ETP-KA/2022-15

Karlsruher Institut für Technologie

A Search for Additional Neutral Higgs Bosons in $\tau\tau$ Final States in pp Collisions at $\sqrt{s} = 13\,\text{TeV}$

Zur Erlangung des akademischen Grades eines

DOKTORS DER NATURWISSENSCHAFTEN

(Dr. rer. nat.)

von der KIT-Fakultät für Physik des
Karlsruher Instituts für Technologie (KIT)
angenommene

DISSERTATION

von

M.Sc. Maximilian Burkart
aus Baden-Baden

Tag der mündlichen Prüfung: 11. November 2022

Referent: Priv.-Doz. Dr. Roger Wolf
Korreferent: Prof. Dr. Günter Quast

Institut für Experimentelle Teilchenphysik

Abstract

Many proposed extensions of the standard model of particle physics (SM) predict an extended Higgs sector with respect to the SM. Experimentally, this extended Higgs sector is accessible through searches for additional spin-0 particles often referred to as additional Higgs bosons. Decays of neutral Higgs bosons to pairs of tau leptons provide the prime search channel for such particles in the minimal supersymmetric SM (MSSM) and other two-Higgs-doublet models (2HDM) of Type II since their couplings are enhanced in wide areas of the parameter space. In this thesis, a search for such decays in the form of a search for an additional resonance above the SM prediction in the $\tau\tau$ mass spectrum, targeting the production via gluon fusion and in association with b quarks, is presented. The results are portrayed in two ways. First, model-independent exclusion limits on the product of the production cross section and the branching fraction for the decay into tau leptons of the targeted resonance are derived. These limits translate into model-dependent exclusion contours in a set of MSSM and 2HDM benchmark scenarios, which are inferred from the full $\tau\tau$ spectrum predicted by the investigated scenario.

Contents

CHAPTER 1

Introduction

The discovery of a Higgs boson reported by the Compact Muon Solenoid (CMS) and ATLAS experiments in 2012 [1, 2] represents the experimental verification of the electroweak unification [3–8] in the standard model of particle physics (SM). Since its discovery, the properties of the observed Higgs boson have been studied extensively [9–17] and individual decay modes have been observed [18–22]. All performed measurements have resulted in observations consistent with the predictions from the SM.

Albeit its large success in describing fundamental particles and their interactions, the SM still has known shortcomings: it does not provide a consistent description of gravitation together with the three other forces of nature; it lacks a mechanism to generate the masses of neutrinos which are treated as massless in the SM; it cannot explain the presence of dark matter. It is thus expected that the SM constitutes a low-energy effective theory of a more inclusive theory. This hypothetical theory might manifest itself at the currently accessible energies at hadron colliders either through deviations of precision measurements from SM predictions or through an extended particle spectrum.

Many hypothesized extensions of the SM imply predictions of an extended Higgs sector consisting of multiple physical Higgs bosons. Among these extensions is supersymmetry [23, 24], a global symmetry that transforms fermions to bosons and vice versa. In the minimal supersymmetric SM (MSSM) [25–29], the Higgs sector of the SM is supplemented by an additional complex scalar doublet field resulting in two additional neutral and charged Higgs bosons. The couplings of the additional neutral Higgs bosons to down-type fermions are enhanced with large values of $\tan\beta$, the ratio of the vacuum expectation values of the two doublet fields, while the couplings to up-type fermions and vector bosons are suppressed. This yields a leading role in searches for these additional neutral Higgs bosons to decays into pairs of heavy down-type fermions. Among these, decays into tau lepton pairs provide the best sensitivity as their experimental signature is easier to separate from backgrounds compared to the signature of decays into b quarks.

In the scope of this thesis, significant contributions have been made to a search for additional Higgs bosons in $\tau\tau$ final states published by the CMS Collaboration [30].

Within the publication and this thesis, which build upon a previous analysis of 2016 data published by the CMS experiment [31], the full Large Hadron Collider (LHC) Run 2 data set collected with the CMS experiment has been analyzed. In the analyzed data, corresponding to an integrated luminosity of $138\,\mathrm{fb}^{-1}$, no signal is found. The results are presented in two ways. First, model-independent exclusion limits on the product of the production cross section and branching fraction for the decay into tau leptons of an additional scalar resonance produced via gluon fusion or in association with b quarks above the $\tau\tau$ mass spectrum arising from SM processes are set. These limits translate into model-dependent exclusion contours in a set of MSSM benchmark scenarios which are derived from the full $\tau\tau$ spectrum predicted by the benchmark scenario. In addition to the results given in the publication, in this thesis model-dependent exclusion contours in terms of selected two-Higgs-doublet models (2HDM) [25, 32] are provided.

In the first chapter of the thesis, the underlying theoretical concepts are introduced and properties of the considered extensions to the SM are discussed. The following chapter describes the structure of the CMS experiment together with the algorithms employed to reconstruct the particles emerging from the proton-proton collisions. The description of the analysis selections, categorization and corrections applied to the reconstructed physics objects constitutes the third chapter of the thesis. It is followed by a discussion of the methods used for the statistical inference of the exclusion limits in both interpretations. Lastly, the obtained results are presented and discussed.

CHAPTER 2

Beyond the Standard Model of Particle Physics

2.1 The Standard Model of Particle Physics

The SM describes the fundamental constituents of matter and their interactions in the form of a relativistic quantum field theory. Interactions are modeled via couplings of fields, the quanta of which are identified as particles. Free relativistic fields are characterized by their mass and spin quantum numbers [33]. Fermion fields exhibit half-integer spin quantum numbers and obey the Fermi–Dirac statistic [34]. They represent the fundamental constituents of matter. Boson fields carry integer spin and obey the Bose–Einstein statistic. Two kinds of boson fields appear in the SM, vector fields which are introduced as mediators of the interactions and a single scalar field introduced to generate the masses of the massive vector fields and fermions dynamically in a gauge invariant way. The mechanism of this mass generation is described in more detail in Section 2.2.

The dynamics of a quantum field theory are completely contained in the Lagrangian density or short Lagrangian of the theory. Equations of motion of the fields are obtained through the Euler-Lagrange equations from the Lagrangian. In the Lagrangian of a theory, invariance under certain groups of transformations of the fields, called symmetries, may be present. In the SM, interactions of the fields are introduced into the theory via the gauge principle, which corresponds to the requirement that specific symmetries also hold locally, i. e., a different strength of the transformations may be applied at each space time point. In order to conserve these local symmetries, additional vector fields need to be introduced for each generator of the group of transformations. They connect the transformations in gauge space at different space time points and are thus interpreted as mediators of the interaction.

The gauge group of the SM Lagrangian consists of the direct product of three gauge groups, which is given by $\mathrm{SU}(3)_\mathrm{C} \times \mathrm{SU}(2)_\mathrm{L} \times \mathrm{U}(1)_Y$. The indices of the gauge groups correspond to the charges that the fields may carry and to which the associated vector bosons couple. The first gauge group $\mathrm{SU}(3)_\mathrm{C}$ describes to the strong interaction, whose vector bosons couple to color charges. The fermions carrying the three color charges associated to the gauge group are called quarks. Six different quark flavors exist for each

color charge, namely the up, down, charm, strange, top, and bottom quarks. Eight gluons mediate the strong interactions. They carry color charges themselves and thus can couple directly to each other. These self-interactions lead to two characteristic properties of the strong interaction, namely confinement and asymptotic freedom. Confinement implies that no free colored particles exist but the quarks form color neutral bound states, the hadrons. The strong interaction is said to be asymptotically free as the coupling decreases with increasing energy scale, thus allowing perturbative calculations above energies around a few hundred MeV.

The weak and electromagnetic interactions are governed by the $\mathrm{SU}(2)_\mathrm{L} \times \mathrm{U}(1)_Y$ gauge group. The $\mathrm{SU}(2)_\mathrm{L}$ gauge group acts only on left-handed fields, while right-handed fields transform as singlets under SU(2) rotations. The charge assigned to the $\mathrm{SU}(2)_\mathrm{L}$ is the weak isospin I, where the upper and lower components of the doublet take on values of $I_3 = 1/2$ and $I_3 = -1/2$, respectively. The hypercharge Y, the fields carry, differs between the left and right-handed fields.

The left-handed quark fields are grouped into three generations of doublet fields. The up and down quark (u and d), the charm and strange quark (c and s), and the top and bottom quark (t and b) fields form one doublet field each. The down, strange, and bottom quark fields constitute the lower components of the doublets and are thus referred to as down-type quarks, whereas the other quark flavors are consequently referred to as up-type quarks. In summary, the quark content of the SM consists of the following left and right-handed fields

$$\begin{pmatrix} u_\mathrm{L} \\ d_\mathrm{L} \end{pmatrix}, \quad \begin{pmatrix} c_\mathrm{L} \\ s_\mathrm{L} \end{pmatrix}, \quad \begin{pmatrix} t_\mathrm{L} \\ b_\mathrm{L} \end{pmatrix}, \quad u_\mathrm{R},\ d_\mathrm{R},\ c_\mathrm{R},\ s_\mathrm{R},\ t_\mathrm{R},\ b_\mathrm{R}\,, \tag{2.1}$$

where the handedness of the fields is indicated by the indices L and R, respectively.

The left-handed quark doublet fields are accompanied by further three generations of left-handed doublet fields, which are neutral under the strong interaction. These fields are referred to as lepton fields and comprise the electrically charged leptons in the lower component of the doublet and their corresponding neutrinos ν_i in the upper component. The electron (e), muon (μ), and tau lepton (τ) fields constitute the three electrically charged leptons. In addition to the left-handed doublet fields, a right-handed field is introduced for each of them, whereas no right-handed neutrino field exists in the SM. The lepton content of the SM thus consists of the fields

$$\begin{pmatrix} \nu_\mathrm{e,L} \\ e_\mathrm{L} \end{pmatrix}, \quad \begin{pmatrix} \nu_{\mu,\mathrm{L}} \\ \mu_\mathrm{L} \end{pmatrix}, \quad \begin{pmatrix} \nu_{\tau,\mathrm{L}} \\ \tau_\mathrm{L} \end{pmatrix}, \quad e_\mathrm{R},\ \mu_\mathrm{R},\ \tau_\mathrm{R}\,. \tag{2.2}$$

The bosonic vector fields associated to the $\mathrm{SU}(2)_\mathrm{L} \times \mathrm{U}(1)_Y$ gauge group are the three fields W^a_μ, where a runs from 1 to 3, and the field B_μ respectively [3]. These fields combine to

the physical boson fields of the electroweak interactions, which are the massive Z and W bosons and the massless photon. The charged W boson fields are formed from orthogonal linear combinations of the W_μ^1 and W_μ^2 fields given by

$$W_\mu^\pm = \frac{1}{\sqrt{2}} \left(W_\mu^1 \mp \mathrm{i} W_\mu^2 \right) . \tag{2.3}$$

They couple only to the left-handed fields and induce weak isospin transitions between the components of the doublets. The Z boson (Z_μ) and photon (A_μ) fields are orthogonal linear combinations of the W_μ^3 and B_μ fields coupling to the third component of the weak isospin I_3 and the weak hypercharge. They are constructed such that photons couple to the electric charge of the fields. The physical Z boson and photon fields are given by

$$\begin{pmatrix} Z_\mu \\ A_\mu \end{pmatrix} = \begin{pmatrix} \cos\theta_\mathrm{w} & -\sin\theta_\mathrm{w} \\ \sin\theta_\mathrm{w} & \cos\theta_\mathrm{w} \end{pmatrix} \begin{pmatrix} W_\mu^3 \\ B_\mu \end{pmatrix} , \tag{2.4}$$

where the electroweak mixing angle or Weinberg angle θ_w has been introduced. As the photon field consists of the W_μ^3 and B_μ fields and couples to the electric charge Q, the electric charge can be expressed in terms of the I_3 and Y quantum numbers via the Gell–Mann–Nishijima formula

$$Q = I_3 + \frac{1}{2} Y . \tag{2.5}$$

Quarks carry fractional electric charges amounting to 2/3 for up-type quarks and $-1/3$ for down-type quarks, whereas the charged leptons carry an electric charge of -1. An illustrative summary of the particles described by the SM is given in Figure 2.1.

In addition to the global continuous symmetries, which have been promoted to local symmetries via the gauge principle, Lagrangians may also exhibit additional global discrete symmetries. Among these discrete symmetries are the charge conjugation C, the parity transformation P and the time inversion T. In the weak interactions, the charge conjugation and parity transformation symmetries are both violated maximally as the W bosons couple only to left-handed fermions and right-handed antifermions. The combination of both symmetries, the CP symmetry, however remains a symmetry of the Lagrangian and the weak interactions are thus CP-conserving. The couplings of the complex scalar Higgs doublet field to the fermions represent the only source of CP-violation in the SM [35].

For a more detailed introduction to quantum field theory, the gauge principle and the Brout–Englert–Higgs mechanism, described in the following, the reader is referred to the standard textbooks [36, 37] or general reviews like [35].

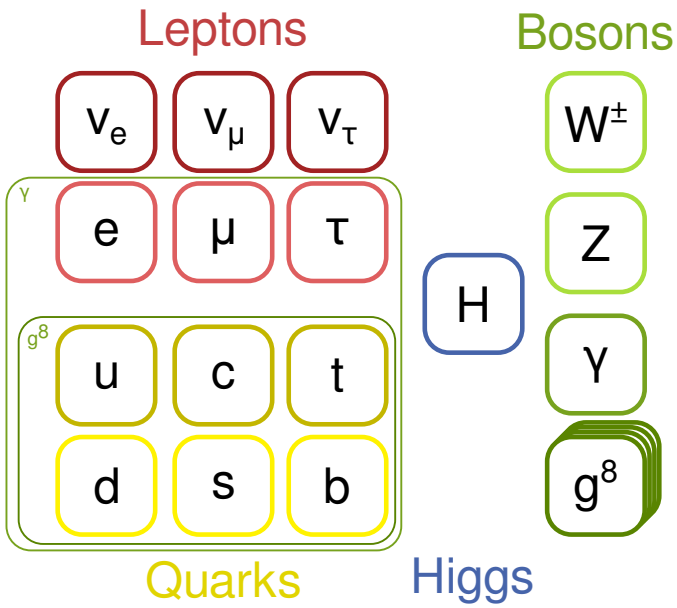

Figure 2.1: Particle content of the SM. Quarks and leptons are shown on the left, whereas bosons are indicated on the right and the physical Higgs boson is depicted in the center. The groupings of the quarks and leptons indicate their association to the left-handed doublets. Down-type fermions are shown in lighter colors than up-type fermions. The green boxes around the fermions indicate to which bosons they couple. The W and Z bosons couple to all shown fermions. The symbol γ denotes the photon.

2.2 The Brout–Englert–Higgs Mechanism

The short range of the weak interactions and observations of the mediating W and Z bosons suggest that these vector bosons are massive. However, mass terms of gauge vector bosons of the form

$$\mathcal{L}_{\text{mass}} = -m_G^2 G_\mu G^\mu \,, \tag{2.6}$$

where G_μ is a generic vector field, are not invariant under transformations of the gauge group the vector boson corresponds to. The Brout–Englert–Higgs mechanism [5–8] allows the dynamic creation of mass terms of this form through spontaneous symmetry breaking (SSB). In spontaneously broken theories the ground state of the theory, i. e., the vacuum, does not respect the symmetry of the theory. In the case of the Brout–Englert–Higgs mechanism, the electroweak symmetry of the $\mathrm{SU(2)_L} \times \mathrm{U(1)}_Y$ gauge group is broken down to the $\mathrm{U(1)}_Q$ symmetry of quantum electrodynamics (QED) through the introduction of an additional complex scalar $\mathrm{SU(2)_L}$ doublet field

$$\Phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} \tag{2.7}$$

with hypercharge $Y = 1$, which acquires a non-zero vacuum expectation value (VEV).

The contribution of this scalar doublet field to the Lagrangian is given by

$$\mathcal{L}_{\text{Higgs}} = |D_\mu \Phi|^2 - V(\Phi^\dagger \Phi) \,, \tag{2.8}$$

where D_μ represents the covariant derivative of the electroweak gauge group $\mathrm{SU(2)_L} \times \mathrm{U(1)}_Y$ defined as

$$D_\mu = \partial_\mu - \mathrm{i}g W_\mu^a \tau^a - \mathrm{i}g' \frac{Y}{2} B_\mu \,. \tag{2.9}$$

Here, g and g' are dimensionless coupling constants and $\tau^a = \sigma^a/2$, where σ^a are the Pauli matrices, correspond to the three generators of the SU(2) symmetry group. The parameter a takes values between 1 and 3 and the sum over all a's is implied in Equation 2.9. The potential of the scalar doublet field of lowest order that allows SSB reads as

$$V(\Phi^\dagger \Phi) = -\mu^2 \Phi^\dagger \Phi + \frac{\lambda}{2} (\Phi^\dagger \Phi)^2 \,, \tag{2.10}$$

where μ would take the role of a mass parameter if μ^2 was negative and λ is a dimensionless coupling constant. Minimizing the potential with respect to the field Φ yields the condition

$$\left\langle \Phi^\dagger \Phi \right\rangle = \frac{\mu^2}{\lambda} =: \frac{v^2}{2} \tag{2.11}$$

for the VEV of Φ. The neutral component of the scalar field Φ takes on this VEV because the acquisition of a VEV in the charged component would break the charge symmetry of

the vacuum. Expanding the scalar field Φ around the VEV and switching to the unitary gauge, in which three degrees of freedom of the scalar doublet field are absorbed in a gauge transformation, the scalar field is given by

$$\Phi = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + h \end{pmatrix} . \tag{2.12}$$

Here, an additional real scalar field h has been introduced as the remaining degree of freedom.

Considering only terms from $|D_\mu \Phi|^2$, which do not contain derivatives and the scalar doublet field after it has acquired its VEV in the unitary gauge, one obtains

$$\begin{aligned} \mathcal{L}_{\text{mass}} &= -\frac{1}{2}\frac{v^2}{4} \left(g^2 (W_\mu^1)^2 + g^2 (W_\mu^2)^2 + (g W_\mu^3 - g' B_\mu)^2 \right) \left(1 + \frac{h}{v} \right)^2 \\ &= -\frac{1}{2}\frac{v^2}{4} \left(2 g^2 W_\mu^+ W^{\mu,-} + (g^2 + g'^2) Z_\mu Z^\mu \right) \left(1 + \frac{h}{v} \right)^2 , \end{aligned} \tag{2.13}$$

where the physical fields $W_\mu^\pm$ and Z_μ have been introduced in the second step. The masses of W and Z bosons corresponding to these fields can then be read directly from the Lagrangian in Equation 2.13 and are given by

$$m_W = g \frac{v}{2} , \qquad m_Z = \sqrt{g^2 + g'^2} \frac{v}{2} . \tag{2.14}$$

Thus, the masses of the W and Z bosons are related via

$$m_W = m_Z \cos \theta_\text{w} , \tag{2.15}$$

where the Weinberg angle θ_w is identified from the definition of the Z boson and photon fields in terms of the couplings g and g'

$$Z_\mu = \frac{1}{\sqrt{g^2 + g'^2}} \left(g W_\mu^3 - g' B_\mu \right) , \tag{2.16}$$

$$A_\mu = \frac{1}{\sqrt{g^2 + g'^2}} \left(g' W_\mu^3 + g B_\mu \right) \tag{2.17}$$

as $\cos \theta_\text{w} = \frac{g}{\sqrt{g^2 + g'^2}}$. The photon thus remains massless and is identified as the gauge boson associated to the resulting $\text{U}(1)_Q$ symmetry of the Lagrangian. As the subscript Q indicates, the conserved quantity is the electric charge which is given in terms of I_3 and Y in Equation 2.5.

Three of the four degrees of freedom of the complex scalar doublet field have been absorbed to give masses to the three massive vector bosons while one physical field, the quanta of which is referred to as the Higgs boson, remains in the theory. In addition to the masses of the weak vector bosons, trilinear and quartic couplings of the Higgs boson

to the massive vector bosons with a coupling strength proportional to the squared mass of the vector bosons are introduced.

Fermion mass terms of the form

$$\mathcal{L}_{\text{mass}} = -m\bar{\psi}\psi\,, \tag{2.18}$$

where ψ denotes a general fermion field, are also not gauge invariant under transformations of the $\mathrm{SU}(2)_\mathrm{L} \times \mathrm{U}(1)_Y$ electroweak gauge group as they couple left to right-handed fields, which transform differently under the gauge transformations. However, the scalar doublet field introduced to create the masses of the massive vector bosons can also be used to give masses to fermions when it acquires its VEV [4]. In the following, the creation of the masses of down-type fermions is illustrated using the electron as example but the mechanism translates directly to other down-type fermions. To create the masses of fermions in a gauge invariant way Yukawa couplings to the scalar doublet field are introduced as

$$\mathcal{L}_{\text{Yukawa}} = -\lambda_\mathrm{e}(\bar{E}_\mathrm{L}\Phi e_\mathrm{R} + \bar{e}_\mathrm{R}\Phi^\dagger E_\mathrm{L})\,, \tag{2.19}$$

where λ_e is a dimensionless coupling constant, $\bar{E}_\mathrm{L} = \begin{pmatrix}\bar{\nu}_{e,\mathrm{L}} & \bar{e}_\mathrm{L}\end{pmatrix}$ the doublet containing the left-handed electron neutrino and electron fields, and e_R the right-handed electron field. After the scalar doublet acquires its VEV, the Yukawa coupling in the unitary gauge is given by

$$\begin{aligned}\mathcal{L}_{\text{Yukawa}} &= -\frac{\lambda_\mathrm{e} v}{\sqrt{2}}(1+\frac{h}{v})(\bar{e}_\mathrm{L}e_\mathrm{R} + \bar{e}_\mathrm{R}e_\mathrm{L}) \\ &= -m_\mathrm{e}\bar{e}e - \frac{m_\mathrm{e}}{v}\bar{e}eh\,,\end{aligned} \tag{2.20}$$

where the electron field $e = e_\mathrm{L} + e_\mathrm{R}$ has been introduced. The first term in Equation 2.20 corresponds to a mass term for the electron with the mass given as $m_\mathrm{e} = \frac{\lambda_\mathrm{e} v}{\sqrt{2}}$ and the second term to a coupling of the electron to the real scalar field h proportional to the mass of the electron.

