \to \mu\tau_e$ channel with additional muons with $p_T > 10$ GeV or electrons with $p_T > 10$ GeV are rejected. In the $H \to \mu\tau_h$ channel events with additional muons with $p_T > 5$ GeV, electrons with $p_T > 10$ GeV or τ_h leptons with $p_T > 20$ GeV are also rejected.

The events are then divided into three categories within each channel according to the number of jets in the event (zero, one, or two jets). Jets are required to have $p_T > 30$ GeV and lie within the range $|\eta| < 4.7$.

The collinear mass (M_{col}) provides an estimate of M_H using the observed decay products. It is constructed using the collinear approximation based on the observation that, since $M_H \gg M_\tau$, the τ decay products are highly Lorentz boosted in the direction of the τ lepton [56]. 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 $\vec{E}_T^{miss}$ in the direction of the visible τ decay products is used to estimate the transverse component of the neutrino momentum ($p_T^{\nu,\,est}$). M_{col} can be then derived from the visible mass of the $\tau - e$ system (M_{vis}) as $M_{col} = M_{vis}/\sqrt{x^{\tau^{vis}}}$, where $x_{\tau^{vis}}$ is the fraction of energy carried by the visible decay products of the τ lepton ($x^{\tau^{vis}} = p_T^{\tau^{vis}}/(p_T^{\tau^{vis}} + p_T^{\nu,\,est})$).

The collinear mass (M_{col}) is used as the final discriminant variable to separate the signal from the background. Figure 1 shows the M_{col} distribution for the data, compared to the esimated background and to the signal, for each of the categories in each channel after the loose selection. The simulated signal for $\mathcal{B}(H \to \mu\tau) = 100\%$ is shown. The backgrounds are estimated using data and simulation as described in Section 5. There is good agreement between data and the background estimation.

The final selection is based on a set of kinematical variables and selection criteria which were optimized for the 8 TeV analysis by maximizing $S/\sqrt{S+B}$ where S and B are the expected signal and background event yields in the mass window $100 < M_{col} < 150$ GeV [23]. This analysis uses the same criteria with looser requirements in the 2-jet categories because of the reduced statistics of the 2015 data sample. The selection criteria for each category, and in each channel, are given in Table 1. The variables used are the lepton transverse momenta p_T^ℓ with $\ell = \tau_h, \mu, e$; azimuthal angles between the leptons $\Delta\phi_{\vec{p}_T^{\ell_1} - \vec{p}_T^{\ell_2}}$; azimuthal angle $\Delta\phi_{\vec{p}_T^{\ell} - \vec{E}_T^{miss}}$; the transverse mass $M_T^\ell = \sqrt{2 p_T^\ell E_T^{miss}(1 - \cos\Delta\phi_{\vec{p}_T^{\ell} - \vec{E}_T^{miss}})}$. Events in the 2-jet category are required to have exactly two jets separated by a large pseudorapidity gap and with a large invariant mass. As such, the zero and one jet categories contain signal events predominantly produced by gluon-gluon fusion and the two jet category is enriched with signal events produced by vector boson fusion. In the $H \to \mu\tau_e$ channel events in which at least one of the jets is identified as coming from a b-quark decay using the combined secondary-vertex b-tagging algorithm [57] are vetoed, to suppress backgrounds from top quark decays.

5 Background Modelling

The largest backgrounds come from misidentified leptons in W+jets and QCD multijet production, and from the $Z \to \tau\tau$ process. The contribution of the misidentified lepton background is estimated using data, while the remaining backgrounds are estimated using simulation.

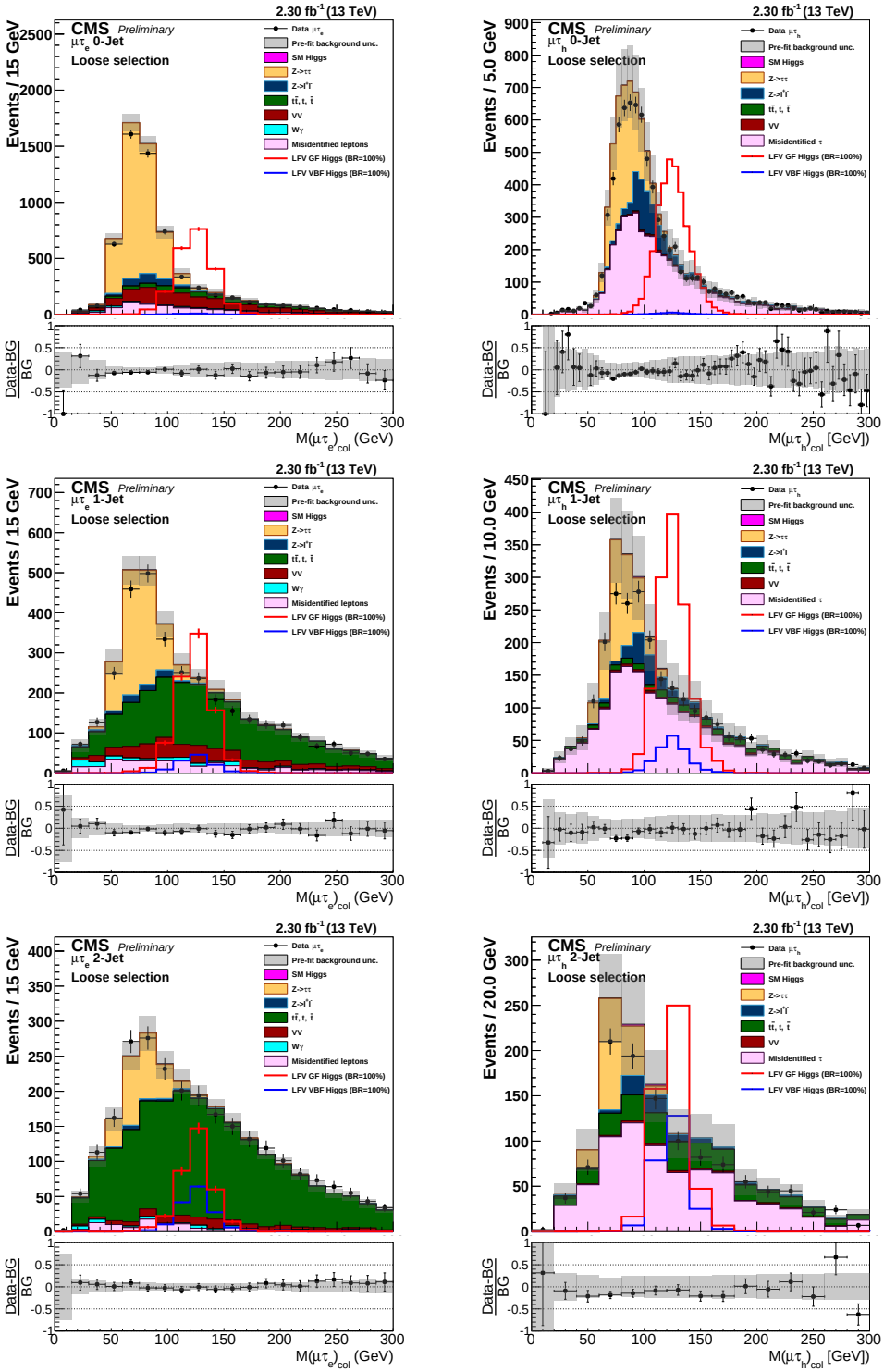

Figure 1: Distributions of the collinear mass M_{col} for signal and background processes after the loose selection requirements, for the LFV $H \to \mu\tau$ candidates, for the different channels and categories, compared to data. For visualization purposes $\mathcal{B}(H \to \mu\tau) = 100\%$ is used for the signal. The shaded grey bands indicate the total uncertainty. The bottom panel in each plot shows the fractional difference between the observed data and the total estimated background. Top left: $H \to \mu\tau_e$ 0-jet; top right: $H \to \mu\tau_h$ 0-jet; middle left: $H \to \mu\tau_e$ 1-jet; middle right: $H \to \mu\tau_h$ 1-jet; bottom left: $H \to \mu\tau_e$ 2-jet; bottom right $H \to \mu\tau_h$ 2-jet.

Table 1: Selection criteria in all the categories used in the analysis

Variable	$H \to \mu\tau_e$			$H \to \mu\tau_h$		
	0-jet	1-jet	2-jet	0-jet	1-jet	2-jet
$p_T^{\mu} >$ [GeV]	50	45	25	45	35	40
$p_T^{e} >$ [GeV]	15	15	15	—	—	—
$p_T^{\tau_h} >$ [GeV]	—	—	—	35	40	40
$M_T^{e} <$ [GeV]	65	65	40	—	—	—
$M_T^{\mu} >$ [GeV]	50	40	15	—	—	—
$M_T^{\tau_h} <$[GeV]	—	—	—	50	35	35
$\Delta\phi_{\vec{p}_T^{\mu}-\vec{p}_T^{\tau_h}} >$	—	—	—	2.7	—	—
$\Delta\phi_{\vec{p}_T^{e}-\vec{E}_T^{miss}} <$	0.5	0.5	—	—	—	—
$\Delta\phi_{\vec{p}_T^{e}-\vec{p}_T^{\mu}} >$	2.7	1.0	—	—	—	—
VBF dijet $\lvert\Delta\eta\rvert >$	—	—	2.5	—	—	2.5
VBF dijet mass > [GeV]	—	—	200	—	—	200

5.1 Misidentified leptons

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 for 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 in an independent collision data set, 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 2 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 with leptons of the same charge, and used for validation of the method. The application is slightly different in the $H \to \mu\tau_e$ and $H \to \mu\tau_h$ channels due to the different lepton isolation requirements used to define the enriched regions in each channel.