To generate the masses of up-type quarks, in principle, a scalar doublet field with different quantum numbers would be necessary to introduce gauge invariant Yukawa terms. In the SM, however the use of a transformed field of the scalar doublet field is sufficient. This field is obtained via

$$\tilde{\Phi} = \mathrm{i}\sigma^2\Phi^*\,. \tag{2.21}$$

The Yukawa couplings generating the quark masses then read as

$$\mathcal{L}_{\text{Yukawa}} = -\lambda_\mathrm{d}\bar{Q}_\mathrm{L}\Phi d_\mathrm{R} - \lambda_\mathrm{u}\bar{Q}_\mathrm{L}\tilde{\Phi}u_\mathrm{R} + \text{h.c.}\,, \tag{2.22}$$

where h.c. indicates the hermitian conjugate of the previous terms, $\bar{Q}_\mathrm{L}$ is the left-handed doublet field containing up- and down-type quark fields, and u_R and d_R correspond to the right-handed singlet fields.

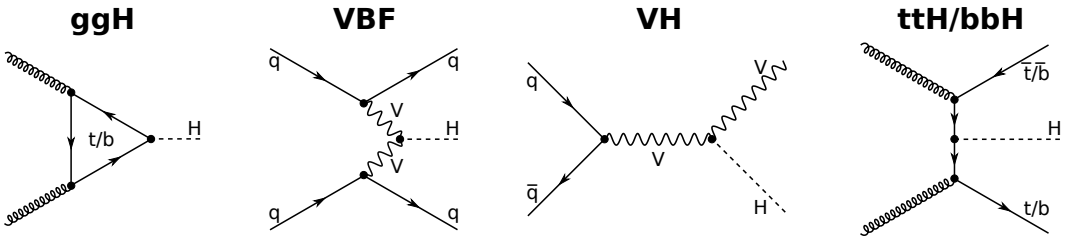

Figure 2.2: Feynman diagrams of the major Higgs boson production mechanisms at hadron colliders. The cross section of the shown production modes decreases from left to right. The production mode with the largest cross section is gluon fusion (ggH), followed by the vector boson fusion (VBF), and vector boson associated (VH) production modes. The production mode with the lowest cross section of the four depicted processes in the SM is the production in association with heavy quarks (ttH/bbH). In the diagrams, straight lines represent fermions, curly lines gluons, wiggly lines vector bosons and dashed lines scalar bosons. The symbols V and q indicate generic massive vector bosons and light flavor quarks, respectively.

Higgs Boson Phenomenology at Hadron Colliders

From the couplings of the Higgs boson to other particles described in the previous paragraphs, multiple possible production modes of Higgs bosons at hadron colliders arise. The possible initial states of all production mechanisms consist of the constituents of the colliding protons, the quarks and gluons. As the strength of the Higgs boson coupling to particles increases with the mass of the particle, all major production modes proceed through the production of heavy particles from the light quarks and massless gluons in the initial state. Example Feynman diagrams of the major production modes of Higgs bosons at hadron colliders are shown in Figure 2.2.

The production mode of Higgs bosons with the largest cross section in the SM is the production via gluon fusion [38]. It proceeds via the annihilation of initial state gluons in a quark loop containing mostly top or bottom quarks. The heavy quarks running in the loop then produce the Higgs boson through the Higgs fermion coupling. The theoretical cross section prediction for Higgs boson production via gluon fusion in the SM amounts to 48^{+3}_{-4} pb at next-to-next-to-next-to leading order (N3LO) precision in the strong coupling constant α_s for a mass of 125.4 GeV [39]. The cross section is dominated by the contribution involving top quarks running in the loop and the total cross section is reduced from the top quark-only contribution through the interference of the top and bottom quark loops [38, 40].

Vector boson fusion (VBF) represents the second most important production mode of SM Higgs bosons. It proceeds via radiation of two massive vector bosons from the incoming quarks, which in turn merge to produce the Higgs boson. The final state of Higgs boson production via VBF thus consists of the Higgs boson and two hadronic jets originating from the initial state quarks [41]. The VBF production mode exhibits a characteristic signature as the two jets originating from the initial state quarks are produced in forward

and backward direction and feature a large separation in pseudorapidity and a high invariant mass. Leptonic decay products of the Higgs boson are isolated from further hadronic activity and more central in the detector. The predicted cross section for the production of the SM Higgs boson via VBF amounts to $3.77 \pm 0.08\,\text{pb}$ at next-to-next-to leading order (NNLO) precision in α_s [39].

The third major production mode of SM Higgs bosons is the production in association with a massive vector boson, also called Higgsstrahlung. It proceeds through the production of an off-shell massive vector boson from the initial state quarks that goes on-shell by radiating off an SM Higgs boson. Leptonic decays of the additional massive vector boson in the final state provide a clean signature of the process. The predicted production cross section amounts to $1.35 \pm 0.03\,\text{pb}$ for the production in association with a W boson and $0.88^{+0.04}_{-0.03}\,\text{pb}$ for the production in association with a Z boson [39]. Both predictions are calculated at NNLO precision in α_s. The given cross section for Higgs boson production in association with Z bosons also includes gluon induced processes, which produce the final state through box diagrams involving top and bottom quarks.

The last major production mode of SM Higgs bosons is the production in association with heavy quarks. The initial state of these processes consists both of quark antiquark pairs and gluons. In the case of initial states involving quarks, the quark antiquark pair annihilates to a gluon that decays again into the heavy quark pair. One of the quarks then radiates off the SM Higgs boson. In the case of initial state gluons, the heavy quark exchanged in the t or u channel radiates off the Higgs boson. The final state in both cases involves the produced SM Higgs boson and two heavy quarks. The cross section prediction for SM Higgs boson production in association with top and bottom quarks amounts to $0.50^{+0.03}_{-0.05}\,\text{pb}$ and $0.48^{+0.10}_{-0.12}$ at NLO and up to NNLO precision in α_s, respectively [39].

Example Feynman diagrams of the different decay modes of the SM Higgs boson are displayed in Figure 2.3. In the notation of Higgs boson decays to charged particles, i. e., W bosons and fermion pairs, in Figure 2.3 and the following the individual charges and particle antiparticle states are not given explicitly but implied. As in Higgs boson production, Higgs boson decays to massive particles are preferred due to the increased coupling. A further factor influencing the branching fractions for the different decay modes is their kinematic accessibility. The branching fractions of the different decay modes of the SM Higgs boson for a mass of 125.4 GeV are displayed in Figure 2.4.

The decay into pairs of b quarks takes a leading role of all decay modes of the SM Higgs boson since b quarks are the heaviest particles for which a decay into real particles is allowed. Decays of the SM Higgs boson to pairs of massive vector bosons are suppressed as at least one of the bosons has to be virtual in the decay. Nonetheless, the $\text{H} \to \text{WW}$ decay, which represents the leading decay mode into vector bosons, exhibits the second largest branching fraction of all decay modes. Furthermore, Higgs boson decays to pairs of massless vector bosons are allowed via loops involving massive particles carrying the

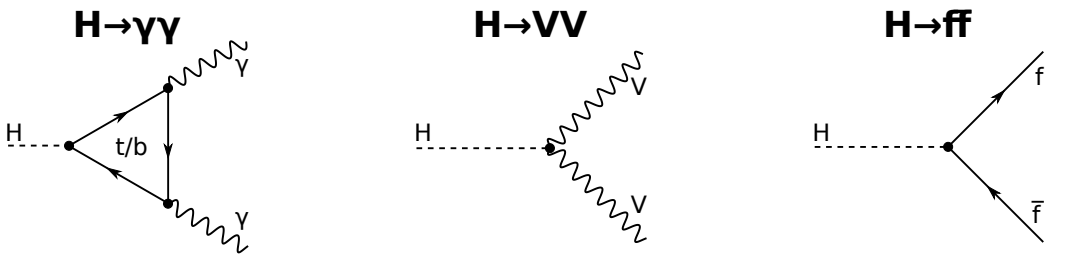

Figure 2.3: Example Feynman diagrams of Higgs boson decays. On the left, an example Feynman diagram of Higgs boson decays to massless vector bosons via loops involving heavy particles is shown. The exemplary shown process is the Higgs boson decay to a pair of photons. In the central diagram, a Feynman diagram of Higgs boson decays to pairs of massive vector bosons is shown, while on the right-hand side a Feynman diagram of the decay of Higgs bosons into fermion pairs is depicted.

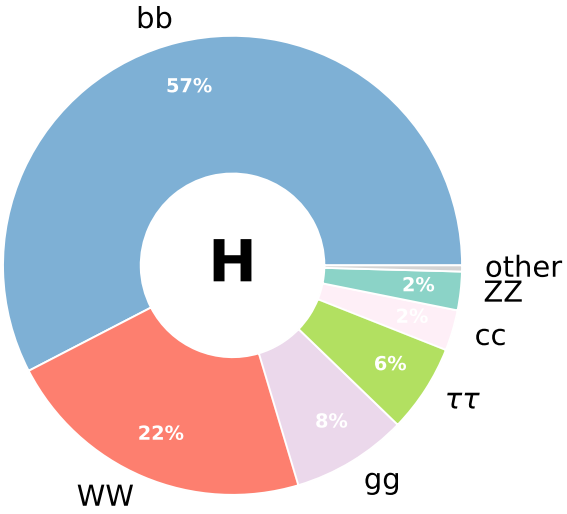

Figure 2.4: Branching fractions of the decay modes of the SM Higgs boson. Decay modes classified in 'other' include the decay to a pair of photons, which represents one of the discovery channels of the SM Higgs boson. This decay mode obtains its large experimental significance through the clean signature with few expected background events. The branching fractions are obtained from [39] for the SM Higgs boson with a mass of 125.4 GeV.

charge to which the bosons couple. As such, decays into pairs of gluons are possible through the same mechanism as in the case of the gluon fusion process. Decays into photon pairs mostly involve top quarks, bottom quarks, or W bosons in the loop.

The experimental accessibility of the different signatures of the decay modes of the SM Higgs boson influences their relevance for its discovery and investigation with respect to their branching fractions. So has the decay mode to a pair of b quarks been the last of the experimentally established decay modes so far [21, 42] as it suffers from large backgrounds from the production of multiple hadronic jets via the strong interaction (QCD multijet events). The same holds for the decay to gluons, which is even harder to observe and will likely not be observed at the LHC albeit exhibiting the third largest branching fraction. Furthermore, the decay into a pair of photons has been one of the discovery channels of the SM Higgs boson with a branching fraction of 0.2 % [39].

2.3 Beyond the Standard Model of Particle Physics

Even though the SM yields a concise description of the known particles and their interactions and so far no significant deviations from its predictions have been observed, some open questions remain [27]:

- The SM is only able to describe three of the four known forces of nature, the strong, the weak, and the electromagnetic force, in the form of gauge theories. The fourth force, gravitation, is not included in the SM.
- The Higgs Yukawa couplings giving mass to the charged leptons and quarks as introduced in Section 2.2 do not generate mass terms for neutrinos. In order to explain the observed oscillations between different neutrino flavors [43], however, at least two of the neutrino generations have to be massive. An inclusion of neutrino masses in the generation of fermion masses via the Higgs field would require the introduction of additional right-handed neutrino fields, which are neutral under all interactions of the SM.
- The SM contains three different interactions with different coupling strengths at the weak interaction scale around 100 GeV [26]. These coupling strengths are not constant but change with the energy scale at which they are evaluated and come closer for larger energies. In the SM, however, there is no single energy scale at which all three running coupling strengths arrive at a common value [35]. Such a unification of the gauge couplings would be theoretically appealing as it allows the explanation of the complicated gauge group of the SM as low-mass effective theory of a unified theory containing only a single gauge group.
- The SM does not provide an explanation for the hierarchy or naturalness problem [26, 27, 44], which describe the observation that the measured mass of the Higgs boson of 125.38 ± 0.14 GeV [14] is 17 orders of magnitude smaller than the Planck scale, which is the energy scale at which the known gauge coupling strengths

become comparable to the strength of gravitation [36]. The Higgs boson mass is subject to radiative corrections, which are quadratically divergent in perturbation theory. As there currently is no indication that the SM may not be valid up to the Planck scale, it naturally provides the cut-off scale for these quadratic divergences. Thus, the bare mass parameter of the Higgs boson has to be carefully adjusted in order to nearly cancel the quadratic divergences and arrive at a physical Higgs boson mass 17 orders of magnitude smaller than the radiative corrections at the Planck scale.

- Cosmological observations suggest the existence of dark matter [45], which accounts for 84.4 % of the matter in the universe [35]. In order to have escaped direct detection so far, it is assumed to only interact weakly with the known particles of the SM. The SM provides no candidate for a large fraction of the dark matter, as neutrinos are not present with a high enough density in the universe to account for all dark matter [45] and make up only for a fraction between 0.5 and 1.6 % [35].

2.3.1 Two-Higgs-Doublet Models

In the SM, the introduction of a single scalar doublet field is sufficient to generate the masses of the W and Z bosons and both up- and down-type fermions. General investigations of extended Higgs sectors, which are obtained through the introduction of additional scalar fields, provide a possibility to search for physics beyond the SM. These investigations can be mostly centered around additional scalar singlet or doublet fields as higher dimensional representations of the gauge group generally would alter the relation between the masses of the W and Z bosons in Equation 2.15 [46], which is in agreement with observation. Two-Higgs-doublet models (2HDM) [25, 32] provide a general extension to the Higgs sector with interesting phenomenological consequences. The following discussion of properties of 2HDMs is based on the discussion given in [46].

With the introduction of an additional complex scalar doublet the number of degrees of freedom in the Higgs sector increases to eight. Three of these degrees of freedom are still needed to take the role of longitudinal polarization of the massive vector bosons. This leaves five physical degrees of freedom in the theory which can be identified as five Higgs bosons with different CP properties and charges. The Higgs bosons emerging from 2HDMs with CP-conserving potential and VEVs are two CP-even Higgs bosons, h and H, where h refers to the lighter of the two, a CP-odd Higgs boson, A, and two charged Higgs bosons, $H^{\pm}$. To provide a model compatible with the observation of a Higgs boson with SM-like properties, one of the CP-even Higgs bosons is usually identified with the observed Higgs boson. To ease the discussions of the properties of the Higgs bosons in the following and throughout the rest of the thesis, the Higgs boson predicted by the SM and the observed Higgs boson are referred to as h_{SM} and H(125), respectively.

The general form of the potential including the scalar doublet fields in a 2HDM is given

by

$$\begin{aligned} V = {} & m_{11}^2 \Phi_1^\dagger \Phi_1 + m_{22}^2 \Phi_2^\dagger \Phi_2 - \left[m_{12}^2 \Phi_1^\dagger \Phi_2 + \text{h.c.} \right] \\ & + \frac{1}{2} \lambda_1 (\Phi_1^\dagger \Phi_1)^2 + \frac{1}{2} \lambda_2 (\Phi_2^\dagger \Phi_2)^2 \\ & + \lambda_3 (\Phi_1^\dagger \Phi_1)(\Phi_2^\dagger \Phi_2) + \lambda_4 (\Phi_1^\dagger \Phi_2)(\Phi_2^\dagger \Phi_1) \\ & + \left\{ \frac{1}{2} \lambda_5 (\Phi_1^\dagger \Phi_2)^2 + \left[\lambda_6 (\Phi_1^\dagger \Phi_1) + \lambda_7 (\Phi_2^\dagger \Phi_2) \right] (\Phi_1^\dagger \Phi_2) + \text{h.c.} \right\} , \end{aligned} \tag{2.23}$$

where m_{11}^2 and m_{22}^2 are free real parameters of dimension mass squared, λ_1, λ_2, λ_3, and λ_4 are dimensionless free real parameters, λ_5, λ_6 and λ_7 are possibly complex dimensionless parameters, and m_{12}^2 is a possibly complex parameter of dimension mass squared [46]. The potential is CP-conserving if a basis in terms of the two scalar doublet fields exists in which m_{12}^2, λ_5, λ_6, and λ_7 are simultaneously real. In the case of 2HDMs, SSB in the electroweak sector is realized when the neutral components of the doublets take on their VEVs. These VEVs are then given by

$$\left\langle \Phi_j^0 \right\rangle = \frac{v_j}{\sqrt{2}} \mathrm{e}^{\mathrm{i}\xi_j} , \tag{2.24}$$

where additionally the relation $v_1^2 + v_2^2 = v^2 = (246\,\text{GeV})^2$ is imposed, the index j takes on values of 1 and 2, and the ξ_j denote arbitrary phases. After SSB, the Higgs sector stays CP-conserving if both VEVs are simultaneously real in a given basis.

The introduction of Yukawa couplings in a general form leads to flavor changing neutral currents (FCNC) mediated by the neutral Higgs bosons, which are not observed in nature, at tree-level. To remove these FCNCs from the theory, an additional discrete symmetry, called $\mathbb{Z}_2$ symmetry, is introduced into the Higgs sector. The requirement of this symmetry has two consequences on the Higgs sector.

Firstly, it reduces the number of free parameters in the potential as it requires the parameters λ_6, λ_7, and m_{12}^2 to be zero. The requirement on the mass parameter m_{12}^2 may be lifted if only a softly broken instead of an exact $\mathbb{Z}_2$ symmetry is imposed. Thus, apart from the parameter v^2, which is fixed through the masses of the W and Z bosons, the CP-conserving 2HDM with softly broken $\mathbb{Z}_2$ symmetry contains seven real free parameters.

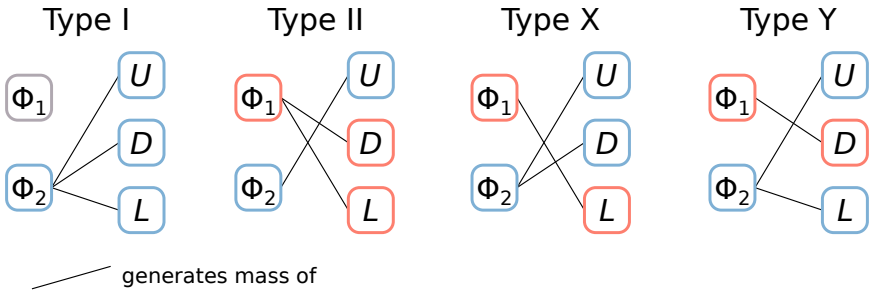

Figure 2.5: Illustration of the definition of the four types of 2HDMs. The symbols Φ_1 and Φ_2 indicate the two Higgs doublet fields, while U, D, and L indicate up- and down-type quarks and charged leptons, respectively. The black lines and the colors indicate which fermion masses are generated by each of the doublet fields.

Secondly, the structure of the Yukawa couplings of the scalar fields to fermions is restricted through the requirement of the $\mathbb{Z}_2$ symmetry to four different types of couplings [47]. The types of 2HDMs are assigned based on the association of the Higgs doublets to the generation of the masses of up-type quarks, down-type quarks and charged leptons:

Type I: All fermion masses are generated through the coupling to the single scalar doublet field Φ_2.

Type II: The masses of the down-type fermions are generated from one scalar doublet field, Φ_1, whereas the masses of the up-type quarks are generated from the other doublet field, Φ_2.

Type X (lepton specific): The lepton masses are generated by a different scalar doublet field than the quark masses. The doublet field giving masses to the leptons is identified as Φ_1.

Type Y (flipped): Down-type quark masses are generated through couplings to a different scalar doublet field with respect to up-type quark and lepton masses. The masses of down-type quarks are generated from the doublet field Φ_1.

An illustration of the definition of the four different types of 2HDMs in terms of the generation of the fermion masses from the Yukawa couplings is displayed in Figure 2.5.

The introduction of a second Higgs doublet and the different coupling structures of the 2HDM types that suppress FCNCs also influence the couplings of quarks and charged leptons to the physical Higgs bosons. Their coupling strengths to the different fermion types with respect to the couplings of h_{SM} for equal masses are given in Table 2.1. They are a function of the parameters α and β, where α is the mixing angle diagonalizing the mass matrix of the CP-even Higgs bosons and β describes the ratio of the VEVs of the two scalar doublet fields as

$$\tan\beta = \frac{\langle \Phi_2^0 \rangle}{\langle \Phi_1^0 \rangle} . \tag{2.25}$$

Table 2.1: Couplings of the interactions of the h, H, and A bosons with quarks and leptons with respect to the coupling of h_{SM} with equal mass for the four different 2HDM types [46]. The notations $\bar{U}U$, $\bar{D}D$, and $\bar{L}L$ refer to the couplings of the Higgs bosons to generic up-type quarks, down-type quarks and charged leptons, respectively. Differences of the couplings of the different 2HDM types with respect to the Type I 2HDM couplings are highlighted in light gray.

	h			H			A		
	$\bar{U}U$	$\bar{D}D$	$\bar{L}L$	$\bar{U}U$	$\bar{D}D$	$\bar{L}L$	$\bar{U}U$	$\bar{D}D$	$\bar{L}L$
Type I	$\frac{\cos\alpha}{\sin\beta}$	$\frac{\cos\alpha}{\sin\beta}$	$\frac{\cos\alpha}{\sin\beta}$	$\frac{\sin\alpha}{\sin\beta}$	$\frac{\sin\alpha}{\sin\beta}$	$\frac{\sin\alpha}{\sin\beta}$	$-\cot\beta$	$\cot\beta$	$\cot\beta$
Type II	$\frac{\cos\alpha}{\sin\beta}$	$-\frac{\sin\alpha}{\cos\beta}$	$-\frac{\sin\alpha}{\cos\beta}$	$\frac{\sin\alpha}{\sin\beta}$	$\frac{\cos\alpha}{\cos\beta}$	$\frac{\cos\alpha}{\cos\beta}$	$-\cot\beta$	$-\tan\beta$	$-\tan\beta$
Type X	$\frac{\cos\alpha}{\sin\beta}$	$\frac{\cos\alpha}{\sin\beta}$	$-\frac{\sin\alpha}{\cos\beta}$	$\frac{\sin\alpha}{\sin\beta}$	$\frac{\sin\alpha}{\sin\beta}$	$\frac{\cos\alpha}{\cos\beta}$	$-\cot\beta$	$\cot\beta$	$-\tan\beta$
Type Y	$\frac{\cos\alpha}{\sin\beta}$	$-\frac{\sin\alpha}{\cos\beta}$	$\frac{\cos\alpha}{\sin\beta}$	$\frac{\sin\alpha}{\sin\beta}$	$\frac{\cos\alpha}{\cos\beta}$	$\frac{\sin\alpha}{\sin\beta}$	$-\cot\beta$	$-\tan\beta$	$\cot\beta$

The couplings of the Higgs bosons to massive vector bosons, g_{VV}, do not depend on the type of the 2HDM. For the neutral CP-even Higgs bosons, h and H, they are given as

$$\frac{g_{hVV}}{g_{h_{SM}VV}} = \sin(\beta-\alpha), \quad \frac{g_{HVV}}{g_{h_{SM}VV}} = \cos(\beta-\alpha) \tag{2.26}$$

and thus add up quadratically to the coupling in the SM, $g_{h_{SM}VV}$. The couplings of the CP-odd Higgs boson A to the massive vector bosons vanish as these couplings are CP-violating.

Since the general CP-conserving 2HDM with softly broken $\mathbb{Z}_2$ symmetry possesses seven free real parameters, its parametric dimensionality is too large to be investigated in a single analysis. Therefore, benchmark scenarios are defined by fixing five of the seven free parameters such that a model with interesting phenomenological properties is obtained. The analysis of these benchmark scenarios is then performed as function of the remaining two free parameters. Each benchmark scenario is investigated with Type I and Type II Yukawa couplings.

For this thesis, the properties of the Higgs sector like the masses of the Higgs bosons and their branching fractions have been calculated using `2HDMC` 1.8.0 [48] given the input parameters of the benchmark scenarios. Furthermore, cross section predictions for the production of the three neutral Higgs bosons in gluon fusion and b quark associated production have been obtained from `SusHi` 1.7.0 [49, 50]. The gluon fusion cross section predictions include analytic predictions at NLO precision in α_s [51], an expansion in the soft limit at NNLO precision in α_s [52], and are supplemented by two loop electroweak corrections [53, 54]. In the case of the h, the predictions additionally include corrections

at N3LO precision in α_s in a series expansion around the production threshold [40, 55, 56]. The cross section of b quark associated production is calculated at NNLO precision in α_s [57]. In the calculations, the running and decoupling of the strong coupling constant α_s and the quark masses is calculated following the analytic formulas given in [58]. The calculation uses the PDF4LHC15 [59] parton distribution function (PDF) set as recommended in [39]. Furthermore, the SM input parameters are set to the values recommended in [39] for the computations.

Three benchmark scenarios are supposed to illustrate the exclusion power of the collected $\tau\tau$ data in the 2HDM parameter space. In all considered benchmark scenarios, two-dimensional scans of the reduced parameter space are performed as function of m_H and $\tan\beta$. Two different basis choices of the Higgs potential in 2HDMs are employed. The first benchmark scenario (Scenario 1) corresponds to the non-alignment scenario described in [46]. It is parameterized in the hybrid basis introduced therein which specifies the Higgs sector in terms of m_h, $\cos(\beta-\alpha)$ and three dimensionless quartic couplings of the Higgs potential, Z_4, Z_5 and Z_7 in addition to the chosen free parameters. In all benchmark scenarios, the parameter m_h is chosen to correspond to the measured mass of H(125). To introduce a small mass hierarchy of

$$m_h < m_H < m_A = m_{H^\pm} \tag{2.27}$$

between the heavy Higgs bosons, the dimensionless couplings are set to $Z_4 = Z_5 = -2$, and $Z_7 = 0$. The non-alignment of the considered scenario is achieved by setting $\cos(\beta-\alpha) \neq 0$, which is realized by a value of $\cos(\beta-\alpha) = 0.01$ for the Type II 2HDM. In the considered Type I 2HDM, the value of $\cos(\beta-\alpha)$ is decreased as function of m_H following the relation $\cos(\beta-\alpha) = 0.1 \cdot (150\,\text{GeV}/m_H)^2$ to prevent exclusions of a large fraction of the parameter space from unitarity, vacuum stability, and perturbativity constraints on the Higgs potential.

Two further benchmark scenarios are introduced to allow a comparison of the obtained exclusion limits with previous [60] (Scenario 2) and recent interpretations [61] (Scenario 3) of data collected with the CMS experiment. Both of these benchmark scenarios are defined in a similar way using the physical basis of the Higgs potential mostly given in terms of the masses of the five Higgs bosons. The fixed parameters of these models are m_h, m_A, $m_{H^\pm}$, $\cos(\beta-\alpha)$, and m_{12}^2, which softly breaks the $\mathbb{Z}_2$ symmetry of the Higgs potential. The two benchmark scenarios differ by the imposed mass hierarchy of the heavy Higgs bosons and the choice of m_{12}^2. In Scenario 2, the mass hierarchy is chosen as $m_A = m_{H^\pm} = m_H + 100\,\text{GeV}$ to open up the possibility for $A \to Zh$ decays [60] while m_H and m_A are chosen to be equal in Scenario 3 [61]. The unitarity, vacuum stability, and perturbativity constraints are not necessarily fulfilled by arbitrary choices of m_H, m_A, and m_{12}^2 [46]. Both models differ also in the choice of m_{12}^2 to guarantee these properties over a large fraction of the parameter space. In Scenario 2, it is chosen to be

$$m_{12}^2 = \max(1-\tan^{-2}\beta, 0) \cdot \frac{1}{2}\sin(2\beta)(m_A^2 + \lambda_5 v^2)\,; \tag{2.28}$$

Table 2.2: Parameters for the 2HDM benchmark scenarios investigated in this thesis. In Scenario 2 and 3, the same parameter set is used for the Type I and Type II couplings of the scenario. The parameter $\cos(\beta-\alpha)$ is abbreviated as $c_{\beta\alpha}$ throughout the table. In all scenarios, the parameters m_{H} and $\tan\beta$ are scanned.

Scenario 1		Scenario 2	Scenario 3
Type I	Type II		
$m_{\mathrm{h}} = 125.38\,\mathrm{GeV}$		$m_{\mathrm{h}} = 125.38\,\mathrm{GeV}$	
$c_{\beta\alpha} = 0.1 \cdot (\frac{150\,\mathrm{GeV}}{m_{\mathrm{H}}})^2$	$c_{\beta\alpha} = 0.01$	$c_{\beta\alpha} = 0.1$	
$Z_4 = -2$		$m_{\mathrm{A}} = m_{\mathrm{H}} + 100\,\mathrm{GeV}$	$m_{\mathrm{A}} = m_{\mathrm{H}}$
$Z_5 = -2$		$m_{\mathrm{H}^\pm} = m_{\mathrm{H}} + 100\,\mathrm{GeV}$	$m_{\mathrm{H}^\pm} = m_{\mathrm{H}}$
$Z_7 = 0$		$m_{12}^2 = 1/2\max(1-\tan^{-2}\beta, 0)$ $\cdot \sin(2\beta)(m_{\mathrm{A}}^2 + \lambda_5 v^2)$	$m_{12}^2 = m_{\mathrm{A}}^2 \frac{\tan\beta}{1+\tan^2\beta}$

in Scenario 3 it is given by

$$m_{12}^2 = m_{\mathrm{A}}^2 \cdot \frac{\tan\beta}{1+\tan^2\beta}\,. \tag{2.29}$$

The parameter choices for the three considered benchmark scenarios are summarized in Table 2.2.

2.3.2 The MSSM Higgs Sector

The introduction of an additional global symmetry that transforms both fermions into bosons and bosons into fermions, supersymmetry (SUSY) [23, 24], is able to provide possible explanations for the last three of the shortcomings of the SM discussed at the beginning of this section. With the introduction of SUSY, the particle content of the theory more than doubles. For each particle in the SM, a superpartner carrying a spin quantum number differing by 1/2 has to be introduced. The scalar sleptons and squarks represent the partners of the leptons and quarks in the SM and fermionic gauginos are introduced as superpartners of the vector bosons. The Higgs sector of the SM has to be adapted to contain an additional doublet field as the generation of the masses of up-type quarks with a single transformed doublet field as in the SM is not invariant under supersymmetric transformations and the fermionic superpartners of a single doublet field would introduce anomalies in the theory [26]. The minimal particle content of supersymmetric models is illustrated in Figure 2.6.

The invariance under exact supersymmetric transformations directly leads to massless scalar bosons since the chiral symmetry requires the fermionic superpartners of the scalar bosons to be massless [26]. Thus, SUSY would render the SM natural and provide an explanation for the hierarchy problem. If realized in nature, however, SUSY must be broken

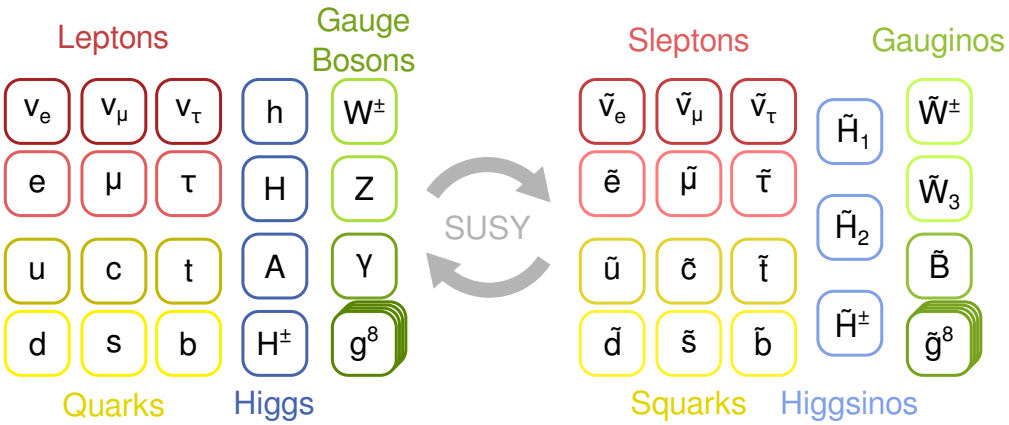

Figure 2.6: Minimal particle content of supersymmetric models. The particles of the SM are shown on the left-hand side, whereas their superpartners, indicated by tildes, are shown on the right-hand side.

as for an exact symmetry the superpartners would have the same mass as their SM partners. So far, no superpartners have been observed, thus they must be more massive than their SM partners. Additionally, a non-vanishing mass in the Higgs sector is required to generate the masses of the massive vector bosons via the Brout–Englert–Higgs mechanism. Furthermore, the additional particles introduced through SUSY can alter the running of the coupling constants in a way such that they arrive at a common value for energy scales shortly below the Planck scale [26]. Thus, SUSY would allow for a unification of the forces.

Lastly, SUSY provides a candidate for weakly interacting massive particle (WIMP) dark matter if an additional conserved quantity is introduced [26]. This quantity is called R-parity and is defined via

$$R_\mathrm{p} = (-1)^{3B+L+2S}\,, \tag{2.30}$$

where B, L, and S refer to the baryon, lepton, and spin quantum numbers respectively. It takes values of $+1$ for SM particles and -1 for their superpartners. Thus, the introduction of R-parity only allows the pairwise creation and annihilation of supersymmetric particles. Therefore, the lightest supersymmetric particle is stable and provides a candidate for WIMP dark matter given that it is electrically and color neutral.

The minimal extension to the SM (MSSM) [25–29] corresponds to the simplest supersymmetric model including the gauge group of the SM [27]. It comprises the minimal particle content as described and illustrated above and the requirement of R-parity conservation. The necessary introduction of an additional scalar doublet field in the MSSM renders its Higgs sector a 2HDM. Invariance under supersymmetric transformations imposes additional requirements on the Higgs sector of the CP-conserving MSSM with respect to the the general CP-conserving 2HDM discussed above [27]. Both consist of the same five physical Higgs bosons, i. e., the two CP-even h and H bosons, the CP-odd A boson,

and the two charged bosons $H^{\pm}$. The additional constraints imposed by SUSY include the form of the Higgs potential, relations between some of the generally unconstrained 2HDM parameters, discussed in Section 2.3.1, and the form of the Yukawa couplings to a Type II 2HDM [62]. At the tree-level, there are only two free parameters describing the properties of the Higgs boson in the MSSM [27], which are usually chosen to correspond to m_A and $\tan\beta$. The masses of the CP-even Higgs bosons at the tree-level are then obtained via [27]

$$m^2_{h,H} = \frac{1}{2}\left[m^2_A + m^2_Z \mp \sqrt{(m^2_A + m^2_Z)^2 - 4m^2_A m^2_Z \cos^2 2\beta}\right] . \tag{2.31}$$

This directly implies that m_h is lower than the mass of the Z boson, whereas m_H is larger than the maximum of m_A and the Z boson mass [27].

Radiative corrections to the Higgs boson masses may raise the lower bound on m_h to values around $m_h \lesssim 140\,\text{GeV}$ [27]. As H(125) has been observed around a mass of 125 GeV, the parameters in most considered models are chosen to raise the lower bound on m_h to masses around 125 GeV for a large fraction of the parameter space. These radiative corrections introduce further MSSM parameters in the predictions of the Higgs boson masses, branching fractions and production cross sections. This renders a full analysis of the parameter space unfeasible similar to the 2HDM case and the experimental analysis is performed in a set of benchmark scenarios [63–66]. The benchmark scenarios have been defined to illustrate certain phenomenological properties of the MSSM by fixing the SUSY parameters, which enter the radiative corrections to specific values [63]. The predictions in the benchmark scenarios are provided in a two-dimensional grid of the free parameters at tree-level usually chosen as m_A and $\tan\beta$, by the LHC Higgs Working Group [67]. The model-dependent exclusion contours presented in this thesis are obtained using the predictions from the M_h^{125} [64] and $M_{h,\text{EFT}}^{125}$ [65] benchmark scenarios.

The M_h^{125} benchmark scenario is characterized by heavy superpartners of the SM particles with masses above 1 TeV [63]. The superpartners therefore play only a minor role in the decays of the heavy neutral Higgs bosons and alter the couplings of h only through loop-induced corrections. This scenario thus corresponds to a Type II 2HDM with couplings of the Higgs bosons constrained by the requirements of SUSY. For low values of $\tan\beta$, the tree-level prediction of m_h decreases and larger radiative corrections are necessary to arrive at predictions compatible with 125 GeV [68]. The radiative corrections are dominated by logarithms of the ratio of the stop quark mass scale over the top quark mass for large values of the stop quark mass parameters at the one-loop level [68]. As a fixed value of 1.5 TeV is used for the stop mass parameter in the M_h^{125} benchmark scenario, the radiative corrections can not be adjusted dynamically and values of m_h incompatible with 125 GeV are obtained for values of $\tan\beta \lesssim 6$.

In the $M_{h,\text{EFT}}^{125}$ scenario, this issue is addressed by setting all squark mass parameters to a common value referred to as the SUSY scale M_S and adjusting this value as function of m_A and $\tan\beta$ to yield a value of m_h around 125 GeV [63, 65]. To obtain these

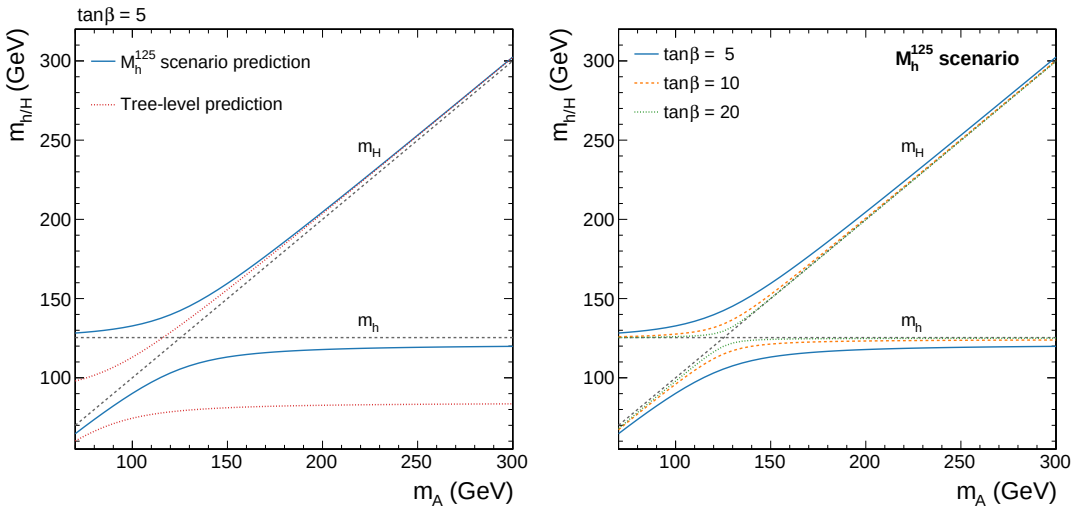

Figure 2.7: Predictions of the masses of the CP-even Higgs bosons in the MSSM. On the left-hand side, a comparison of the tree-level predictions of the Higgs boson masses and predictions including radiative corrections in a benchmark model are shown for a fixed value of $\tan\beta = 5$. On the right-hand side, the predictions in the benchmark model are shown for multiple values of $\tan\beta$. The diagonal gray dashed lines represent values of $m_{\mathrm{h/H}} = m_{\mathrm{A}}$, while the dashed horizontal lines indicate values of $m_{\mathrm{h/H}} = 125.38\,\mathrm{GeV}$. The considered benchmark scenario is the M_{h}^{125} scenario described in [63] and the predictions are obtained from [67].

values of m_{h} for low values of $\tan\beta$ and m_{A}, large values of M_{S} up to $10^{16}\,\mathrm{GeV}$ are necessary [65]. For these large values of M_{S}, the logarithmic corrections to prediction of m_{h} become large as well. They are resummed in the form of an effective field theory (EFT), in which the effective theory valid below the SUSY scale corresponds to a 2HDM, where all heavy supersymmetric degrees of freedom have been integrated out [65]. Consequently, the $M_{\mathrm{h,EFT}}^{125}$ scenario represents a low-$\tan\beta$ extension of the M_{h}^{125} scenario [63].

The effect of the radiative corrections on m_{h} is illustrated in the left panel of Figure 2.7, where a comparison of the predictions for the masses of the CP-even Higgs bosons at tree-level and in the M_{h}^{125} benchmark scenario [63] is shown. The comparison is shown as function of m_{A} for a fixed value of $\tan\beta = 5$. On the right-hand side of Figure 2.7, a comparison of the mass predictions for multiple values of $\tan\beta$ is shown in the M_{h}^{125} benchmark scenario. With increasing values of $\tan\beta$, the masses of both CP-even Higgs bosons take on the values of their upper and lower bounds faster as function of m_{A}.

As the Higgs sector of the MSSM corresponds to a Type II 2HDM, the coupling structure of the MSSM follows the one described in Section 2.3.1 with the parameter

$$\alpha = \frac{1}{2}\arctan\left(\tan 2\beta \frac{m_{\mathrm{A}}^2 + m_{\mathrm{Z}}^2}{m_{\mathrm{A}}^2 - m_{\mathrm{Z}}^2}\right), \quad -\frac{\pi}{2} \leq \alpha \leq 0 \tag{2.32}$$

given as function of m_{A} and $\tan\beta$ at tree-level [27]. An interesting phenomenological

limit of the Higgs sector in the MSSM is the decoupling limit which is reached for $m_A \gg m_Z$ [27]. With increasing values of $\tan\beta$, the MSSM Higgs sector reaches the decoupling limit for lower values of m_A.

In the decoupling limit, the heavy neutral Higgs bosons and the charged Higgs bosons decouple from the light neutral Higgs boson and become degenerate in mass. This mass degeneracy of the heavy neutral Higgs boson is also illustrated in Figure 2.7 for large values of m_A. In addition, the decoupling limit affects the couplings of the CP-even Higgs bosons. The couplings of h approach the couplings of h_{SM} and thus the couplings of H to vector bosons vanish. In general, the couplings of the two neutral heavy Higgs bosons become similar. Their couplings to down-type fermions grow proportionally to $\tan\beta$, while the couplings to up-type quarks decrease proportional to $\tan^{-1}\beta$. Since the direct couplings of the heavy neutral Higgs bosons to the vector bosons vanish, gluon fusion and heavy quark associated production constitute the main production modes for the heavy neutral Higgs bosons. With increasing values of $\tan\beta$ the cross section of the b quark associated production increases and turns into the major production mode.