In the $H \to \mu\tau_e$ channel, region I is the signal region in which an isolated muon and an isolated electron are required. In region III all the analysis selection criteria are applied to the data sample except that one of the leptons is required to be non-isolated. This creates a region enriched with misidentified leptons. There are two types of events in this region: those with an isolated muon and a non-isolated electron and those with an isolated electron and a non-isolated muon. There is a negligible number of signal events in region III. Regions II and IV are defined using the same selection criteria as regions I and III, respectively, but same-sign leptons instead of opposite-sign leptons are required. It has been verified than the kinematic distributions of the same-sign samples are very similar to the opposite-sign samples.

The sample in region III is dominated by W+jets and QCD multijets, with small contributions from prompt leptons that are subtracted using simulation. The misidentified electron background in region I is estimated by multiplying the event yield in region III by a factor f_e, where f_e is the ratio of non-isolated to isolated electrons. It is computed in an independent data sample $Z \to \mu\mu + X$, where X is a PF object identified as an electron, in bins of p_T and η. In the estimation of f_e, background sources of prompt leptons, predominantly WZ and ZZ, are subtracted from the $Z \to \mu\mu + X$ sample using simulation. The misidentified μ background is computed in the same way. The technique is validated by using the same-sign data from regions II and IV. In Fig. 2 (left), the observed data yield in region II is compared to the estimate obtained by scaling the region IV sample by the measured misidentification rates. The region II sample is dominated by misidentified leptons but also includes small contributions of true

Table 2: Definition of the regions used to estimate the misidentified lepton background. The different regions have different requirements for the isolation and the relative charge of the two leptons $\ell_1^{\pm}$ and $\ell_2^{\pm}$, which can be e, μ or τ_h.

Opposite-sign leptons	Same-sign leptons
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)
$\ell_2^{\mp}$(non-isolated)	$\ell_2^{\pm}$(non-isolated)

leptons originating from vector boson decays, estimated with simulated samples.

In the $H \to \mu\tau_h$ channel, the τ_h candidate can arise from a misidentified jet coming from a number of sources, predominantly W+jets and QCD multijets, but also $Z \to \mu\mu$+jets and $t\bar{t}$. The misidentification rate f_{τ_h} is measured in $Z \to \mu\mu + X$ events selected in data, where X is an object identified as a τ_h candidate that passes a loose isolation requirement. f_{τ_h} is defined as the fraction of these loosely isolated τ_h candidates that also pass the tighter isolation requirement used to define the signal region. This misidentification rate measured in $Z \to \mu\mu + X$ collision data is compared to that measured in $Z \to \mu\mu + X$ simulation and found to be in good agreement.

The enriched background regions (III and IV) are defined by requiring the presence of τ_h candidates that pass the looser isolation requirement, but do not pass the tight isolation requirement. Small background sources of prompt leptons are subtracted using simulation. The misidentified background yield in the signal region (region I) is estimated by multiplying the event yield in region III by a factor $f_{\tau_h}/(1 - f_{\tau_h})$. The procedure is validated with same-sign $\mu\tau_h$ events (region II) using region IV to model the misidentified background, in the same way as for the $H \to \mu\tau_e$ channel. Figure 2 (right) shows the data in region II compared to the estimate obtained by scaling region IV by the misidentification fractions.

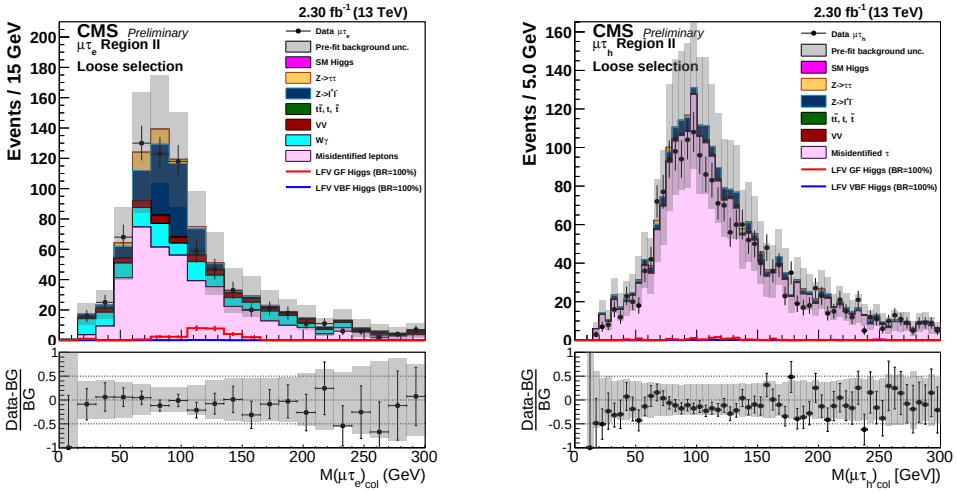

Figure 2: Distributions of M_{col} for region II compared to the estimate obtained by scaling the region IV sample by the measured misidentification fractions. The bottom panel in each plot shows the relative difference between the observed data and the estimate. Left: $H \to \mu\tau_e$. Right: $H \to \mu\tau_h$.

5.2 Other backgrounds

The background contribution from $Z \to \tau\tau$, $Z \to \mu\mu$, $t\bar{t}$, WW, ZZ, Wγ and single top quark production is estimated with simulation. The simulated events are corrected for residual discrepancies between data and simulation. These discrepancies, which are related to the electron and muon triggering, identification, and isolation, are determined using the tag-and-probe technique in $Z \to ll$ data [49, 55]. The background contribution coming from SM H decays in the $H \to \tau\tau$ channel is also estimated with simulation. This background is suppressed by the kinematic selection criteria and M_{col} peaks below 125 GeV.

6 Systematic Uncertainties

A profiled likelihood method is used to derive all results assuming the asymptotic approximation [58], based on the distributions of M_{col} for the signal and the various background sources. Systematic uncertainties are represented by nuisance parameters, some of which only affect the background and signal normalizations, while others also affect the shape of the M_{col} distributions. To set upper bounds on the branching fraction the CL_s method [59, 60] is used.

The uncertainties in the electron and muon selections (trigger, identification, and isolation efficiencies) are estimated using the tag and probe technique in $Z \to \mu\mu, ee$ data [49, 50]. The uncertainty in the τ_h efficiency is estimated to be 10%, consistent with [23, 51]. The misidentified μ, e and τ_h uncertainties are estimated from the agreement with data in a control region (region II).

The shape and normalization uncertainties related to jet energy scale are computed for each background by calculating the differences in yield and shape when the jet energy scale is altered by $\pm 1\sigma$, which results in a shift of 5% to 10% in jet p_T. A similar procedure is employed to estimate the shape and normalization uncertainties due to τ_h energy scale. A misidentified lepton shape uncertainty is computed by varying the lepton fake rate factors by $\pm 1\sigma$.

The theoretical uncertainties are calculated by adding the change in acceptance, when the factorization and renormalization scales are shifted by a factor of two, to the uncertainty in the Higgs production cross sections.

All systematic uncertainties are summarized in Table 3.

7 Results

After applying the full selection criteria, a maximum likelihood fit is performed in the M_{col} variable. Each systematic uncertainty is used as a nuisance parameter in the fit. The event yields in the mass range $100 < M_{\text{col}} < 150$ GeV are shown in Table 4 and the distributions of the signal and background contributions after the full selection and the fit are shown in Fig. 3. The different channels and categories are combined to set a 95% CL upper limit on the branching fraction of LFV H decay in the $\mu\tau$ channel, $\mathcal{B}(H \to \mu\tau)$.

The observed and median expected 95% CL upper limits on the $\mathcal{B}(H \to \mu\tau)$, assuming the Higgs boson mass to be 125 GeV, are given for each category in Table 5. Combining all the categories, an expected upper limit of $\mathcal{B}(H \to \mu\tau) < (1.62 \pm 0.58)\%$ is obtained. The observed upper limit is $\mathcal{B}(H \to \mu\tau) < 1.20\%$. The limits are also summarized graphically in Fig. 4. The best-fit branching fractions per category are also given in Table 5. The combination of all the categories yields a best-fit branching ratio of $\mathcal{B}(H \to \mu\tau) = -0.76^{+0.81}_{-0.84}\%$.

Table 3: Systematic uncertainties in the expected event yield. All uncertainties are treated as correlated between the categories, except those which have two values indicated. In this case the first value is correlated as above, while the second value (following the $\oplus$ symbol) represents an uncorrelated uncertainty for each individual category. The total uncertainty in a given category is the sum in quadrature of the two values.

Systematic uncertainty	$H \to \mu\tau_e$	$H \to \mu\tau_h$
Muon trigger/ID/isolation	3%	3%
Electron trigger/ID/isolation	3%	—
Hadronic τ efficiency	—	10%
b-tagging veto	3%	—
$Z \to \tau\tau$ background	10%$\oplus$5%	10%$\oplus$5%
$Z \to \mu\mu$, ee background	10%$\oplus$5%	10%$\oplus$5%
Misidentified μ, e background	40%$\oplus$10%	—
Misidentified τ_h background	—	30%$\oplus$10%
WW, ZZ background	10%$\oplus$5%	10%$\oplus$5%
$t\bar{t}$ background	20%$\oplus$5%	20%$\oplus$5%
$W + \gamma$ background	10%$\oplus$5%	—
Single top production background	10%	10%
Jet energy scale	3-20%	3-20%
Hadronic τ energy scale	—	3%
Misidentified lepton shape	$\pm\sigma$	$\pm\sigma$
Theory uncertainty	10%	10%
Luminosity	2.7%	2.7%

Table 4: Event yields in the signal region in the range $100 < M_{col} < 150\,\text{GeV}$. The expected contributions are normalized to an integrated luminosity of $2.3\,\text{fb}^{-1}$. The LFV Higgs boson signal indicated corresponds to $B(H \to \mu\tau) = 1\%$, with the expected SM Higgs boson cross section.