In the case of the decay modes, not only their relevance is altered with respect to h_{SM} but also the number of allowed decay modes increases for the heavy neutral Higgs bosons. Additionally allowed decay modes include decays to top quarks, which open up for Higgs boson masses above $m_{H/A} = 2m_t$, decays to lighter neutral Higgs bosons as well as decays to supersymmetric partners of the SM particles. In most considered models, the decays to SUSY particles play only a minor role as the masses of these particles are assumed to be large since they have escaped observations so far. For large values of $\tan\beta$, the decays to b quarks and τ leptons dominate and thus provide prime candidates for the search for heavy neutral Higgs bosons in the MSSM.

The differences of the branching fractions of the neutral Higgs bosons for values of $\tan\beta = 7$ and $\tan\beta = 14$ and $m_A = 600\,\text{GeV}$ in the M_h^{125} benchmark scenario are illustrated in Figure 2.8. The branching fractions of h roughly correspond to the branching fractions of h_{SM} shown in Figure 2.4. Their differences for the selected values of $\tan\beta$ mainly originate from differences in the prediction of m_h at which they are calculated. For the heavy neutral Higgs bosons and $\tan\beta = 7$, the decay to a top quark pair constitutes a significant fraction of the total decay width, whereas it only plays a minor role for $\tan\beta = 14$.

The predictions for physical observables, i. e., the masses, partial decay widths and production cross sections, in the benchmark scenarios are obtained from three different programs [63]. `FeynHiggs` [69] is used to calculate the masses of the Higgs bosons in fixed order perturbation theory and additional resummation of leading logarithms in the ratio of the SUSY and electroweak scales. The precision of the calculations ranges up to two-loop corrections in specific approximations and partial resummation of next-to-next-to-leading logarithms (NNLL). In the $M_{h,\text{EFT}}^{125}$ scenario, the corrections are calculated at the one-loop level supplemented with the resummation in the 2HDM EFT. For a more

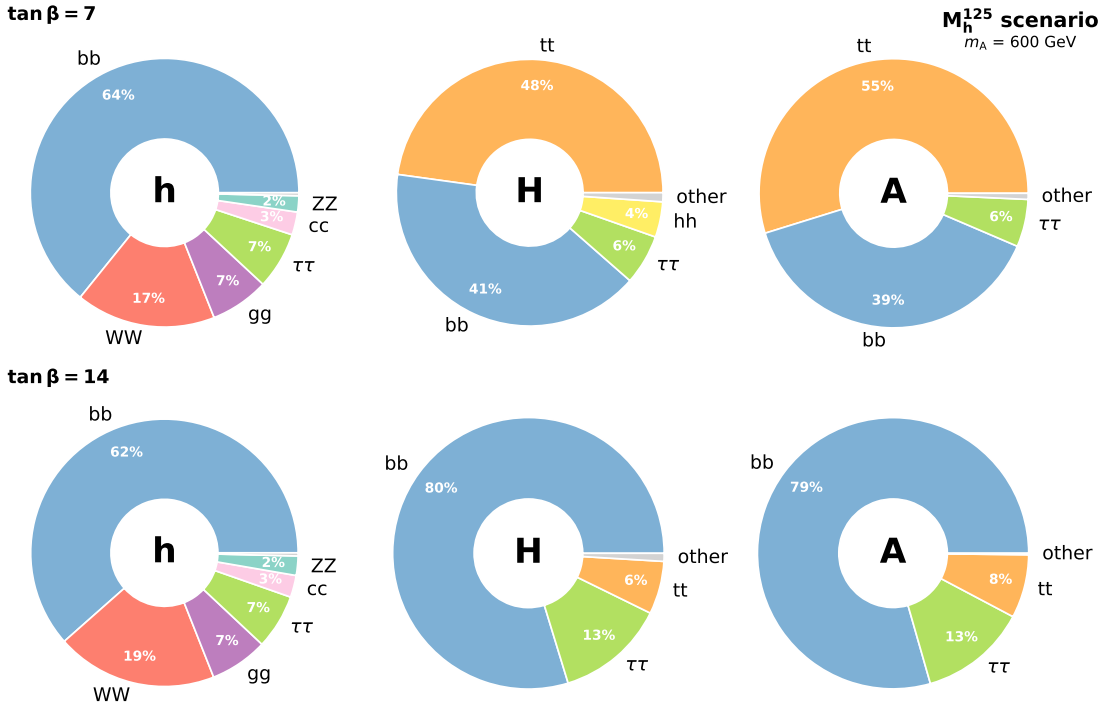

Figure 2.8: Predictions of the branching fractions of the neutral Higgs bosons in the M_{h}^{125} benchmark scenario [63]. Decay modes with branching fractions below 2 % are labeled as 'other'. The predictions are shown for $m_{\mathrm{A}} = 600\,\mathrm{GeV}$ and $\tan\beta = 7$ (upper row) or $\tan\beta = 14$ (lower row). The branching fractions are shown separately for the decays of the light CP-even, the heavy CP-even, and the CP-odd neutral Higgs bosons from left to right. The branching fraction predictions in the M_{h}^{125} benchmark scenario are obtained from [67].

detailed description of the calculations and their currently available precision, the reader is referred to [63]. The uncertainty in the mass predictions due to missing higher order corrections is assumed to amount to $\pm 3\,\text{GeV}$ [70].

In the M_{h}^{125} scenario, the partial decay widths are calculated using a combination of the programs `FeynHiggs` and `HDECAY` [71], which are used for different decay modes. In particular, the decay width of the decay into a pair of tau leptons is calculated with `FeynHiggs`, which is also used for all decay width predictions in the $M_{\text{h, EFT}}^{125}$ scenario. The calculation of the production cross sections in gluon fusion and b quark associated production are performed with `SusHi` [49, 50], as in the case of the 2HDM predictions. The precision in the calculation of the gluon fusion cross section reaches up to N3LO QCD in an expansion around the production threshold for h [63], whereas the cross section in b quark associated production is obtained from a reweighting of the predictions for h_{SM} with equal mass with the couplings expected for the considered parameter point in the benchmark scenario [63].

CHAPTER 3

The CMS Experiment

The presented search is performed using proton-proton (pp) collision data collected by the Compact Muon Solenoid (CMS) experiment [72]. The CMS detector is a multi-purpose detector targeting the detection of final state particles produced in pp collisions and is located at the Large Hadron Collider (LHC) [73].

The LHC is a proton and lead-ion collider at the European Council for Nuclear Research (CERN) near Geneva, Switzerland. It accelerates protons to energies of 6.5 TeV in two counter-rotating beams, which are brought to collision at four interaction points. It consists of eight straight sections and eight arcs. The interaction points and the radio-frequency (RF) cavities used to accelerate the particles are located at the straight sections. Protons are accelerated in a system of pre-accelerators to energies of 450 GeV before being injected into the LHC, where the final acceleration to 6.5 TeV takes place. As the protons are accelerated in RF cavities, they are not accelerated in continuous beams but are separated in bunches consisting of 1.15×10^{11} protons each [73]. In both beams, 2808 bunches are circulating with a minimal bunch spacing of 25 ns.

Around the LHC ring, three further interaction points are located at the straight sections equipped with other detectors, the ATLAS [74], LHCb [75], and ALICE [76] experiments. The ATLAS experiment and exploits the same physics programme as the CMS experiment. The LHCb and ALICE experiments are more specialized. LHCb is focused on the investigation of CP-violation and rare decays of mesons which contain b quarks [75]. The investigation of the quark-gluon plasma and the strong interaction in heavy-ion collisions constitutes the goal of the ALICE experiment [76].

3.1 Design of the CMS Experiment

The CMS experiment exhibits the typical structure common to all multi-purpose high-energy physics experiments with equal beam energies. It is constructed in layers of different subdetector components around the interaction point, where the collisions of the protons occur. The different detector layers together with the signatures of different particles in the detector are shown in Figure 3.1. A superconducting solenoid, constituting

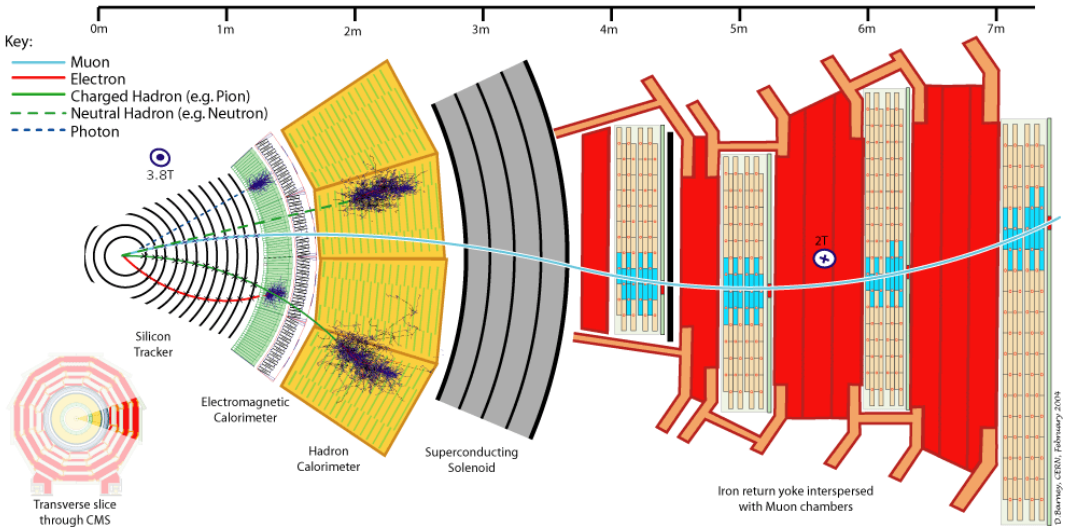

Figure 3.1: Structure of the CMS detector. The different sub-detectors of the CMS experiment and their layout are shown. Additionally, the signatures from different particles in the sub-detectors are indicated. Taken from [77].

one of the central features of the CMS experiment, creates the strong magnetic field of 3.8 T bending the trajectories of charged particles [72]. Inside the solenoid an all-silicon tracking detector, the electromagnetic (ECAL), and the hadron (HCAL) calorimeter are placed. The iron return yoke, which is interspersed with the muon system, forms the outermost part of the CMS experiment.

The tracking detector constitutes the innermost layer and is used to reconstruct the tracks of charged particles and the vertices from which they originate. The magnetic field from the superconducting solenoid bends the trajectories of charged particles and thus allows the measurement of their momenta in the plane perpendicular to the incident beams from the reconstructed tracks. The ECAL and HCAL destructively measure the energy of electrons, photons, and hadrons, respectively. The ECAL is placed inside of the HCAL as the hadronic interaction length λ_I is larger than the radiation length of the electromagnetic interaction, X_0, and thus hadrons traverse the ECAL, whereas electrons and photons are fully absorbed. Muons, which are minimum ionizing particles, traverse the tracking detector, the calorimeters, and the solenoid depositing only a small fraction of their energy. Thus, the muon detectors are placed inside the iron return yoke around the solenoid to identify muons and measure their momenta.

In the following, a short introduction to the coordinate system employed by the CMS experiment is given and afterwards the different detector subsystems are described in more detail. A more detailed description of the CMS detector is given in [72].

Coordinate System

The origin of the CMS coordinate system is placed at the nominal interaction point, where the pp collisions are expected to take place. It is a right-handed coordinate system with the x-axis pointing to the center of the LHC and the y-axis pointing upwards. The z-axis then points in the direction of the counterclockwise circulating beam inside the LHC. As the CMS detector exhibits a cylindrical symmetry, the cartesian coordinate system is not used directly in most cases but a cylindrical coordinate system is defined. The radial component and the azimuthal angle φ of the coordinate system are defined in the x-y plane longitudinal to the beam axis, where φ is measured starting from the x-axis. The radial component of the momentum of particles, the transverse momentum p_{T}, is calculated as

$$p_{\mathrm{T}} = \sqrt{p_x^2 + p_y^2}\,. \tag{3.1}$$

The polar angle ϑ is measured with respect to the z-axis and the pseudorapidity η is defined via

$$\eta = -\ln\left(\tan\frac{\vartheta}{2}\right)\,. \tag{3.2}$$

The angular separation ΔR of two physics objects is often expressed as function of η and ϕ angle via

$$\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\varphi)^2}\,. \tag{3.3}$$

Tracking Detector

The tracking detector of the CMS experiment is composed of multiple layers of silicon modules with two different segmentations. The pixel detector constitutes the inner part of the tracking detector. It consists of silicon modules exhibiting a two-dimensional (2D) segmentation with a pixel size of $100 \times 150\,\mu\mathrm{m}$ [72]. The pixel detector has been upgraded after the 2016 data-taking period [78]. This results in 74 % of the total analyzed data that have been taken with the upgraded pixel detector. Before the upgrade, the pixel detector consisted of three layers of modules in the barrel region, that were complemented by two endcap disks at each side of the detector [72]. During the upgrade, an additional layer of pixel modules has been added and the innermost layers in the barrel and endcap regions have been brought closer to the interaction point. After the upgrade, the pixel detector in total consists of 124 million pixels [78].

Around the pixel detector, the silicon strip detector is located. It consists of 10 layers of silicon strip modules in the barrel region and 3 plus 9 disks in the inner and outer endcap regions [72]. The modules in the barrel and endcap regions are mounted parallel and perpendicular to the beam axis, respectively. While the sensors of the pixel detector natively yield 2D information on single particle hits, in some layers of the strip detector the modules are mounted back-to-back under a small angle to obtain a 2D resolution as

well [72]. With the upgrade of the pixel tracking detector, charged particles cross at least four detection layers in the pixel detector in addition to the nine detection layers in the strip tracking detector. At least four of the crossed detection layers in the strip detector additionally yield 2D information on the hit position. Both the pixel and strip detectors exhibit a pseudorapidity coverage up to $|\eta| = 2.5$ [72].

The pixel detector reaches single hit resolutions of 10 to 20 µm [78], while the single hit resolution of the strip detector ranges from 23 to 53 µm in the different detector layers [72]. This translates to a transverse momentum resolution of 1–2 % for reconstructed tracks with transverse momentum of $p_{\mathrm{T}} = 100\,\mathrm{GeV}$ up to pseudorapidities of $|\eta| < 1.6$ [72]. The impact parameter resolution, which is of importance for the identification of tau lepton decays and b quark-initiated jets (b jets) described in the following, amounts to 10 µm for high momentum tracks [72]. The impact parameter corresponds to the shortest distance of a reconstructed track to the point of the collision from which it emerged. Its resolution is limited by the single hit resolution of the first pixel layer [72]. Material in the tracking detector limits the performance of the track reconstruction through multiple scattering, bremsstrahlung of electrons and pair conversion of photons. The material inside the tracking detector quantified in units of X_0 by the material budget ranges from $0.4\,X_0$ in the central region to a maximum value of $1.8\,X_0$ around $|\eta| = 1.4$ [72].

Electromagnetic Calorimeter

The ECAL measures the energy of electrons, positrons, and photons via the electromagnetic showers that these particles induce. The measurement is performed through the conversion of the energy deposited in the shower to scintillation light. The ECAL is built as homogeneous calorimeter and consists of 75 848 lead tungstate ($PbWO_4$) crystals that simultaneously act as absorber and scintillation material [72]. The barrel region covers a range of $|\eta| < 1.479$ and is equipped with 61 200 crystals. The η coverage of the ECAL endcaps extends up to $|\eta| = 3$. Avalanche photodiodes and vacuum phototriodes convert the scintillation light emitted by the $PbWO_4$ crystals into electronic signals in the barrel and endcap regions, respectively. The length of the crystals in the barrel and endcap regions corresponds to around $25\,X_0$.

The ECAL is supported by a preshower detector, which improves the identification of neutral pions in the high-pseudorapidity region of $1.653 < |\eta| < 2.6$ [72]. The preshower detector is constructed as a sampling calorimeter employing two layers of lead absorbers and silicon strip sensors each to measure the energy of the induced shower. The total depth of both layers amounts to $3\,X_0$.

The energy resolution of calorimeters is composed of three terms with different functional dependence on the energy of the incident particle [72]. It is given by

$$\frac{\sigma_E}{E} = \frac{S}{\sqrt{E}} \oplus \frac{N}{E} \oplus C\,, \tag{3.4}$$

where the $\oplus$ sign represents quadratic addition of the operands. The first term in Equation 3.4 is the stochastic term, the second term the noise term and the third the constant term. The stochastic term parameterizes fluctuations in the shower-development and the readout of the photomultipliers [72]. The noise term includes the effects of electronic and digitization noise on the energy resolution and the constant term models effects that increase linearly with the energy of the incident particle. The values of the constants in Equation 3.4 have been measured in a test beam setup to $S = 2.8\,\%\,\mathrm{GeV}^{1/2}$, $N = 0.13\,\mathrm{GeV}$, and $C = 0.3\,\%$ [79]. Thus, the constant term dominates the energy resolution for energies above 85 GeV, whereas the contribution of the noise term is the largest for energies below 20 GeV.

Hadronic Calorimeter

The HCAL is used to measure the energy of charged and neutral hadrons produced in the proton collisions at the center of the CMS detector. It is placed between the ECAL and the superconducting solenoid [72]. The HCAL consists of three parts with different η coverages. In the central detector region, the HCAL barrel with a η coverage up to $|\eta| < 1.3$ is placed. The HCAL endcap and forward calorimeters extend this range to values of $|\eta| < 3$ and $|\eta| < 5.2$, respectively.

The barrel and endcap calorimeters are built as sampling calorimeters consisting of brass absorber layers and plastic scintillator tiles. The forward calorimeter is built from steel absorber plates instrumented with quartz fibers, which create photons via Cherenkov radiation. The size of the HCAL is limited by its placement inside the superconducting solenoid. The depth of the HCAL barrel in the central region amounts to $5.8\,\lambda_\mathrm{I}$ and increases with $|\eta|$ to values of $10.6\,\lambda_\mathrm{I}$ at $|\eta| = 1.3$. The ECAL contributes another $1.1\,\lambda_\mathrm{I}$ to the total length of the calorimeter. The total length of the HCAL endcap including the ECAL amounts to $10\,\lambda_\mathrm{I}$.

To increase the total thickness of the calorimeter in the barrel region and to prevent leakage of the hadronic showers beyond the calorimeter, the hadronic outer calorimeter is placed around the superconducting solenoid. It consists of an additional layer of scintillator tiles and uses the superconducting solenoid as absorber material. In the central detector region, where the total thickness of the calorimeter is the smallest, two layers of scintillator tiles and an additional iron absorber are used to further increase the thickness of the calorimeter. Including the outer calorimeter, the minimal total thickness of the calorimeter is $11.8\,\lambda_\mathrm{I}$ [72].

The combined energy resolution of the ECAL and HCAL barrel for pions has been measured in a test beam setup to amount to [80]

$$\frac{\sigma_E}{E} = \frac{111\,\%\,\mathrm{GeV}^{1/2}}{\sqrt{E}} \oplus 9\,\%\,. \tag{3.5}$$

Muon System

The muon systems constitutes the outermost layer of the CMS experiment. It consists of three different detector systems located inside the iron return yoke of the magnet [72]. In the barrel region, four stations containing drift tube (DT) chambers are placed in the yoke and cover an η range up to $|\eta| < 1.2$. In the first three stations counted from the interaction point, a three-dimensional measurement of the muon coordinates is performed, whereas in the fourth layer only the radial and azimuthal coordinates are measured. The DT chambers are constructed to reach a coordinate resolution of 100 µm.

In the endcaps of the muon systems, the coordinates of the muons are measured with four stations of cathode strip chambers (CSC). The CSCs cover an η range of $0.9 < |\eta| < 2.4$. The DT chambers and CSCs are complemented with resistive plate chambers (RPC) that cover an η region up to $|\eta| < 1.6$. The RPCs deliver a fast response and exhibit a good time resolution and are thus used for fast decisions in the trigger system. In the barrel region, six layers of RPCs are mounted around the DTs, while in the endcap regions three layers are mounted.

Trigger and Data Acquisition

During LHC Run 2, the LHC has been operated at a bunch crossing frequency of 40 MHz. However, the reconstruction and data storage infrastructure of CMS can only sustain an event rate of 1 kHz [81]. Additionally, only a small fraction of the 40 million collisions per second is interesting for searches for additional neutral Higgs bosons or analyses of other physics processes. The CMS experiment employs a two-tiered trigger system to reduce the event rate in two steps and select events of interest for physics analyses [81].

The level-1 (L1) trigger [82] constitutes the first tier of the CMS trigger system. It is a hardware based trigger consisting of multiple field programmable gate arrays (FPGA). The L1 trigger itself is composed of two subsystems, which independently reconstruct electron, photon, jet, and hadronically decaying tau lepton candidates from energy deposits in the calorimeters and muons from detector signals in the muon system. The reconstructed objects from both subsystems are combined to decide whether the full detector information from the event should be extracted and the event further reconstructed by the high-level trigger (HLT) [81]. The L1 trigger reduces the event rate by a factor of roughly 400 from 40 MHz to 100 kHz, which corresponds to the maximum rate at which the full event information can be extracted from the CMS experiment [81].

The HLT is a software based trigger running a version of the offline reconstruction algorithms, described in the following section, which is optimized for faster processing, on a server farm close to the detector. The HLT consists of multiple sequences of modules reconstructing physics objects and directly imposing selection criteria on produced quantities of the objects. Before the time consuming optimized track reconstruction and particle flow algorithm, described in the following section, are executed, preselections

based on fast reconstructed quantities like local energy deposits in the calorimeters are performed. The HLT reduces the output rate of the L1 trigger further to an output rate of 1 kHz [82]. Events selected by at least one of the defined module sequences are passed on to the offline reconstruction facility, where the reconstruction of physics objects takes place, and the reconstructed events are stored persistently for analysis.

3.2 Reconstruction of Physics Objects

Particle Flow Algorithm

The Particle Flow (PF) algorithm [83] combines the measurements in different detector systems to identify single particles and yield a global description for each recorded event. The basic elements entering the PF algorithm are the reconstructed tracks of charged particles reconstructed from hits in the inner tracking detector and the muon system, and clusters of energy deposits in the calorimeters.

Tracks of the charged particles are reconstructed following an iterative procedure to achieve high tracking efficiencies while maintaining small misidentification rates [83, 84]. Each track reconstruction iteration consists of three steps following the Kalman filtering (KF) algorithm [85]. First, a track seed is built from a small set of hits in the pixel and the strip tracking detectors for later iterations. Starting from the seeds, further hits compatible with the reconstructed track are searched for and associated to the track if they are found to be compatible. Lastly, a fit of the track parameters is performed to extract the origin, transverse momentum, and direction of the track. As hits already used to form a track in earlier iterations are vetoed for further iterations, the probability to combine unrelated hits to a track seed decreases with the number of completed iterations. Thus, looser criteria can be imposed on the track seeds in later iterations to increase the overall efficiency of the tracking algorithm. In addition, the reconstructed tracks are subject to quality criteria to reduce the misidentification rate before using them in the PF algorithm .

The positions of primary vertices are calculated from the reconstructed tracks of charged particles using an adaptive vertex fit [86], which clusters the reconstructed tracks to vertices based on their position in z direction along the beam line. Multiple pp collisions occur at a single bunch crossing leading to additional reconstructed vertices from which particles with lower transverse momenta emerge. The identification of the primary vertex (PV) is crucial to reduce the impact of the additional collisions on the analysis and, e. g., to correctly identify b jets. To identify the PV, the reconstructed tracks of charged particles associated to each vertex are clustered into track jets using the anti-k_{T} algorithm [87, 88] with a distance parameter of $R = 0.4$. The vertex with the largest $\sum p_{\mathrm{T}}^2$ is identified as the PV [89]. The transverse momenta of track jets, remaining tracks not clustered in the jets, and the missing transverse momentum at each vertex contribute to the calculation of $\sum p_{\mathrm{T}}^2$. Each of the contributions may enter the sum with a different weight.

The clustering of energy deposits in the calorimeters is performed separately for the barrel and endcap regions of both the ECAL and HCAL and the two layers of the ECAL preshower detector [83]. Local cluster seeds are identified as calorimeter cells with local maxima of the deposited energy. Topological clusters of energy deposits are built by iteratively combining calorimeter cells with energy deposits above a certain threshold surrounding these local cluster seeds. Single calorimeter clusters are then reconstructed from the topological clusters using a Gaussian mixture model. As only cells above a specific energy threshold are used for the cluster reconstruction, the energy of the clusters is biased to smaller values. Thus, the cluster energies are calibrated, separately for the ECAL and HCAL, using isolated photons and neutral hadrons respectively, in simulation.

To distinguish the types of particles produced in the proton collisions in the reconstruction, reconstructed PF elements from the different detector systems are combined to PF blocks. The combination of PF elements to PF blocks is performed using a linking algorithm based on the distance of the PF elements [83]. Reconstructed tracks are linked separately to ECAL preshower, ECAL, or HCAL clusters by interpolating the reconstructed track into the respective calorimeter. Links between clusters in different calorimeters are established if the position of the cluster in the more granular calorimeter falls into the second cluster. Particle candidates are reconstructed sequentially from PF blocks and elements, starting with muons, followed by electrons and isolated photons. In the next step, charged and neutral hadrons, and photons in their vicinity are reconstructed. After each reconstruction step, the used PF blocks are removed from the set of PF blocks considered for the following reconstruction steps. Lastly, a postprocessing step takes place which uses global event information to reduce the effect of particle misidentifications on the missing transverse momentum.

Muons

Muon candidates are reconstructed from tracks in the inner tracking detector and track segments reconstructed in the muon system [90]. Track segments are a combination of hits in a DT or CSC detector and serve as seeds for the tracking algorithm in the muon system. The tracking algorithm combines hits in all three muon systems to standalone muon tracks.

Two further types of muons are defined as well, global muons and tracker muons. Global muons are obtained through the combination of a standalone muon track and a track in the inner tracking detector if the track parameters of both tracks extrapolated to a common surface match within their uncertainties [83]. The track properties of the global muon track are then extracted from a combined fit of the hits in the muon detector and the inner tracking detector. The combination of both detector subsystems in the reconstruction of global muon tracks leads to an improved transverse momentum resolution for muon transverse momenta larger than 200 GeV.

The third type of muons are tracker muons, which are reconstructed from tracks in the inner tracking detector that are compatible with track segments in the muon system. The compatibility of the track and the track segment is quantified by their distance in a plane transverse to the beam direction. The efficiency of the muon reconstruction as either global or tracker muon amounts to 99 % [90]. To be identified as PF muons, the reconstructed muon candidates are required to exhibit an isolation relative to its transverse momentum smaller than 0.1, where the isolation is calculated as the sum of the transverse momenta of reconstructed tracks and the energy deposits in the calorimeter in a cone with radius $\Delta R = 0.3$ around the direction of the muon.

Two different sets of identification criteria are defined to identify muons directly produced in a hard interaction (prompt muons) and muons from light and heavy flavor hadron decays [90]. To be identified by the loose identification criteria, muon candidates are required to be either tracker muons or global muons identified as PF muons. To be identified as medium muons, selected muons are required to pass certain selection criteria imposed on the quality of the reconstructed track and the compatibility of the tracks in the inner tracking detector and the muon system in addition to the loose identification criteria. The complete list of the imposed criteria on the muons is given in [90]. The medium muon identification criteria target a selection efficiency of 99.5 % for genuine muons [90].

These identification criteria are supplemented with isolation criteria [90] imposed on the muons. The isolation of a muon candidate is defined as the sum of the transverse momenta of all PF candidates originating from the event primary vertex within a cone of radius $\Delta R = 0.4$ around the muon direction. Charged hadrons can be directly associated to the PV and their contribution to the isolation of the muon candidate directly computed as the sum of their transverse momenta. This is not the case for neutral particles, and thus their contribution to the isolation sum is overestimated. To correct for this additional bias, the contribution of neutral particles from pileup to the isolation is subtracted from the total contribution from neutral particles. The combined relative isolation of the muon candidate is thus given by

$$I_{\text{rel}} = \frac{1}{p_{\text{T}}} \left[I_{\text{ch}} + \max\left(I_{\text{n}} + I_{\gamma} - I_{\text{PU}}, 0 \right) \right] , \tag{3.6}$$

where I_{ch} represents the contribution from charged hadrons originating from the primary vertex, and I_{n} and I_{γ} the contributions from neutral hadrons and photons, respectively. Contributions from neutral particles from pileup I_{PU} are estimated as half the contribution from charged particles not originating from the PV $I_{\text{PU}} = 1/2 \sum p_{\text{T}}^{\text{ch,PU}}$. Loose and Tight working points (WP) of the isolation criterion are defined by requiring the relative isolation to be below 0.15 and 0.25, corresponding to selection efficiencies of 95 % and 98 %, respectively [90].

Electrons and Positrons

The signature of electrons in the detector consists of a track in the inner tracking detector and energy deposits in the ECAL [91]. Electrons traversing the tracking detector can also interact with the material in the tracker and emit bremsstrahlung. The energy loss of the electron due to bremsstrahlung depends on the η of the electron as the material budget increases for larger η. As opposed to the electron trajectories, the trajectories of the emitted bremsstrahlung photons are not bend in the magnetic field and thus constitute tangents to the trajectory of the electron. This leads to a spread of the energy deposits in azimuthal direction, that is not directly associated to the reconstructed track of the electron.

The reconstruction of electrons is based on the reconstruction of PF blocks [91]. Adaptations to the reconstruction of calorimeter clusters and the tracking algorithm are introduced in the PF algorithm to account for the bremsstrahlung losses of the electrons. Superclusters of energy deposits are built from multiple clusters in the electromagnetic calorimeter to collect the full energy of the initial electron and the bremsstrahlung photons. They are constructed by combining adjacent clusters in a region around a seeding cluster, that has a larger extension in the ϕ than in the η direction.

Two different approaches are combined to seed the reconstruction of electron candidates [83]. The ECAL-based approach reconstructs electron seeds with large transverse momenta with an efficiency larger than 95 % [91]. It starts from clusters in the ECAL and tries to find hits in the innermost layers of the tracking detector that are compatible with an electron or positron with the energy and position of the cluster. The tracker-based seeding approach is introduced to increase the reconstruction efficiency for low-p_{T} electrons and electrons contained in hadronic jets. It propagates tracks from the iterative tracking procedure to the inner ECAL surface and links them to clusters localized at the propagated position. Electron seeds from both approaches are passed on to the track fitting procedure for electrons, which uses a Gaussian-sum filter (GSF) [92] with twelve components to reconstruct the electron tracks in presence of energy losses through bremsstrahlung [83]. A more detailed description of the electron reconstruction is given in [91].

The calibration of the reconstructed electron energy is performed in three steps exploiting a multivariate (MVA) regression method based on boosted decision trees (BDT) [91]. In the first and second step, a correction to the energy response and resolution of the associated supercluster is derived. The combination of the calibrated energy of the supercluster and the momentum of the GSF track via a weighted average constitutes the final step of the calibration procedure. The inclusion of the GSF track in the energy calibration improves the energy resolution for transverse momenta below 15 GeV [91].

Two different algorithms are employed to identify electrons reconstructed by the PF algorithm and discriminate them against backgrounds arising from photon conversions in the tracker, misidentification of hadrons, or from heavy flavor hadron decays [91]. The first

selection algorithm consists of a set of eight requirements that are applied sequentially and is further referred to as cut-based identification. The used selection criteria include requirements on shower-shape, the compatibility of the GSF track with the energy and position of the supercluster, the number of missing hits in the tracking detector, the combined relative isolation as defined below, and a discriminant to suppress photon conversions. Four different WPs with efficiencies ranging from 70 to 95 % are defined. The WP with the highest selection efficiency is employed in the analysis to veto events with additional electrons.

The second identification algorithm, referred to as MVA-based identification algorithm, is based on multiple BDTs used in distinct p_{T} and η regions. The set of input variables is extended with respect to the variables used for the cut-based isolation criteria. The additional input variables include the track-to-cluster matching variables evaluated at the ECAL surface, further cluster-shape and track-quality variables, and the track momentum difference between entering and leaving the tracking detector relative to the track momentum when entering the tracking detector. The combined relative isolation is not used in the MVA-based identification algorithm. Two WPs with efficiencies of 80 and 90 % are defined. Both identification algorithms are optimized on data sets simulated with the detector conditions from the 2017 data-taking year and applied to all data-taking years. The MVA-based identification criteria are trained on simulated Z + jet events, whereas the cut-based identification algorithm is optimized on simulated $\mathrm{t\bar{t}}$ events.

As the MVA-based identification criteria do not include isolation information, requirements on the combined relative isolation are imposed on identified electrons to discriminate them against electrons from heavy flavor hadron decays in hadronic jets. The definition of the isolation criteria is similar to the one imposed on muons given in Equation 3.6 [91]. Different from the case of muon isolation, the isolation requirement of electrons is computed in a cone with radius $\Delta R = 0.3$ and the contribution of pileup is mitigated using a different method. The contribution from pileup is estimated as $I_{\mathrm{PU}} = \rho A_{\mathrm{eff}}$, where ρ is the median of the transverse energy density and A_{eff} the effective area of the isolation region. It corresponds to the area of the isolation cone weighted by a factor to account for the pseudorapidity dependence of the energy distribution from pileup.

Hadronic Jets

Hadronic jets consist of collimated bunches of particles produced in the pp collisions. Energy deposits in the ECAL and HCAL associated with multiple tracks form the signature of hadronic jets in the detector. The reconstruction of hadronic jets is performed by clustering PF candidates with the anti-k_{T} algorithm [87, 88] with a radius parameter of $R = 0.4$. To suppress PU contributions to the reconstructed jets, charged hadron candidates associated to a different vertex than the PV are removed in the clustering process [93].

The measured energy response of the reconstructed jets is calibrated in a multi-step procedure [93]. A different combination of the correction steps is used to correct the jets in data and in simulation. A more detailed description of the steps of the jet energy response corrections is given in [93]. In a first step, the energy of the reconstructed jets is corrected for the contributions of neutral particles from pileup in the previous, the same, and the following bunch crossings. The correction factor is derived in simulated events by clustering jets in the same event with and without overlayed pileup collisions as function of p_{T} and η of the jets. Residual differences of the pileup corrections between data and simulation are corrected by measuring the energy density in an event using randomly placed cones to collect energy deposits. In the next step, the jet response in data and simulation is corrected as function of the p_{T} and η of the jets. The correction is derived as the ratio of the reconstructed jet energy over the energy of the matched jet clustered from simulated particles.

Residual differences between data and simulation are corrected by exploiting momentum conservation in the transverse plane. The protons in the beams, and thus also their colliding constituents, exhibit negligible momentum in the transverse plane. Therefore, the vectorial sum of the momenta of all final state particles in the transverse plane must also vanish. Corrections to the jet energy response are then derived based on the balance of the considered jet with a reference object. First, the response in different η regions is calibrated with respect to the response in the central η region ($|\eta| < 1.3$) using the missing transverse momentum projection fraction (MPF) in dijet events. Afterwards, the jet energy response is corrected as function of the transverse momentum in $\mathrm{Z}/\gamma + \mathrm{jet}$ events by exploiting the balance of the Z boson with the reconstructed jet in the transverse plane.

Hadronically Decaying Tau Leptons

With a mass of 1.78 GeV [35], the tau lepton is the heaviest known lepton. It decays into electrons, muons, or hadrons and one or more corresponding neutrinos before reaching the first layer of the detector. A list of the decay modes and their branching fractions is given in Table 3.1. The main hadronic decay modes include decays into one or three charged hadrons and neutral pions. Neutral pions directly decay further into a pair of photons. The photons may convert in electron positron pairs while traversing the inner tracking detector. The strong magnetic field separates the electrons and positrons in azimuthal direction due to their opposite charges.

The experimental signature of the leptonic decay modes corresponds to the signature of prompt electrons or muons with an increased impact parameter corresponding to the mean life time of tau leptons of $c\tau = 87\,\mu\mathrm{m}$ [35]. The electrons or muons originating from the decays are reconstructed as described above. The signature of a hadronic tau lepton decay is a narrow hadronic jet of low particle multiplicity, that is isolated from any further hadronic activity.

Table 3.1: Tau lepton decay modes and branching fractions $\mathcal{B}$. The decay modes are listed for the negatively charged tau lepton. The charge conjugated decay modes of positively charged tau antileptons exhibit the same branching fraction. The symbol $\mathrm{h}^{\pm}$ denotes generic charged hadrons. The given branching fractions are taken from [35].

Decay mode	Intermediate resonance	$\mathcal{B}$ (%)
$\mathrm{e}^{-}\bar{\nu}_{\mathrm{e}}\nu_{\tau}$		17.8
$\mu^{-}\bar{\nu}_{\mu}\nu_{\tau}$		17.4
h^{-}		11.5
$\mathrm{h}^{-}\pi^{0}$	$\rho(770)$	25.9
$\mathrm{h}^{-}\pi^{0}\pi^{0}$	$\mathrm{a}_1(1260)$	9.5
$\mathrm{h}^{-}\mathrm{h}^{+}\mathrm{h}^{-}$	$\mathrm{a}_1(1260)$	9.8
$\mathrm{h}^{-}\mathrm{h}^{+}\mathrm{h}^{-}\pi^{0}$		4.8
Other		3.3

Hadronic decays of tau leptons, referred to as τ_{h} candidates in the following, are reconstructed using the hadrons-plus-strips (HPS) algorithm [94–96]. This algorithm starts from jets reconstructed with the anti-k_{T} algorithm and builds decay mode hypotheses from the particles contained in a cone of radius $\Delta R = 0.5$ around the direction of the jet. These decay mode hypotheses are constructed from charged hadrons and strips in the η-φ plane, which aim to collect the photons and electrons originating from decays of neutral pions. Charged hadron candidates are required to exhibit a transverse momentum larger than 0.5 GeV and to originate from the PV with a transverse impact parameter $|d_{xy}| < 0.1\,\mathrm{cm}$.

The strips are reconstructed from electron and photon candidates inside the cone around the jet in an iterative procedure [96]. This procedure starts from the electron or photon candidate with the largest p_{T} and adds further candidates to the strip if they fall in a $\Delta\eta$-$\Delta\varphi$ window around the position of the strip. The $\Delta\eta$-$\Delta\varphi$ window is dynamically adjusted taking the transverse momentum of the considered electron or photon candidate and the strip into account. The window sizes are given by

$$\Delta\eta = f(p_{\mathrm{T}}^{\mathrm{e}/\gamma}) + f(p_{\mathrm{T}}^{\mathrm{strip}}) \quad \text{and} \tag{3.7}$$

$$\Delta\varphi = g(p_{\mathrm{T}}^{\mathrm{e}/\gamma}) + g(p_{\mathrm{T}}^{\mathrm{strip}})\,, \tag{3.8}$$

where f and g are both functions of the form $a \cdot p_{\mathrm{T}}^{-b}$ with different constants a and b, which have been optimized such that the strip contains 95 % of the electron and photon candidates originating from the tau lepton decay in simulation. The maximum values of $\Delta\eta$ and $\Delta\varphi$ are set to 0.15 and 0.3, respectively, where the larger extent in φ direction has been chosen to collect the electrons and positrons separated by the magnetic field. The new position of the strip is then calculated as the p_{T} weighted average of all electron and photon candidates contained in the strip. The procedure ends when no further candidate is found that either falls into the reconstructed strip or seeds the reconstruction of a

further strip.

Four different decay mode hypotheses are constructed from charged hadron candidates and the reconstructed strips. The decay mode hypotheses include one prong decays ($h^{\pm}$), one prong decays in association with additional neutral pions ($h^{\pm}\pi^0$, $h^{\pm}\pi^0\pi^0$), three prong decays ($h^{\pm}h^{\mp}h^{\pm}$), and three prong decays in association with neutral pions ($h^{\pm}h^{\mp}h^{\pm}\pi^0$). Constructed decay mode hypotheses with an absolute value of the charge not equal to unity, or with charged hadron candidates or strips outside the signal cone of the tau lepton are rejected, where the signal cone is defined as a cone of radius $R_{\text{sig}} = 0.05 < 3\,\text{GeV}/p_{\text{T}}^{\tau_{\text{h}}} < 0.1$ around the direction of the tau lepton candidate, i. e., a fixed size of $R_{\text{sig}} = 0.1$ is used for $p_{\text{T}}^{\tau_{\text{h}}} > 30\,\text{GeV}$. For the decay modes proceeding via intermediate resonances, additional requirements are imposed on the reconstructed visible mass of the candidate, $m_{\tau_{\text{h}}}$, to decrease the misidentification probability while maintaining a high reconstruction efficiency. The requirements for the different decay mode hypotheses are:

- $h^{\pm}\pi^0$: $0.3\,\text{GeV} - \Delta m_{\tau_{\text{h}}} < m_{\tau_{\text{h}}} < 1.3\,\text{GeV}\sqrt{p_{\text{T}}^{\tau_{\text{h}}}/100\,\text{GeV}} + \Delta m_{\tau_{\text{h}}}$
- $h^{\pm}\pi^0\pi^0$: $0.4\,\text{GeV} - \Delta m_{\tau_{\text{h}}} < m_{\tau_{\text{h}}} < 1.2\,\text{GeV}\sqrt{p_{\text{T}}^{\tau_{\text{h}}}/100\,\text{GeV}} + \Delta m_{\tau_{\text{h}}}$
- $h^{\pm}h^{\mp}h^{\pm}$: $0.8 < m_{\tau_{\text{h}}} < 1.5\,\text{GeV}$,

where $\Delta m_{\tau_{\text{h}}}$ represents the change in the visible mass of the candidate originating from the addition of e/γ candidates to the strip. For the one prong decay modes, the upper boundary on the mass requirement is relaxed for transverse momenta above 100 GeV to account for resolution effects. The maximum values of the upper boundaries are chosen to be 4.2 and 4 GeV for the $h^{\pm}\pi^0$ and $h^{\pm}\pi^0\pi^0$ decay modes, respectively.

The reconstructed decay mode hypothesis with the largest transverse momentum is chosen as the τ_{h} candidate. Thus, there is at most one τ_{h} candidate reconstructed per jet seeding the HPS algorithm. Even though the $h^{\pm}\pi^0$ and $h^{\pm}\pi^0\pi^0$ decay modes are reconstructed in slightly different ways, they are analyzed together and referred to as $h^{\pm}\pi^0$ decay mode in the following. The overall reconstruction efficiency of the HPS algorithm for τ_{h} candidates with transverse momenta above 30 GeV amounts to roughly 70 %, where the highest reconstruction efficiency is achieved for the $h^{\pm}$ decay mode [97].