Sample	$H \to \mu\tau_e$			$H \to \mu\tau_h$		
	0-Jet	1-Jet	2-Jets	0-Jet	1-Jet	2-Jets
misidentified leptons	12.2	5.2	2.8	232.3	54.7	4.7
$Z \to \tau\tau$	14.4	10.6	1.7	5.3	2.3	0
ZZ, WW	10.7	4.6	3.2	3.2	2.0	0.3
$W\gamma$	1.2	3.4	0.9	—	—	—
$Z \to$ ee or $\mu\mu$	1.9	2.2	0.3	79.1	11.9	0.1
$t\bar{t}$	1.4	21.8	18.6	1.3	5.4	1.1
t, $\bar{t}$	0.4	4.1	1.7	0.3	2.2	0.2
SM H background	0.4	0.4	0.4	1.1	0.7	0.3
sum of backgrounds	42.6	52.2	29.6	322.5	79.3	6.6
LFV Higgs boson signal	7.1	3.7	1.9	13.8	4.7	1.2
Observed data	33	41	31	315	77	7

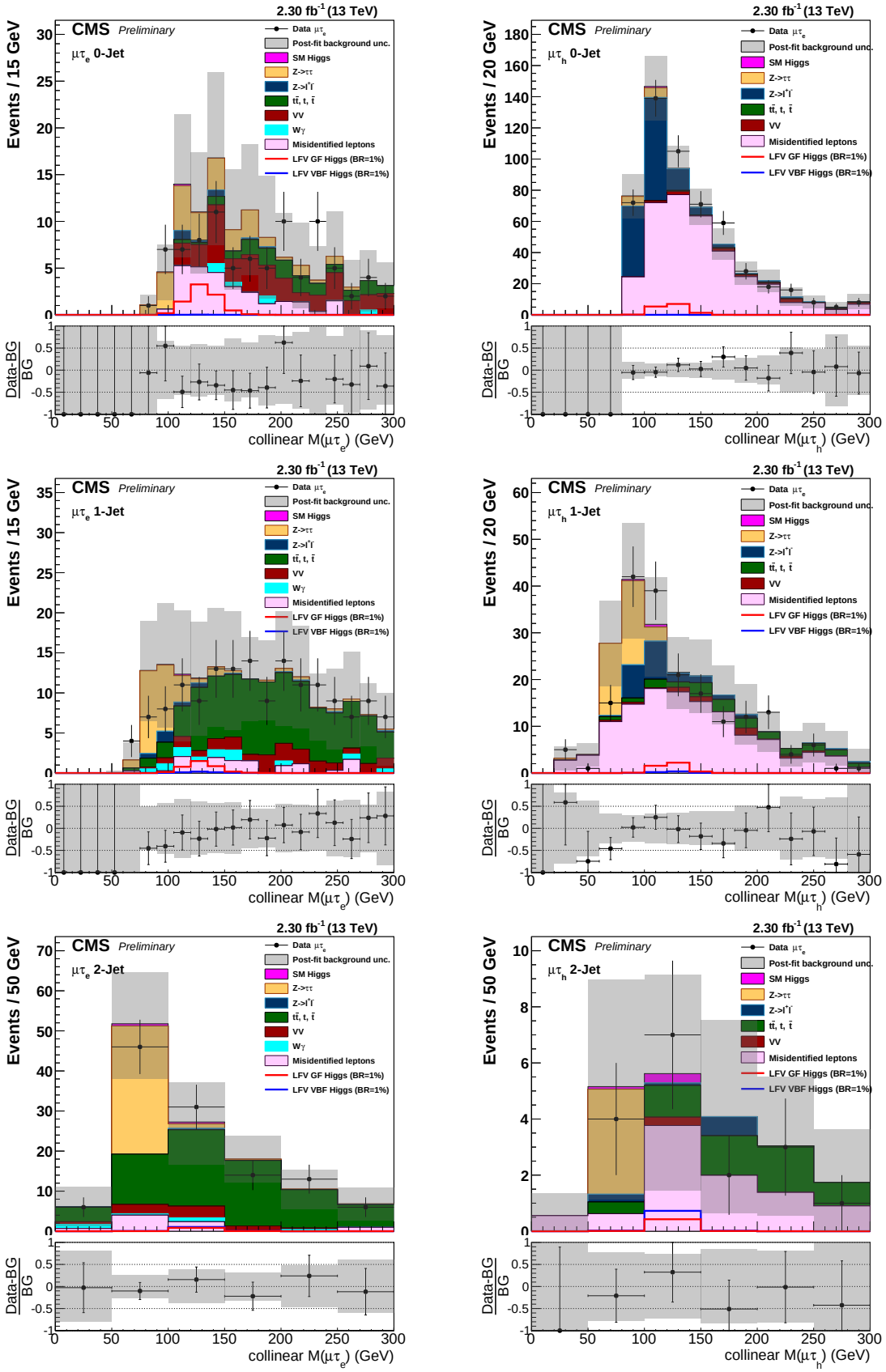

Figure 3: Distribution of the collinear mass M_{col} in the different channels and categories compared to the signal and background estimation. The background is normalized to the best-fit values from the signal plus background fit while the signal is normalized to $\mathcal{B}(\text{H} \to \mu\tau) = 1\%$. The bottom panel in each plot shows the fractional difference between the observed data and the fitted background. Top left: $\text{H} \to \mu\tau_{\text{e}}$ 0-jet; top right: $\text{H} \to \mu\tau_{\text{h}}$ 0-jet; middle left: $\text{H} \to \mu\tau_{\text{e}}$ 1-jet; middle right: $\text{H} \to \mu\tau_{\text{h}}$ 1-jet; bottom left: $\text{H} \to \mu\tau_{\text{e}}$ 2-jet; bottom right $\text{H} \to \mu\tau_{\text{h}}$ 2-jet.

The observed limit on the branching ratio is slightly tighter than the $\mathcal{B}(\mathrm{H} \to \mu\tau) < (1.51 \pm 0.83)\%$ limit obtained using the 19.7 fb^{-1} data sample at 8 TeV analyzed in [23].

Table 5: The observed and expected upper limits and the best-fit branching fractions for different n-jet categories for the $\mathrm{H} \to \mu\tau$ process.

Expected limits				
	0-jet (%)	1-jet (%)	2-jets (%)	Combined (%)
$\mu\tau_h$	<4.17	<4.89	<6.41	<2.98
$\mu\tau_\mathrm{e}$	<2.24	<4.36	<7.31	<1.96
$\mu\tau$	<1.62 %			
Observed limits				
	0-jet (%)	1-jet (%)	2-jets (%)	Combined (%)
$\mu\tau_h$	<4.24	<6.35	<7.71	<3.81
$\mu\tau_\mathrm{e}$	<1.33	<3.04	<8.99	<1.15
$\mu\tau$	<1.20 %			
Best-fit branching fractions				
	0-jet (%)	1-jet (%)	2-jets (%)	Combined (%)
$\mu\tau_h$	$0.12^{+2.02}_{-1.91}$	$1.70^{+2.41}_{-2.52}$	$1.54^{+3.12}_{-2.71}$	$1.12^{+1.45}_{-1.40}$
$\mu\tau_\mathrm{e}$	$-2.11^{+1.30}_{-1.89}$	$-2.18^{+1.99}_{-2.05}$	$2.04^{+2.96}_{-3.31}$	$-1.81^{+1.07}_{-1.32}$
$\mu\tau$	$-0.76^{+0.81}_{-0.84}\%$			

8 Limits on lepton flavour violating couplings

The constraint on $\mathcal{B}(\mathrm{H} \to \mu\tau)$ can be interpreted in terms of LFV Yukawa couplings [4]. The LFV decays $\mathrm{H} \to \mathrm{e}\mu$, $\mathrm{e}\tau$, $\mu\tau$ arise at tree level from the assumed flavour violating Yukawa interactions, $Y_{\ell^\alpha \ell^\beta}$ where ℓ^α, ℓ^β denote the leptons, $\ell^\alpha, \ell^\beta = \mathrm{e}, \mu, \tau$ and $\ell^\alpha \neq \ell^\beta$. The decay width $\Gamma(\mathrm{H} \to \ell^\alpha \ell^\beta)$ in terms of the Yukawa couplings is given by:

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

and the branching fraction by:

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

The SM H decay width is assumed to be $\Gamma_\mathrm{SM} = 4.1\,\mathrm{MeV}$ [61] for $M_\mathrm{H} = 125\,\mathrm{GeV}$. The 95% CL constraint on the Yukawa couplings derived from $\mathcal{B}(\mathrm{H} \to \mu\tau) < 1.20\%$ and the expression for the branching fraction above is:

$$\sqrt{|Y_{\mu\tau}|^2 + |Y_{\tau\mu}|^2} < 3.16 \times 10^{-3}.$$

Figure 5 compares this result to the constraints from previous indirect measurements.