Multiple physics objects may be wrongly identified as τ_{h} candidates. Hadronic jets initiated from quarks or gluons exhibit the largest misidentification probability because their experimental signature is similar to the signature of hadronic tau lepton decays. The misidentification of these hadronic jets asτ_{h} candidates is referred to as jet $\rightarrow \tau_{\text{h}}$ throughout the rest of the thesis. Electrons ($e \rightarrow \tau_{\text{h}}$) and muons ($\mu \rightarrow \tau_{\text{h}}$) also show potential for misidentification as τ_{h} when occurring in combination with energy deposits in the HCAL. These additional energy deposits can be created for example from electrons forming electromagnetic showers leaking into the HCAL or nuclear interactions of muons

within the HCAL.

To efficiently identify hadronic tau lepton decays and discriminate them against jet $\rightarrow \tau_\mathrm{h}$, $\mathrm{e} \rightarrow \tau_\mathrm{h}$, and $\mu \rightarrow \tau_\mathrm{h}$ misidentification, the DeepTau [97] algorithm is employed. It combines low-level reconstruction information in the form of quantities related to single reconstructed particles with information on the reconstructed τ_h candidate and event level quantities in a convolutional neural network. A more detailed description of the network architecture and the input variables is given in [97]. The neural network consists of three different subnetworks processing the low-level information, in two grids of different sizes, and the high-level information separately. In a final step, the outputs from the three subnetworks are combined and passed to a dense network with an output node for each of the four possible targets.

The outputs of the final neural network can be interpreted as estimates for the probability y_α of the reconstructed τ_h candidate to originate from a genuine tau lepton decay ($\alpha = \tau$), a quark- or gluon-initiated hadronic jet ($\alpha = \mathrm{jet}$), an electron ($\alpha = \mathrm{e}$), and a muon ($\alpha = \mu$). The probability estimates are then combined via

$$D_\alpha(\boldsymbol{y}) = \frac{y_\tau}{y_\tau + y_\alpha} \tag{3.9}$$

to obtain the DeepTau discriminants against jets (D_jet), electrons (D_e), and muons (D_μ). For each discriminant, multiple WPs are defined based on the expected efficiency to identify genuine hadronic tau lepton decays. The medium WP of the D_jet discriminant for example corresponds to an expected efficiency of 70 % to identify genuine tau lepton decays [97]. Taking also the reconstruction efficiency for τ_h candidates into account, this yields an efficiency around 50 % for the reconstruction and identification of genuine hadronic tau lepton decays.

Missing Transverse Momentum

Particles that only interact weakly, e. g., neutrinos, do not deposit energy in any of the detector systems and can thus not be detected directly. As the colliding initial state partons do not exhibit a sizable transverse momentum, momentum conservation ensures that also the sum of the transverse momenta of all final state particles must be zero. Thus, the presence of weakly interacting particles and the sum of their transverse momenta can be inferred from the $\vec{p}_\mathrm{T}$ that is missing to combine the $\vec{p}_\mathrm{T}$ of all visible particles to zero. In events in which more than one weakly interacting particle is present, the inference of the number of present particles and the association of $\vec{p}_\mathrm{T}^{\,\mathrm{miss}}$ to single particles is not possible.

The raw missing transverse momentum is calculated from the reconstructed PF candidates as the negative vectorial sum of their weighted transverse momenta [98]

$$\vec{p}_\mathrm{T}^{\,\mathrm{miss,\,raw}} = -\sum_{i=1}^{N_\mathrm{particles}} w_i \vec{p}_\mathrm{T}^{\,i}\,, \tag{3.10}$$

following the pileup per particle identification method [99, 100]. This method assigns weights to each reconstructed particle indicating the probability of the particle to originate from the PV. For charged particle candidates, a direct association of the particle to the PV is possible through the requirement of compatibility of the reconstructed track with the PV. Thus, charged particles are assigned a weight of $w_i = 1$ if they are associated with the PV and $w_i = 0$ otherwise.

The calculation of the weight for neutral particle candidates is based on the local shape variable α_i defined via

$$\alpha_i = \log \sum_{j \neq i, \Delta R_{ij} < 0.4} \left(\frac{p_{\mathrm{T},j}}{\Delta R_{ij}} \right)^2 , \tag{3.11}$$

where for particles inside the acceptance of the inner tracker ($|\eta| < 2.5$), the sum runs over all charged particles associated to the PV. Outside of the acceptance of the inner tracking detector, the sum runs over all PF candidates. In both cases, only PF candidates inside a cone of radius $R = 0.4$ around the considered neutral particle candidate contribute to α_i. The compatibility for the neutral particle candidate with originating from pileup is then quantified via

$$\text{signed}\,\chi_i^2 = \frac{(\alpha_i - \bar{\alpha}_{\mathrm{PU}})|\alpha_i - \bar{\alpha}_{\mathrm{PU}}|}{\mathrm{RMS}_{\mathrm{PU}}^2} , \tag{3.12}$$

where $\bar{\alpha}_{\mathrm{PU}}$ and $\mathrm{RMS}_{\mathrm{PU}}$ correspond to the median and root-mean-square of the α_i distribution estimated from charged particle candidates originating from pileup in the corresponding event. Lastly, the value of χ_i^2 is translated to a weight by means of the cumulative distribution function for the χ^2 distribution with one degree of freedom

$$w_i = F_{\chi^2,\mathrm{dof}=1}(\text{signed}\,\chi_i^2) . \tag{3.13}$$

Neutral particle candidates with $w_i < 0.01$ or $w_i p_{\mathrm{T},i} > (A + B N_{\mathrm{vtx}})$ are excluded from the calculation of $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$ in Equation 3.10, where N_{vtx} corresponds to the number of reconstructed vertices in the event and A and B are free parameters optimized to yield the best energy resolution of $p_{\mathrm{T}}^{\mathrm{miss}}$ and jets clustered using the PUPPI method as function of η.

To improve the resolution of the missing transverse momentum, the improved energy response and resolution of the jet energy calibration is exploited. The raw $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$ is corrected for the difference in the corrected and uncorrected transverse momentum of jets exceeding transverse momenta of 15 GeV according to

$$\vec{p}_{\mathrm{T}}^{\,\mathrm{miss,\,corr}} = \vec{p}_{\mathrm{T}}^{\,\mathrm{miss,\,raw}} - \sum_{j=1}^{N_{\mathrm{jets}}} (\vec{p}_{\mathrm{T},j}^{\,\mathrm{corr}} - \vec{p}_{\mathrm{T},j}) . \tag{3.14}$$

Events with anomalously large $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$ may arise from noise in the readout electronics, dead cells in the calorimeters, or interactions of muons from the beam halo in the calorimeters [98]. The beam halo consists of muons produced in proton collisions inside a single bunch inside the beam pipe but traversing the detector parallel to the beam pipe. Multiple filters are used to identify events with anomalously large reconstructed $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$ and reject them from being used in analyses. These filters exhibit an efficiency of 85–90 % and a mistag rate below 0.1 % [98].

b Quark-Initiated Jets

Identification of b jets exploits differences in the structure of jets initiated by b quarks, lighter flavor quarks, or gluons. These differences arise through decay properties of hadrons involving b quarks and different fragmentation and hadronization properties in b jets [101]. The mean lifetime of hadrons involving b quarks amounts to approximately $c\tau \approx 500\,\mu\mathrm{m}$ [35]. Thus, these hadrons may travel a few millimeters before decaying and forming a displaced secondary vertex from which additional tracks emerge. The higher mass and harder fragmentation process of hadrons involving b quarks leads to a larger p_{T} of their decay products with respect to the jet axis. In around 20 % of the cases, hadrons involving b quarks decay weakly with charged leptons in the decay chain and thus give rise to non-isolated leptons inside jets.

The DeepJet algorithm [102] is used to distinguish b jets from jets initiated from quarks of lighter flavor or gluons. It combines input quantities from up to 25 charged and 25 neutral PF candidates inside the considered jet, up to four reconstructed secondary vertices, and global jet and event properties in a recurrent neural network. Separate convolutional layers are used to generate features from the three different low-level object types before passing them on to an independent layer of recurrent nodes for each object type. Lastly, the outputs of the three independent subnetworks are combined with the high-level inputs in a dense deep neural network to classify the jets in six different categories. Three of these categories correspond to jets identified to originate from b quarks with the additional distinction in b jets, jets containing two hadrons involving b quarks and leptonic b hadron decays inside the jet. In the presented analysis, these three categories are combined to result in an inclusive b jet identification estimate. A more detailed description of the algorithm and its performance is given in [102].

CHAPTER 4

Search for Additional Neutral Higgs Bosons

The studies described in this thesis have been performed as contributions to the search for additional neutral Higgs bosons ϕ in $\tau\tau$ final states published by the CMS Collaboration [30]. In this chapter, the structure of the performed search is described. The additional neutral Higgs bosons are assumed to manifest themselves as resonant excesses above the background expectation in mass spectra of pairs of tau leptons. The full data set recorded by the CMS experiment in the LHC Run-2 data-taking campaign in the years 2016, 2017, and 2018 is analyzed. It corresponds to an integrated luminosity of $138\,\mathrm{fb}^{-1}$ [103–105].

For the presented interpretations of the search results, neutral Higgs boson decays to down-type fermions are of special importance since the involved couplings are enhanced for large values of $\tan\beta$. This is the case for the MSSM as well as for generic 2HDMs of Type II. The better experimental tractability with respect to decays to quarks and the higher branching fraction with respect to decays into lighter leptons yields a dominant role in these searches to the decay in tau leptons.

As discussed in Section 3.2, tau leptons decay further in electrons, muons, or one or more hadrons, and the corresponding neutrino(s) before reaching the detector. Grouping the hadronic decay modes, the decays of both tau leptons give rise to six possible final states for $\tau\tau$ events. Sorted by their branching fractions, these final states comprise $\tau_\mathrm{h}\tau_\mathrm{h}$, in which both tau leptons decay hadronically, $\mathrm{e}\tau_\mathrm{h}$ and $\mu\tau_\mathrm{h}$ involving the decay of one τ lepton into an electron or muon, and $\mathrm{e}\mu$, ee, and $\mu\mu$, in which both tau leptons decay into electrons or muons. The branching fractions of the $\tau\tau$ system to each of these final states are shown in Figure 4.1. Out of the six possible final states, four are considered in this search. The ee and $\mu\mu$ final states are not analyzed as they provide only a negligible contribution to the sensitivity of the search because they exhibit only a small branching ratio and suffer from large backgrounds from the production of prompt lepton pairs. In total, thus, roughly 94 % of all possible $\tau\tau$ final states are analyzed.

Depending on the final state, the experimental signature of $\tau\tau$ events consists of up to two well isolated electrons or muons, up to two narrow hadronic jets originating from

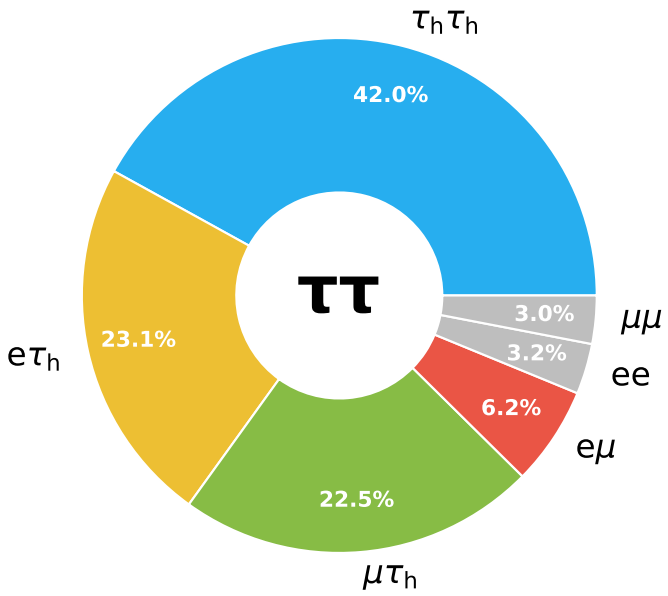

Figure 4.1: Branching fractions to final states of $\tau\tau$ events. The numbers are obtained by multiplying the respective branching fractions for tau lepton decays into electrons, muons, and hadrons. The ee and $\mu\mu$ final states not analyzed in the presented analysis are shown in gray. The branching fractions for single tau leptons are taken from [35].

the hadronic tau lepton decays, and missing transverse energy from the neutrinos. With at least two neutrinos present in the $\tau_h\tau_h$ final state and up to four neutrinos in the eμ final state, an association of the missing transverse momentum to single tau leptons is not possible, which prevents the full reconstruction of the of the individual τ leptons and complicates the estimate of the $\tau\tau$ invariant mass.

The results of the search are presented in two different interpretations:

Model-independent ϕ boson search: A search for a not further specified resonance ϕ decaying into a pair of tau leptons in addition to H(125) is performed for a set of hypothesized mass points of ϕ. For each mass point, the signal hypothesis for the given mass is compared to the prediction from all relevant SM background processes including H(125), for which properties as in the SM are assumed. The signal hypotheses are constructed directly from the signal samples, which have been simulated for the corresponding mass points. The search is performed separately for the two major production modes of additional neutral Higgs bosons in 2HDMs of Type II at the LHC, namely gluon fusion (ggϕ) and b quark associated production (bbϕ).

Exploration of dedicated benchmark scenarios: The compatibility of selected MSSM and 2HDM benchmark scenarios with the data is tested. For this purpose, a comparison between the compatibility with the data of the considered benchmark scenario and the SM is performed. This comparison is quantified in the form of statistical hypothesis tests for each benchmark scenario described in Section 2.3. In contrast to the model-independent ϕ boson search, the signal model consists of a multi-resonance structure. The exploration of the benchmark scenarios fully exploits this multi-resonance structure by incorporating the complete spectrum of $\tau\tau$ pairs reconstructed with the CMS detector in the comparison. In particular, information on the properties of h is inferred directly from the data in its mass range using the best known methods. For this, the categories used for the model-independent ϕ search are combined with a slightly adapted version of [15]. This ansatz has first been proposed in [106].

4.1 Background Processes

At the LHC, the hard scattering process occurs at length scales that can not be resolved with the detectors. The physics processes occurring in a given proton-proton collision can only be inferred from the kinematic properties of the final state particles, which are reconstructed from their interactions with the detector. Analyses thus select reconstructed final state configurations expected from the signal process. There are two possibilities how backgrounds may arrive at a configuration of final state particles identical to the signal process. The first one covers cases where the same final state particles are created in a different physics process. These backgrounds are referred to as irreducible backgrounds as there are no experimental means to discriminate between those and signal. Reducible backgrounds on the other hand arise from originally different final states that resemble the signal signature through misidentification in the reconstruction. Background sources for $\tau\tau$ events include the production of genuine tau lepton pairs, the production of prompt leptons, and misidentification of other reconstructed objects, namely leptons and jets originating from quarks or gluons. To each of these background sources multiple physics processes contribute.

The major background for tau lepton pair production is the resonant production of Z bosons or photons in Drell–Yan processes [107, 108]. Exemplary Feynman diagrams of this process are shown in Figure 4.2. Depending on the subsequent decay of the resonance, this background constitutes an irreducible (decays to tau lepton pairs) or a reducible (decays to electrons or muons ($\mathrm{Z} \to \ell\ell$)) background of the analysis. The symbol ℓ is used to refer to electrons and muons here and throughout the rest of the thesis. In the latter case, the experimental signature of the signal process may be obtained, e. g., if one of the leptons is misidentified as hadronic tau lepton decay. The irreducible background from the decay to pairs of genuine tau leptons is estimated using the τ-embedding method [109]. The method is described in more detail in Section 4.4.2. The $\ell \to \tau_\mathrm{h}$ background contribution is estimated from simulation.

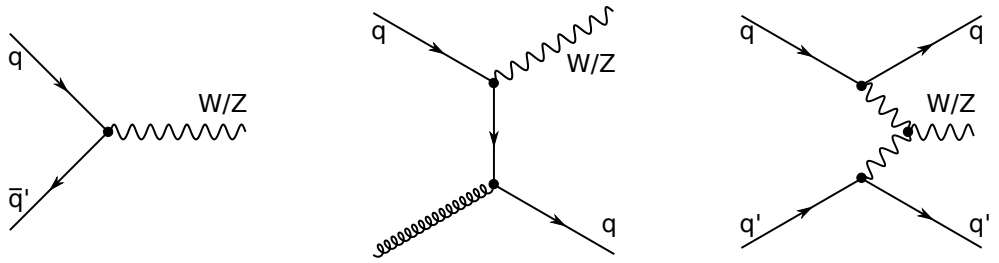

Figure 4.2: Example Feynman diagrams of W and Z boson production at hadron colliders.The subsequent decays of the produced bosons are not shown. The process shown on the right is suppressed with respect to the other processes because it involves multiple electroweak couplings.

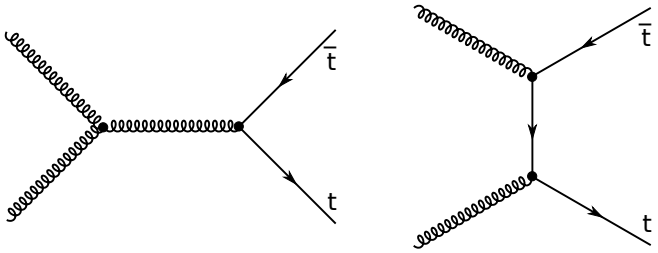

Figure 4.3: Example Feynman diagrams of the production of top quark pairs. The subsequent decays of the top quark and antiquark are not shown in the diagrams.

The production of top quark pairs ($t\bar{t}$) constitutes a further major background of the analysis. Feynman diagrams of the process are displayed in Figure 4.3. The top (anti)quark subsequently decays into a W boson and an b (anti)quark in more than 99 % of the cases [35]. Depending on the subsequent decays of the W bosons, there are three ways $t\bar{t}$ production contributes to the background. Genuine tau lepton pairs arise if both W bosons decay in tau leptons. This happens in roughly 1 % of all $t\bar{t}$ decays [35]. This background is estimated using the τ-embedding method. Secondly, one or both W bosons may decay into quarks. This process contributes to the background if at least one jet is misidentified as τ_h. The F_F-method [31, 110] further discussed in Section 4.4.3 is used to estimate this background. The third contribution arises from decays of the W bosons into electrons or muons. This contribution is sizable in the semi-leptonic final states. It becomes a major background in the eμ final state. It is estimated from simulation. The $t\bar{t}$ background is among the dominant backgrounds in events containing at least one reconstructed b jet but only plays a minor role in event selections without this requirement.

The production of W bosons in association with quark- or gluon-initiated jets (W + jets) constitutes a further background of the analysis. The Feynman diagrams of this process are similar to the ones displayed in Figure 4.2 with the Z boson replaced by the W boson. As the W boson decays into at most one lepton, this background enters the analysis as

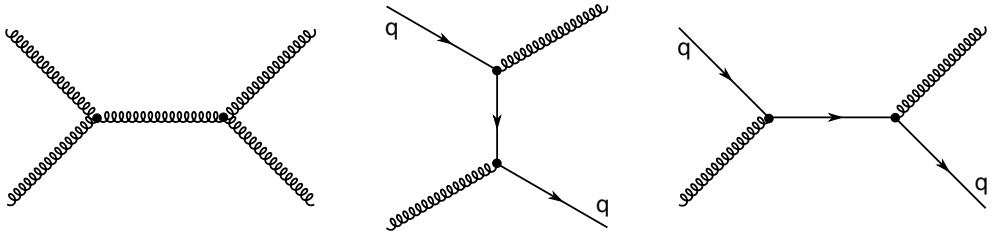

Figure 4.4: Example Feynman diagrams of the QCD multijet background.

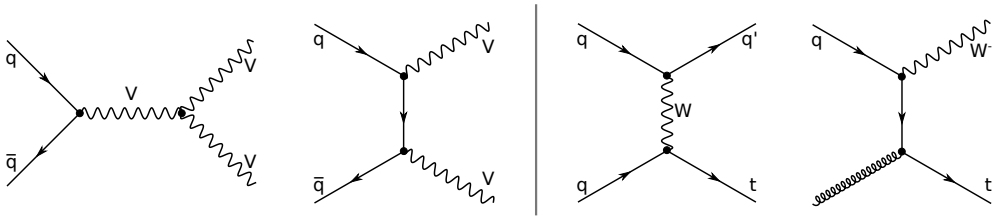

Figure 4.5: Example Feynman diagrams of diboson and single top quark production. The two diagrams on the left show diboson production, whereas the two diagrams on the right show the production of single top quarks via electroweak interactions. The second diagram from the right shows single top quark production in the t channel. The rightmost diagram displays the production of single top quarks in association with W bosons.

reducible background if a jet is misidentified as τ_h in the eτ_h and $\mu\tau_\mathrm{h}$ final states, or a light lepton in the eμ final state. The F_F-method provides an estimate of this background in the first case, whereas it is estimated from simulation in the latter.

The production of at least two gluon- or quark-initiated jets via the strong interaction in QCD multijet events represents another reducible background of the analysis. Feynman diagrams of the process are shown in Figure 4.4. These events are produced abundantly at the LHC and enter the analysis if jets are misidentified as τ_h, electron, or muon. This process is the major source of the jet $\rightarrow \tau_\mathrm{h}$ background in the $\tau_\mathrm{h}\tau_\mathrm{h}$ final state. Its contribution to the analysis is estimated using the F_F-method in final states involving hadronically decaying tau leptons and a similar method described in Section 4.4.4 in the eμ final state.

A minor background of the analysis arises from the production of pairs of massive vector bosons (diboson) or single top quarks or antiquarks (single t quark) in electroweak processes. Exemplary Feynman diagrams of the processes are shown in Figure 4.5. Depending on the bosons involved in the process and their decays, final states consisting of up to four genuine leptons are possible, while only the subset of processes with at most two genuine leptons in the final is selected. The final states arising from single t quark production show similarities to the W + jets background in the case of t-channel production or the diboson and $\mathrm{t\bar{t}}$ backgrounds when produced in association with a W boson. The diboson and single t quark backgrounds may enter the analysis in the same

three ways as discussed for the $t\bar{t}$ background and are estimated using the same methods.

The production of H(125) and subsequent decays in tau leptons poses a minor background for signal in the mass range of H(125). The production modes are the same as for the additional neutral Higgs bosons described in more detail in Chapter 2 but the relevance of the different production modes differs. In the mass range around 125.38 GeV there are no experimental means to distinguish between signal and this background. Additionally, subsequent decays of H(125) to pairs of W bosons (H(125) $\rightarrow$ WW) constitute a minor background in the eμ final state. These backgrounds are estimated from simulation.

4.2 Event Selection

In order to reduce the contribution of backgrounds to the analysis, while still maintaining a large acceptance for signal events, selection criteria are imposed on the events collected by the CMS experiment. The event selection and categorization is based on physics object candidates reconstructed as described in Chapter 3. The events are categorized exclusively into the four analyzed final states based on the number of selected leptons present in the event.

In the $\mu\tau_h$ final state, the presence of exactly one muon and at least one τ_h candidate is required. An event has to be selected by a single muon trigger, a $\mu\tau_h$ pair trigger or a single τ_h trigger. To facilitate the combination of the efficiencies of the different triggers, all triggers involving muon selections are used in exclusive phase space regions based on the transverse momenta of the offline reconstructed muon and τ_h candidate. For events containing τ_h candidates with high transverse momenta, both the single muon and single τ_h triggers are used. The different kinematic regions the triggers are used in are illustrated in Figure 4.6.

The muon candidate is required to have a transverse momentum of $p_T > 21$ GeV and to be reconstructed within the pseudorapidity range $|\eta| < 2.1$. In case the event has been selected by a single muon trigger, the transverse momentum requirement imposed on the muon is increased to 23 GeV for the 2016 data-taking year and 25 GeV for the 2017 and 2018 data-taking years. Additionally, the muon is required to pass the medium WP of the identification algorithm and exhibit a relative isolation smaller than 0.15.

The transverse momentum of τ_h candidates is required to exceed 30 GeV. This threshold is increased to $p_T > 32$ GeV if the event is selected by the $\mu\tau_h$ pair trigger in the 2017 and 2018 data-taking years and to $p_T > 120$ GeV(180 GeV) for the 2016 (2017–2018) data-taking years if the event is selected by the single τ_h trigger. In events recorded by a trigger selecting hadronically decaying tau leptons, the τ_h candidate is required to be reconstructed within $|\eta| < 2.1$. Otherwise, this requirement is relaxed to $|\eta| < 2.3$.

Furthermore, reconstructed τ_h candidates are required to pass the Medium, Tight, and

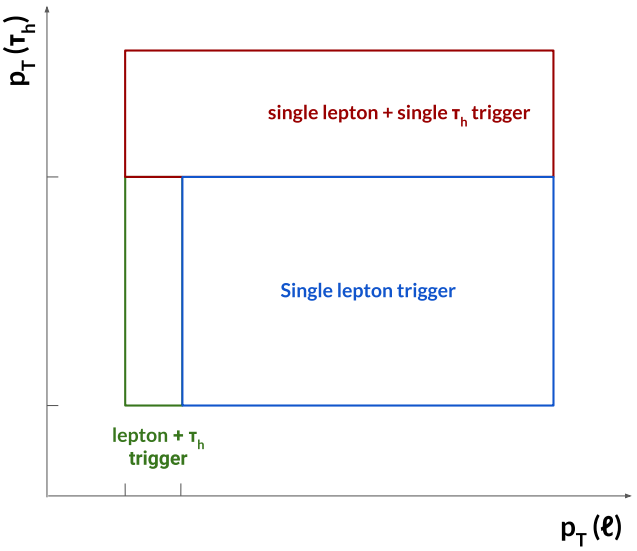

Figure 4.6: Schematic view of the kinematic regions with different trigger selections in the $\ell\tau_{\mathrm{h}}$ final states.

VVLoose WPs of the D_{jet}, D_{μ}, and D_{e} discriminants, as discussed in Section 3.2, respectively. Requirements on the absolute values of the longitudinal and transverse impact parameters, $|d_z|$ and $|d_{xy}|$, are imposed on the muon and hadronically decaying tau leptons to ensure that both tau lepton candidates originate from the PV. For the muons these requirements are $|d_z| < 0.2\,\mathrm{cm}$ and $|d_{xy}| < 0.045\,\mathrm{cm}$, respectively, while for the τ_{h} only the requirement on $|d_z|$ is imposed.

Events containing exactly one electron and at least one hadronic tau lepton decay are categorized into the $\mathrm{e}\tau_{\mathrm{h}}$ final state. These events are selected by single electron, $\mathrm{e}\tau_{\mathrm{h}}$ pair, and single τ_{h} triggers following the same strategy as in the $\mu\tau_{\mathrm{h}}$ final state with tighter requirements imposed on the p_{T} of the electron than on the muon. The transverse momentum requirement of $p_{\mathrm{T}} > 25\,\mathrm{GeV}$ on the electron candidate is increased to $p_{\mathrm{T}} > 26\,\mathrm{GeV}$ ($p_{\mathrm{T}} > 28\,\mathrm{GeV}$, $p_{\mathrm{T}} > 33\,\mathrm{GeV}$) for the 2016 (2017, 2018) data-taking periods for events selected by the single electron trigger. Electron candidates are required to be reconstructed within $|\eta| < 2.1$, to pass the 90 % efficiency WP of the MVA-based electron identification algorithm and to exhibit $I_{\mathrm{rel}} < 0.15$. Hadronically decaying tau leptons are subject to the same requirements as in the $\mu\tau_{\mathrm{h}}$ final state, except for the WPs of the D_{ℓ} discriminant that are chosen to be the VLoose and the Tight WP for D_{μ} and D_{e}, respectively. Electron candidates are subject to the same requirements on the transverse and longitudinal impact parameters as the muon candidates.

Events containing at least two hadronically decaying tau lepton candidates are selected in the $\tau_{\mathrm{h}}\tau_{\mathrm{h}}$ final state. The single τ_{h} and $\tau_{\mathrm{h}}\tau_{\mathrm{h}}$ pair triggers are employed to select the events. The τ_{h} candidates are required to have transverse momenta of $p_{\mathrm{T}} > 40\,\mathrm{GeV}$

and to be reconstructed within $|\eta| < 2.1$. The p_{T} requirement is increased to 120 GeV (180 GeV) when the tau lepton is selected by the single τ_{h} trigger in the 2016 (2017–2018) data-taking years. Furthermore, the hadronically decaying tau leptons are required to pass the Medium, VLoose, and VVLoose WPs of the D_{jet}, D_{μ}, and D_{e}, respectively. The same requirements on the longitudinal impact parameter as in the $\ell\tau_{\mathrm{h}}$ final states are imposed.

In the eμ final state, events containing a reconstructed muon and electron candidate are selected exploiting triggers selecting eμ pairs with asymmetric p_{T} thresholds. Electron and muon candidates are selected mostly based on the criteria imposed in the $\mu\tau_{\mathrm{h}}$ and eτ_{h} final states. Both candidates are required to exhibit transverse momenta of at least 15 GeV, while the threshold is increased to 24 GeV for the lepton with the higher p_{T} threshold at the trigger level. The requirement on the pseudorapidity as well as the requirement on the relative isolation of the muon are relaxed to $|\eta| < 2.4$ and $I_{\mathrm{rel}} < 0.2$.

In the final states involving τ_{h} candidates, ambiguities in the selection of the tau lepton pair can arise from the fact that more than one reconstructed τ_{h} candidate may be present in an event. In such cases, the τ_{h} candidate with the larger value of D_{jet} is chosen. In cases in which this value is the same for both candidates, the candidate with the larger transverse momentum is chosen. The candidates building the tau lepton pair are required to have an angular separation of $\Delta R > 0.5$ and to be oppositely charged in all final states. Events also have to pass the filters for spurious missing transverse momentum as described in Section 3.2.

Vetoes on additional electrons and muons are imposed in all final states to prevent the selection of the same event in multiple final states. In the $\mu\tau_{\mathrm{h}}$ and eτ_{h} final states, these vetoes are supplemented with vetoes on muon and electron pair candidates, respectively, to suppress the backgrounds from Z boson production in decays to electrons or muons. Electrons and muons considered for the veto on additional light leptons are subject to looser selection criteria than the ones used for the final state selection. Electrons are required to have transverse momenta of $p_{\mathrm{T}} > 10$ GeV, to be reconstructed within $|\eta| < 2.5$, to pass the 90 % efficiency WP of the MVA-based electron identification, to pass the $|d_z|$ and $|d_{xy}|$ requirements as described above and to exhibit a relative isolation smaller than 0.3. Selection criteria of $p_{\mathrm{T}} > 10$ GeV, $|\eta| < 2.4$, medium identification requirements, and a relative isolation below 0.3 are imposed on the muons.

For the veto of electron pairs, the selection criterion on the transverse momentum is increased to $p_{\mathrm{T}} > 15$ GeV and they are required to pass the cut-based electron identification algorithm. For muons the same p_{T} threshold as for electrons is used and the identification criterion is loosened to the Loose WP. To be considered in the lepton pair veto, the leptons are required to be oppositely charged and to exhibit an angular separation of $\Delta R > 0.15$.

4.3 Categorization and Final Discriminant

The selected events in each final state are further categorized to increase the sensitivity to signal processes while collecting the $\mathrm{t\bar{t}}$ background in a distinct region to control this background from data. The categorization of events follows the one introduced in [31]. Firstly, events are divided in No b tag and b tag categories based on the number of b jets in the event to target the major production modes $\mathrm{gg}\phi$ and $\mathrm{bb}\phi$ separately. For events to be categorized in the b tag category, the presence of at least one jet reconstructed within a pseudorapidity of $|\eta| < 2.4$ (2.5) for the 2016 (2017–2018) data-taking years and identified by the Medium WP of the DeepJet [102] algorithm is required.

In the $\tau_\mathrm{h}\tau_\mathrm{h}$ final state, no further subcategorization is applied. In the $\mathrm{e}\tau_\mathrm{h}$ and $\mu\tau_\mathrm{h}$ final states, events are further split in two categories based on the transverse mass m_T of the ℓ and $\vec{p}_\mathrm{T}^{\,\mathrm{miss}}$ system. The transverse mass is defined as

$$m_\mathrm{T}(\vec{p}_\mathrm{T}^{\,\mathrm{O}_1}, \vec{p}_\mathrm{T}^{\,\mathrm{O}_2}) = \sqrt{2 p_\mathrm{T}^{\mathrm{O}_1} p_\mathrm{T}^{\mathrm{O}_2} (1 - \cos\Delta\varphi)}\,, \tag{4.1}$$

where $\Delta\varphi$ is the separation of the considered reconstructed objects, O_1 and O_2, i. e., the lepton and $\vec{p}_\mathrm{T}^{\,\mathrm{miss}}$ in this case, in azimuthal angle. The Tight-m_T category is defined by $m_\mathrm{T} < 40\,\mathrm{GeV}$, while the Loose-$m_\mathrm{T}$ category is defined by $40 < m_\mathrm{T} < 70\,\mathrm{GeV}$. Signal events are expected to mainly contribute to the Tight-m_T category as the direction of the missing transverse momentum and the transverse momentum of the lepton are expected to coincide more than for the dominant backgrounds [31]. The Loose-m_T categories are introduced to increase the signal acceptance without increasing the number of background events in the more sensitive Tight-m_T categories and thus increase the combined sensitivity of the analysis.

Events in the $\mathrm{e}\mu$ final state are subdivided in three categories in the *No b tag* and *b tag* categories each, based on the D_ζ variable [111], defined as

$$D_\zeta = p_\zeta^{\mathrm{miss}} - 0.85 \cdot p_\zeta^{\mathrm{vis}}, \qquad p_\zeta^{\mathrm{miss}} = \vec{p}_\mathrm{T}^{\,\mathrm{miss}} \cdot \hat{\zeta}, \qquad p_\zeta^{\mathrm{vis}} = (\vec{p}_\mathrm{T}^{\,e} + \vec{p}_\mathrm{T}^{\,\mu}) \cdot \hat{\zeta}\,, \tag{4.2}$$

where $\hat{\zeta}$ is the normalized angle bisector between the electron and the muon momenta in the transverse plane. Low, Medium, and High-D_ζ categories are defined via the requirements of $-35 < D_\zeta < -10\,\mathrm{GeV}$, $-10 < D_\zeta < 30\,\mathrm{GeV}$, and $30\,\mathrm{GeV} < D_\zeta$, respectively. The Medium-D_ζ category is expected to be the most sensitive category, while the Low- and High-D_ζ categories are used to increase the signal acceptance. A dedicated control region to constrain the background arising from the $\mathrm{t\bar{t}}$ background is defined in the $\mathrm{e}\mu$ final state by requiring at least one b jet and $D_\zeta < -35\,\mathrm{GeV}$. The categorization is depicted in Figure 4.7.

The distributions of the D_ζ and m_T variables used for the further categorization after a maximum likelihood fit to the data are shown in Figure 4.8. The data is indicated by black markers. In the upper panels, the filled templates represent the contributions from

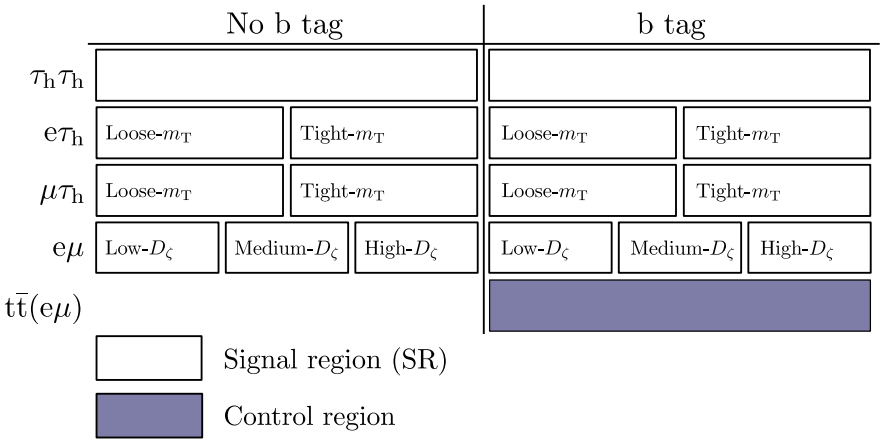

Figure 4.7: Categorization targeting the search for an additional heavy neutral resonance ϕ. In final states, in which the global No b tag and b tag categories are further categorized, the labels of the further categories are given in the corresponding regions. Taken and modified from [30].

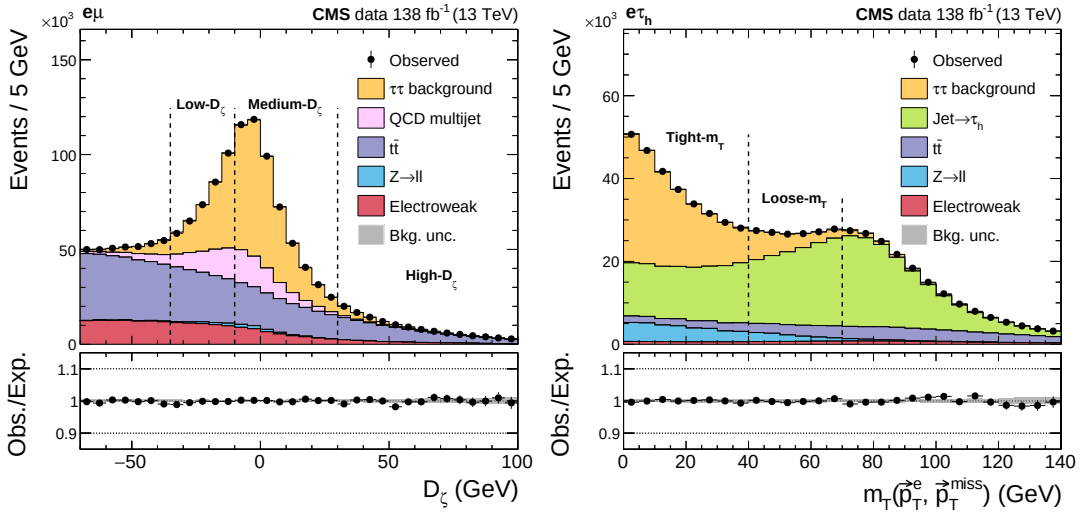

Figure 4.8: Distributions of the D_ζ and m_T variables used for the further categorization in the $e\mu$ and $e\tau_h$ final states. On the left-hand side, the distribution of the D_ζ variable is shown in the $e\mu$ final state. On the right-hand side, the distribution of the $m_T(\vec{p}_T^{\,e}, \vec{p}_T^{\,miss})$ variable in the $e\tau_h$ final state is shown. The requirements for the different subcategories are indicated with vertical dashed lines. The distributions are shown after a fit of the statistical model to the data.

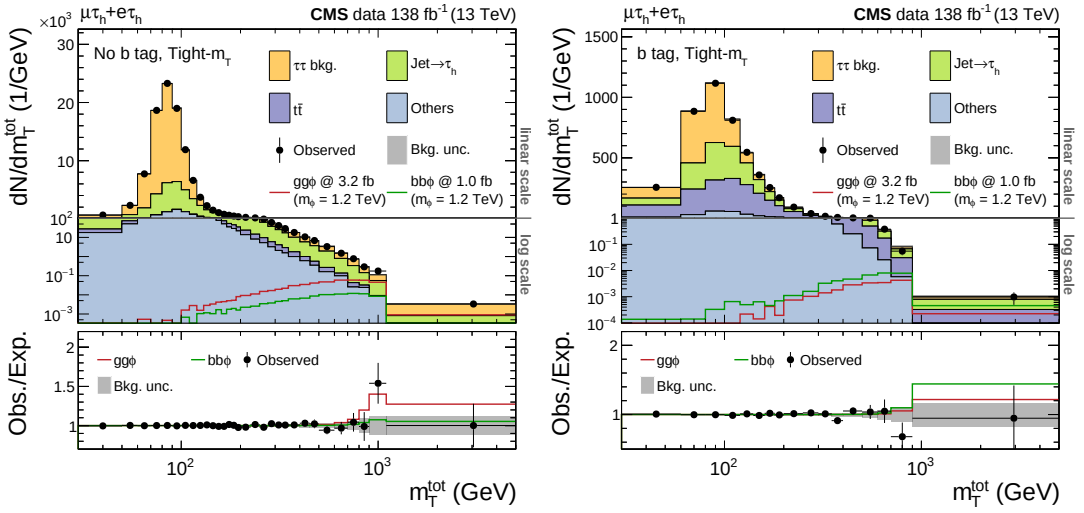

Figure 4.9: Distributions of the total transverse mass $m_{\mathrm{T}}^{\mathrm{tot}}$ in the No b tag Tight-m_{T} (left) and b tag Tight-m_{T} (right) categories. The distributions are shown combined over the three data-taking years and the $\mu\tau_{\mathrm{h}}$ and $\mathrm{e}\tau_{\mathrm{h}}$ final states after the fit of the background model to the data. The figures are modified versions of figures published in [30].

the different background sources to the distribution. The gray shaded area in all panels represents the background uncertainty after the fit to the data. In the lower panels, the ratio of the data over the expectation from all backgrounds is shown. The requirements imposed on the variables for further categorization are indicated as black dashed vertical lines in the upper panels.

The mass of the $\tau\tau$ system is a natural choice to discriminate signal from background processes. A complete reconstruction of the $\tau\tau$ system is compromised by the presence of at least two neutrinos from the decays of the tau leptons. Estimators for the mass of the $\tau\tau$ system or of variables correlated with the mass of the $\tau\tau$ system are used instead. In this analysis, the the total transverse mass defined as

$$m_{\mathrm{T}}^{\mathrm{tot}} = \sqrt{m_{\mathrm{T}}^2(\vec{p}_{\mathrm{T}}^{\,\tau_1}, \vec{p}_{\mathrm{T}}^{\,\tau_2}) + m_{\mathrm{T}}^2(\vec{p}_{\mathrm{T}}^{\,\tau_1}, \vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}) + m_{\mathrm{T}}^2(\vec{p}_{\mathrm{T}}^{\,\tau_2}, \vec{p}_{\mathrm{T}}^{\,\mathrm{miss}})} \tag{4.3}$$

is used and its distributions enter the statistical inference described in Chapter 5. The total transverse mass has first been used as discriminating variable in BSM Higgs boson searches in $\tau\tau$ final states in [112]. Exemplary distributions of the total transverse mass in the $\mu\tau_{\mathrm{h}}$ and $\mathrm{e}\tau_{\mathrm{h}}$ final states are shown in Figure 4.9, the full set of distributions in each final state is shown in Appendix A. The advantage of the use of $m_{\mathrm{T}}^{\mathrm{tot}}$ as discriminating variable over a likelihood-based mass estimate $m_{\tau\tau}$ [113] is that it tends to estimate lower masses for QCD multijet backgrounds [112]. This leads to an improved signal-to-background ratio in the tail of the distribution and an increased sensitivity of the analysis on high-mass signals.