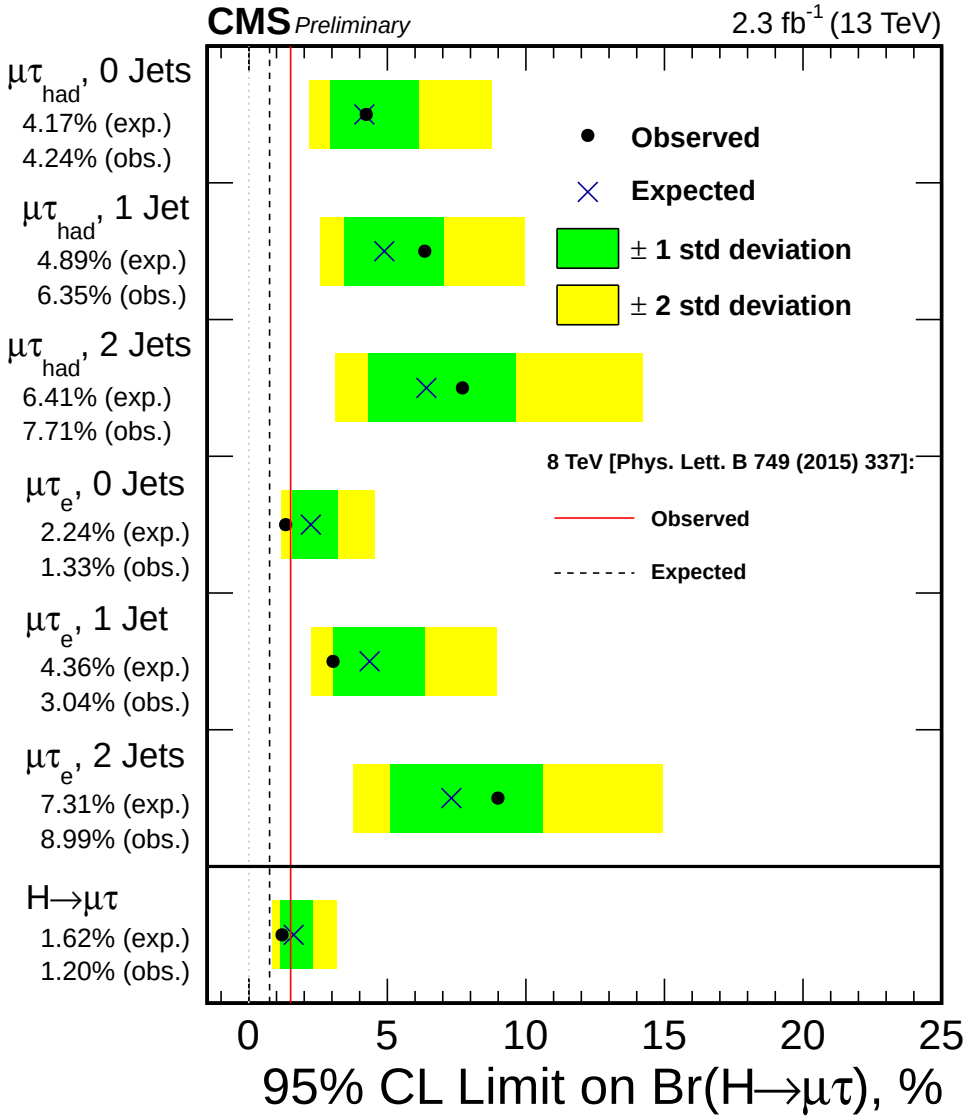

Figure 4: Observed and expected 95% CL upper limits on the $\mathcal{B}(\mathrm{H} \to \mu\tau)$ for each individual category and combined. The solid red and dashed black vertical lines correspond, respectively, to the observed and expected 95% CL upper limits obtained at $\sqrt{s} = 8$ TeV [23].

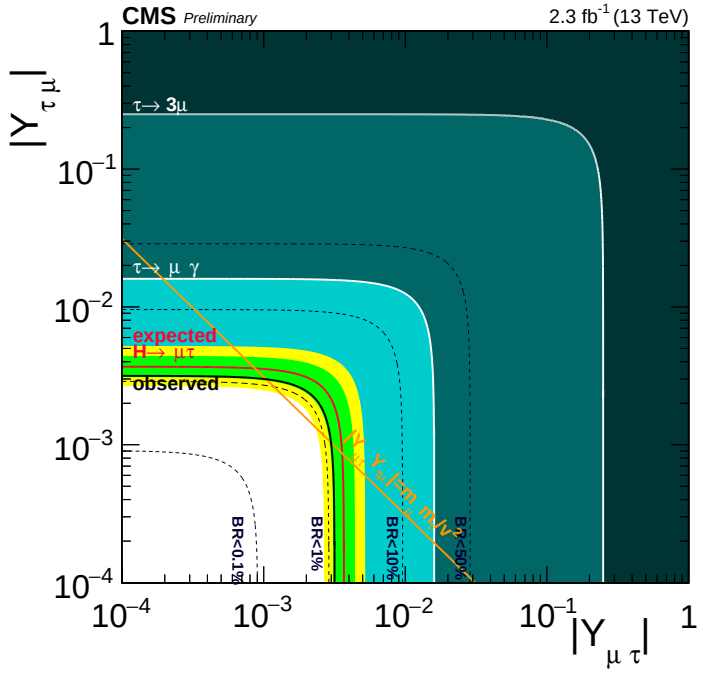

Figure 5: Constraints on the flavour violating Yukawa couplings, $|Y_{\mu\tau}|$ and $|Y_{\tau\mu}|$. The black dashed lines are contours of $\mathcal{B}(\mathrm{H} \to \mu\tau)$ for reference. The expected limit (red solid line) with one standard deviation (green) and two standard deviation (yellow) bands, and observed limit (black solid line) are derived from the limit on $\mathcal{B}(\mathrm{H} \to \mu\tau)$ from the present analysis. The shaded regions are derived constraints from null searches for $\tau \to 3\mu$ (dark green) and $\tau \to \mu\gamma$ (lighter green). The light blue region indicates the additional parameter space excluded by our result. The purple diagonal line is the theoretical naturalness limit $Y_{ij}Y_{ji} \leq m_i m_j / v^2$.

9 Conclusions

A direct search for lepton flavour violating decays of the Higgs boson in the $\mathrm{H} \to \mu\tau$ channel is described. The data sample used in the search was collected in proton-proton collisions at $\sqrt{s} = 13$ TeV with the CMS experiment at the LHC and corresponds to an integrated integrated luminosity of 2.3 fb^{-1}. No excess is observed. The best-fit branching fraction is $\mathcal{B}(\mathrm{H} \to \mu\tau) = -0.76^{+0.81}_{-0.84}\%$ and an upper limit of $\mathcal{B}(\mathrm{H} \to \mu\tau) < 1.20\%$ (1.62% expected) is set at 95% CL.

At $\sqrt{s} = 8$ TeV a small excess was observed, corresponding to 2.4σ, with an analysis based on an integrated luminosity of 19.7 fb^{-1} that yielded an expected 95% CL limit on the branching fraction of 0.75%. More data are needed to make definitive conclusions on the origin of that excess.

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

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

[6] J. L. Diaz-Cruz and 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.

[7] 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.

[8] A. Arhrib, Y. Cheng, and O. C. 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] 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.

[10] 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.

[11] 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.

[12] 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.

[13] 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.

[14] 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.

[15] 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.

[16] 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.

[17] J. 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.

[18] 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.

[19] 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.

[20] 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.

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

[22] 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.

[23] 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.

[24] CMS Collaboration, "Search for lepton-flavour-violating decays of the Higgs boson to $e\tau$ and $e\mu$ at $\sqrt{s} = 8$ TeV", *CMS Physics Analysis Summary* **CMS-PAS-HIG-14-040** (2015).

[25] ATLAS Collaboration, "Search for lepton-flavour-violating $H \to \mu\tau$ decays of the Higgs boson with the ATLAS detector", (2015). arXiv:1508.03372. Submitted to JHEP.

[26] ATLAS Collaboration, "Search for lepton-flavour-violating decays of the Higgs and Z bosons with the ATLAS detector", arXiv:1604.07730.

[27] 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.

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

[29] 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.

[30] K. Olive et al., "Review of Particle Physics", *Chin. Phys. C* **38** (2014) 090001, doi:10.1088/1674-1137/38/9/090001.

[31] 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.

[32] 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.

[33] 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.

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

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

[36] 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.

[37] 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.

[38] 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.

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

[40] 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.

[41] T. Sjöstrand, S. Mrenna, and P. Skands, "A Brief Introduction to PYTHIA 8.1", *Comput. Phys. Commun.* **178** (2007) 852, doi:10.1016/j.cpc.2008.01.036, arXiv:0710.3820.

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

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

[44] 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.

[45] 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).

[46] CMS Collaboration, "Particle flow reconstruction of 0.9 TeV and 2.36 TeV collision events in CMS", *CMS Physics Analysis Summary* **CMS-PAS-PFT-10-001** (2010).

[47] 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).

[48] 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.

[49] 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.

[50] CMS Collaboration, "Performance of CMS muon reconstruction in pp collision events at $\sqrt{s} = 7$ TeV", (2012). arXiv:1206.4071. Submitted to *J. Inst.*

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

[52] 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.

[53] 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.

[54] 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.

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

[56] R. K. Ellis, I. Hinchliffe, M. Soldate, and 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.

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

[58] ATLAS and CMS Collaborations, LHC Higgs Combination Group, "Procedure for the LHC Higgs boson search combination in Summer 2011", Technical Report ATL-PHYS-PUB 2011-11, CMS NOTE 2011/005, 2011.

[59] 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.

[60] 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.

[61] 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.