In the MSSM and the CP-conserving 2HDM, the Higgs sector comprises three Higgs bosons, a light Higgs boson, h assumed to be equivalent to H(125) in most cases and two additional neutral heavy Higgs bosons. The categorization described above is optimized to increase the sensitivity of the analysis to the additional heavy Higgs bosons. The hypothesized signal models, however, also predict differences in the couplings of h relative to H(125). To increase the sensitivity to these differences, the categorization described above is complemented with the categorization from the analysis of the coupling of the observed Higgs boson to tau leptons [15]. This categorization is based on the output score of a neural net (NN) trained to separately identify H(125) in gluon fusion and VBF production. The categorization is composed of a single signal and multiple background categories per final state. In the signal category, a two-dimensional discriminant built from the output scores of the NN for the two analyzed production modes is used. In the background categories, the output score for the given background is used directly. To ensure the orthogonality between both sets of categories, a split is introduced based on $m_{\tau\tau}$, as introduced in [113]. The categories from [15] are used for $m_{\tau\tau} < 250\,\text{GeV}$. Since the event selection in [15] excludes events containing b jets, the global b tag categories described above do not need to be modified. The global No b tag categories are modified by the requirement of $m_{\tau\tau} > 250\,\text{GeV}$.

4.4 Background and Signal Estimation Methods

To extract information about the signal processes from the collected data, the expected signal and background processes are modeled using three different methods. The major backgrounds are estimated from data, namely by the τ-embedding [109] and the F_{F}-method [31, 110]. The latter only applies for the eτ_{h}, $\mu\tau_{\text{h}}$ and $\tau_{\text{h}}\tau_{\text{h}}$ final states, which involve hadronic tau lepton decays. In the eμ final state, a method similar to the F_{F}-method is used. The remaining background processes are modeled from simulation of the corresponding processes from first principles. The background composition in each final state considered in the analysis is shown in Figure 4.10. Depending on the analyzed final state, the fraction of backgrounds estimated from data varies. In the $\tau_{\text{h}}\tau_{\text{h}}$ final state, more than 95 % of the backgrounds are estimated from data. This fraction decreases to roughly 90 % and 83 % in the $\mu\tau_{\text{h}}$ and eτ_{h} final states and is lowest in the eμ final state where still slightly more than half of the background contributions are estimated from data.

4.4.1 Monte Carlo Simulation

The simulation of physics processes at hadron colliders involves multiple scales at which specific processes occur. The asymptotic freedom of QCD allows the generation of events from matrix elements at a fixed order in perturbation theory at large momentum transfers. This hard process is generated using either `MADGRAPH5_aMC@NLO` 2.2.2 (2.2.4, 2017–2018) [114] or `POWHEG` v2 [115–117] at LO or NLO precision depending on the simulated process as detailed below. On the other hand, the final state of an event can only consist of leptons and stable color-neutral hadrons due to the confining property of QCD at lower

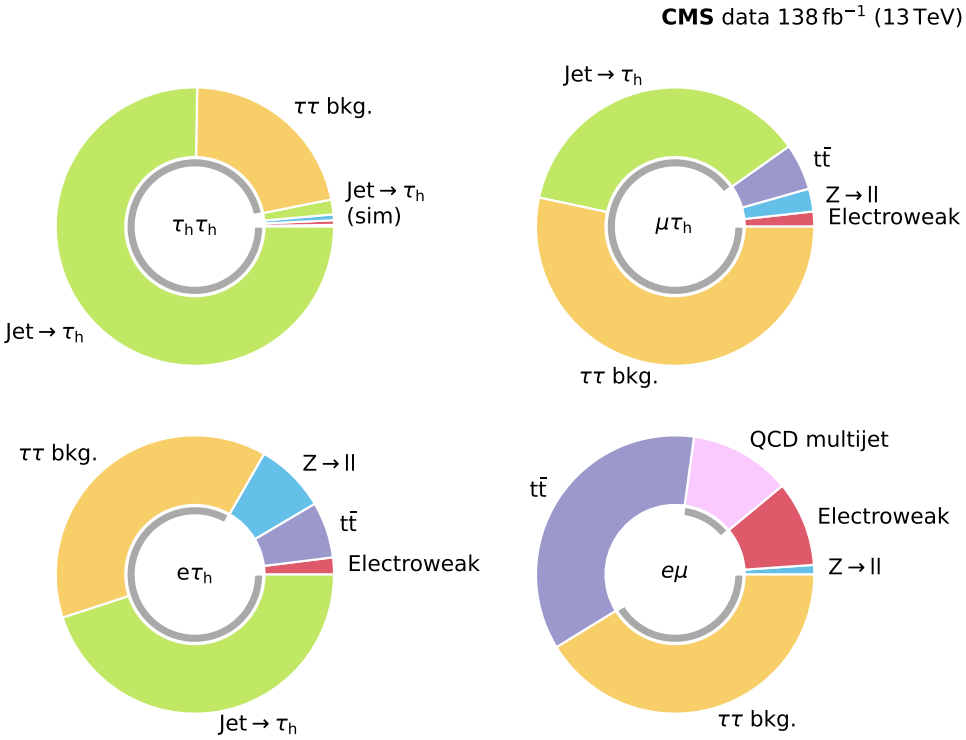

Figure 4.10: Composition of background events in each final state. The background composition is derived for the final event selection before any fit to the data is performed. The compositions in the $\tau_h\tau_h$, $\mu\tau_h$, $e\tau_h$, and $e\mu$ final states are shown clockwise starting with the $\tau_h\tau_h$ final state in the top left corner. The inner gray wedges indicate backgrounds estimated from data. The fraction of backgrounds estimated from data ranges from around 97 % in the $\tau_h\tau_h$ final state to 53 % in the $e\mu$ final state. In the $\mu\tau_h$ and $e\tau_h$ final state 90 and 83 % of the backgrounds are estimated from data.

energies. The hadronization of the final state partons takes place at energies significantly lower than the energy of the hard scattering process and is governed by phenomenological models. The transition from the large energy scale of the simulated event to the lower energies of hadronization is modeled through sequential splittings of partons into two partons with smaller energies, the parton shower. The parton shower, hadronization, and the underlying event, additional soft collisions of further partons inside the colliding protons, are modeled using `PYTHIA` 8.230 [118]. In case of the events generated by `MADGRAPH5_aMC@NLO`, a matching of the contributions generated from the matrix element and the parton shower is performed using the MLM [119] (FxFx [120]) matching schemes at LO (NLO). The NNPDF3.0 [121] (2016) and NNPDF3.1 [122] (2017– 2018) (PDFs) are used to model the content of the protons. The underlying event is modeled using the CUETP8M1 [123] and CP5 [124] `PYTHIA` tunes for the 2016 and 2017–2018 data-taking periods, respectively. The stable particles contained in the final simulated event are then interfaced to the Geant4 [125] detector simulation toolkit, which simulates the

interactions of the particles with the detector up to the level of electronic signals. The simulated event record is then present in the same format as the data recorded with the CMS experiment and is reconstructed using the same algorithms.

The same set of simulated background data sets as in [15] is used. The $t\bar{t}$ background is simulated at NLO QCD precision with POWHEG [126]. Before comparing the predictions to data, the obtained distributions are reweighted to the predictions of the cross sections in NNLO precision in QCD. The backgrounds arising from diboson and single t quark production are simulated with MADGRAPH5_aMC@NLO and POWHEG [127, 128] respectively at NLO precision in QCD. The distributions of diboson processes and single t quark production in the t channel exchange are reweighted to calculations at NLO precision in QCD, while the single t quark production in association with a W boson is reweighted to NNLO precision. The kinematic distributions of backgrounds arising from the production of Z and W bosons in association with jets are simulated at LO in QCD using MADGRAPH5_aMC@NLO. Inclusive predictions are combined with predictions binned in the number of jets produced in association with the corresponding electroweak boson. The obtained distributions are reweighted to the cross section predictions calculated at NNLO QCD. The production of the SM Higgs boson in gluon fusion [129], VBF [130] and VH [131] production is simulated using POWHEG. The subsequent decays of the Higgs boson in tau leptons or pairs of W bosons is simulated together with the parton shower in PYTHIA.

To investigate signals in a large mass range, data sets for 31 mass hypotheses in the range from 60–3500 GeV have been simulated in the gluon fusion and b quark associated production modes using POWHEG [129, 132]. The PDF4LHC15 [59] PDF set is used to model the proton contents at NLO precision in α_s for both production modes. For $bb\phi$ production the four flavor scheme is adopted in the used PDF set [133]. The predictions for the signal processes have been validated in the scope of [133]. The construction of the specific signal models from the simulated samples for the different interpretations of the data is described further in Section 5.3.

4.4.2 The Embedding Method

The embedding method [109] is a method to model the backgrounds arising from genuine tau lepton pairs from data. It exploits lepton universality to form hybrid events composed of selected $\mu\mu$ events and simulated tau lepton decays. The energy deposits of the muons in the recorded events are replaced by energy deposits of simulated tau lepton decays. Advantages of the use of the embedding method with respect to the full event simulation include a better description of the data and reduced statistical fluctuations of the resulting data set due to an increased number of available events. The improved data-expectation agreement originates from the direct use of pileup, the underlying event, and the fragmentation of jets from data. These event components are usually difficult to model in simulation [109].

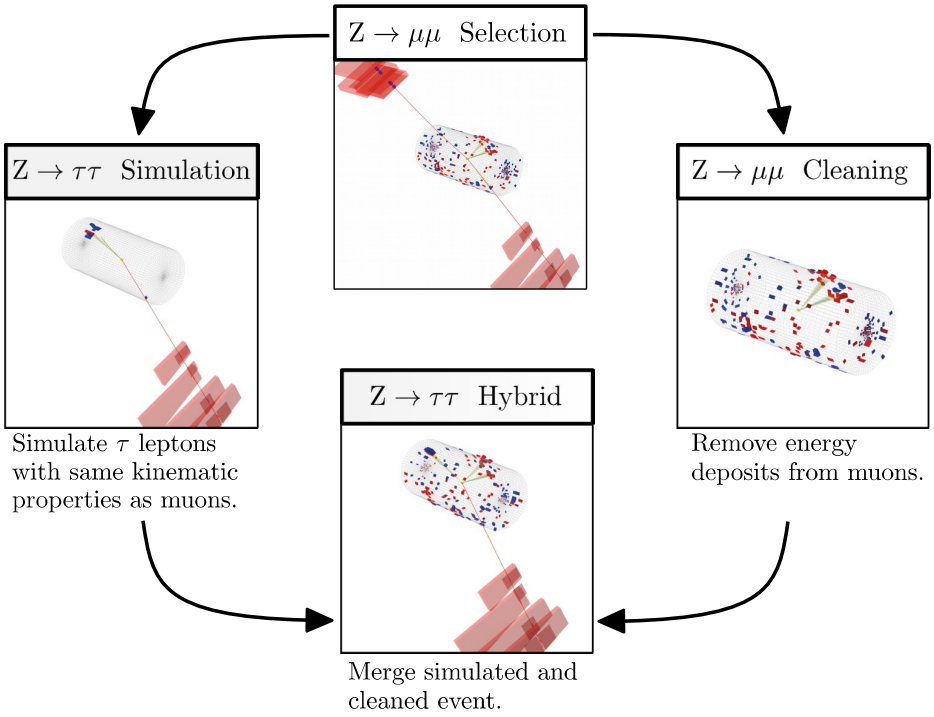

Figure 4.11: Sketch of the procedure of the embedding method. First, $\mu\mu$ events are selected in data (upper box). Afterwards, the reconstructed muons are removed from the event (right box) and tau lepton decays are simulated with the kinematic properties of the reconstructed muons (left box). Lastly, the cleaned $\mu\mu$ event is merged with the simulated tau leptons to get the final hybrid event (lower box). Taken from [109].

The procedure of the embedding method is sketched in Figure 4.11. It is performed in four steps, i. e., the selection, cleaning, simulation, and merging steps.

The selection of an observed $\mu\mu$ data set constitutes the first step of the embedding procedure. The selection requirements are chosen to guarantee a high purity of the selected sample with minimal bias on the resulting background prediction introduced by the selection. The offline selection is applied to events selected with $\mu\mu$ pair triggers with asymmetric p_{T} thresholds and a loose requirement imposed on the invariant mass of the muon pair. It comprises the selection of a pair of oppositely charged reconstructed global muons exceeding transverse momentum requirements of 17 (8) GeV imposed on the leading (sub-leading) muon candidate and an invariant mass of the $\mu\mu$ system above 20 GeV. In case of ambiguities in the selection of the muon pair, the pair with an invariant mass closest to the mass of the Z boson is chosen. Inclusively more than 97 % of the selected events originate from a Z boson decay to a pair of muons [109]. With the requirement of the presence of at least one b jet in the event this fraction reduces to 69 %, while the

fraction of selected $\mathrm{t\bar{t}}$ events increases to 26 %. The fraction of diboson and single t quark events is significantly lower in both selections, i. e., below 1 % inclusively and around 2 % in events with b jets.

During the cleaning step, hits in the tracking and muon systems associated to the reconstructed track of the global muon are removed from the event record. Energy deposits in the calorimeter cells, which are crossed by the reconstructed track of the muon interpolated to the calorimeter, are removed as well [134]. This procedure may lead to additionally reconstructed photons or neutral hadrons if a larger calorimeter cluster is split into smaller clusters when removing calorimeter cells. In general, these additionally reconstructed objects exhibit only low energies and are poorly reconstructed and therefore pose no problem to the general procedure [109].

In the simulation step, tau lepton decays are simulated starting from the kinematic properties of the selected muons in an empty detector using `PYTHIA`. To account for the difference in mass between muons and tau leptons, the four-momenta of the muons are boosted into their center-of-mass frame. Their energy is then set to half of the reconstructed mass and the magnitude of the three-momentum set according to the energy momentum relation before the four-momenta are boosted back into the laboratory frame. The tau lepton decays are then simulated using these modified momenta. Separate data sets for each analyzed final state of $\tau\tau$ decays are created by enforcing the decays of the tau leptons. These data sets are commonly referred to as τ-embedded data sets in the following. Additional data sets are also created by replacing the reconstructed muons by simulated muons or electrons. These samples are referred to as μ- and e-embedded data sets. They are used for the validation of the method and to derive corrections for the reconstructed objects.

In the final step, the selected $\mu\mu$ events, with hits and energy deposits from the selected muons removed from the event record, are merged with the hits and energy deposits from the simulated tau lepton decays. The merging is performed at an early stage of the reconstruction procedures at the level of reconstructed tracks, calorimeter clusters and muons. The remaining steps of the reconstruction procedure, in particular the PF algorithm, are then performed using the merged reconstructed objects.

To use the embedded data sets in an analysis, two effects have to be corrected first [109]. Firstly, the simulation of the interactions of particles with the detector is a CPU intensive task. The neutrinos present in the decay of tau leptons reduce the transverse momentum of the visible decay products with respect to the transverse momentum of the selected muons. Additionally, in physics analyses stricter selection requirements are imposed on the kinematic properties of the visible decay products in some final states due to requirements of the triggers used, e. g., in the $\tau_\mathrm{h}\tau_\mathrm{h}$ final state, where a minimal threshold of 35 GeV is imposed on the p_T of the reconstructed τ_h candidate. To prevent the costly simulation of the detector response for events that will not be selected by any analysis, kinematic requirements are imposed on the decay products of the tau leptons before

entering the detector simulation. To increase the efficiency of the additional requirements 1000 simulations of the tau lepton decays are performed and the last successful trial is passed on to the detector simulation. To obtain the correct normalization of the background prediction for the analysis of the data, entering events are weighted by the fraction of successful trials multiplied by the branching fraction of the decays under consideration.

Lastly, the trigger selections, offline reconstruction and selection algorithms applied to the $\mu\mu$ events in the selection step are not fully efficient. To obtain unbiased kinematic distributions and the correct normalization from the embedded data sets, the effect of these efficiencies must be corrected in the selected data sets. The necessary corrections are measured using a tag-and-probe technique as a function of the transverse momentum and pseudorapidity of the muons following the procedure outlined in Section 4.5. The efficiency of the trigger selections amounts to roughly 80 %, while the reconstruction and selection efficiency generally exceeds 95 % [109].

4.4.3 The F_{F}-Method

The F_{F}-method provides an estimate of the backgrounds originating from jets misidentified as hadronically decaying tau leptons from data. The method has first been introduced in analyses targeting $\tau\tau$ final states involving hadronically decaying tau leptons by the CMS Collaboration in [31, 110]. A modified version of the method adapted to the selections and categorizations of the presented analysis is used [30]. The distribution of backgrounds from jets misidentified as hadronically decaying tau leptons in the signal region (SR) is derived from a sideband region (application region, AR) selecting non-isolated hadronically decaying tau leptons. The AR is defined by requiring the τ_{h} candidates to pass the VVVLoose WP of the DeepTau D_{jet} discriminant while rejecting isolated τ_{h} candidates to keep the AR orthogonal to the SR. The observed jet $\rightarrow \tau_{\mathrm{h}}$ contribution in the AR is then extrapolated to the signal region using independent extrapolation factors for the three major sources constituting the background. The background contribution in the SR is thus given by

$$N_{\mathrm{SR}} = N_{\mathrm{AR}} \cdot F_{\mathrm{F}}\,, \tag{4.4}$$

where N_{SR} corresponds to the predicted number of events in the SR and N_{AR} corresponds to the number of observed events in the AR after subtracting the expected contributions from simulated events or from the embedding method. The extrapolation factor F_{F} is defined as

$$F_{\mathrm{F}} = \sum_i f_i F_{\mathrm{F}}^i, \quad i = \mathrm{QCD,W{+}jets,t\bar{t}}\,, \tag{4.5}$$

depending on the final state. It is obtained from the sum of the individual extrapolation factors F_{F}^i of the three sources of jets misidentified as hadronically decaying tau leptons weighted with the expected respective share of each source, f_i, in the total jet $\rightarrow \tau_{\mathrm{h}}$ background. The f_i are derived separately for each analysis category as function of the

discriminating variable $m_{\mathrm{T}}^{\mathrm{tot}}$ in the AR [30].

In the $\tau_{\mathrm{h}}\tau_{\mathrm{h}}$ final state, QCD multijet events constitute the dominant contribution to the background arising from jets misidentified as τ_{h} candidates. Here, only the extrapolation factor for QCD multijet events is used to estimate the jet $\rightarrow \tau_{\mathrm{h}}$ background. As the AR requirement is only imposed on the leading τ_{h} candidate in the event, the F_{F}-method covers only the background arising from misidentification of the leading τ_{h} candidate. The background arising from misidentification of the sub-leading tau lepton, in events in which the leading tau lepton is a genuine tau lepton, is estimated from simulation. These events originate from the $\mathrm{t\bar{t}}$ and $\mathrm{W}+\mathrm{jets}$ backgrounds as in the background from QCD multijet production both reconstructed tau leptons are misidentified. It contributes a fraction of slightly more than 2 % to the total jet $\rightarrow \tau_{\mathrm{h}}$ background.

The individual extrapolation factors are measured in dedicated determination regions (DR_i) for each of the three sources of misidentification [30]. They are measured as the ratio of the number of isolated over non-isolated τ_{h} candidates after subtracting residual contributions from other processes estimated from simulation and τ-embedded events. The extrapolation factors for the jet $\rightarrow \tau_{\mathrm{h}}$ background from QCD multijet production are determined from events containing tau lepton pairs with charges of same sign. In the $\tau_{\mathrm{h}}\tau_{\mathrm{h}}$ final state, this requirement leads to a sufficiently pure determination region, while in the semi-leptonic final states a further requirement of $I_{\mathrm{rel}} > 0.05$ is imposed on the relative isolation of the lepton. The determination region for the measurement of the extrapolation factors for the $\mathrm{W}+\mathrm{jets}$ background is defined by requiring the transverse mass of the lepton to be larger than $m_{\mathrm{T}} > 70\,\mathrm{GeV}$ and the absence of b jets in the event. The extrapolation factors for the background arising from top quark pair production are estimated from simulation. This backgrounds constitutes less than 3 % of the total expected jet $\rightarrow \tau_{\mathrm{h}}$ background in the No b tag categories while its fraction increases up to 65 % in the b tag categories.

The extrapolation factors are measured as a function of the p_{T} of the (leading) τ_{h} in the $\mathrm{e}\tau_{\mathrm{h}}$ and $\mu\tau_{\mathrm{h}}$ ($\tau_{\mathrm{h}}\tau_{\mathrm{h}}$) final states. The measurement is performed in categories defined by the number of jets with $p_{\mathrm{T}} > 20\,\mathrm{GeV}$ and $|\eta| < 2.1$ as well as the ratio of the transverse momentum of the misidentified jet over the p_{T} of the hadronically decaying tau leptons, $p_{\mathrm{T}}^{\mathrm{jet}}/p_{\mathrm{T}}^{\tau_{\mathrm{h}}}$. The number of jets with $p_{\mathrm{T}} > 20\,\mathrm{GeV}$ and $|\eta| < 2.1$ is used as proxy for the number of b jets to increase the available number of events used in the multijet categories in the derivation of the F_{F}^{i}.

Corrections to the extrapolation factors are applied to account for two effects. Firstly, the variables, as a function of which the extrapolation factors are measured, cover only the dominant functional dependence of the extrapolation factors. Further functional dependencies are taken into account by deriving corrections within the DR_i as function of additional variables. The corrections to the extrapolation factor for the QCD multijet background are measured as function of the angular separation of the tau leptons ΔR in the $\tau_{\mathrm{h}}\tau_{\mathrm{h}}$ final state, while they are derived as function of the projection of $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$ to

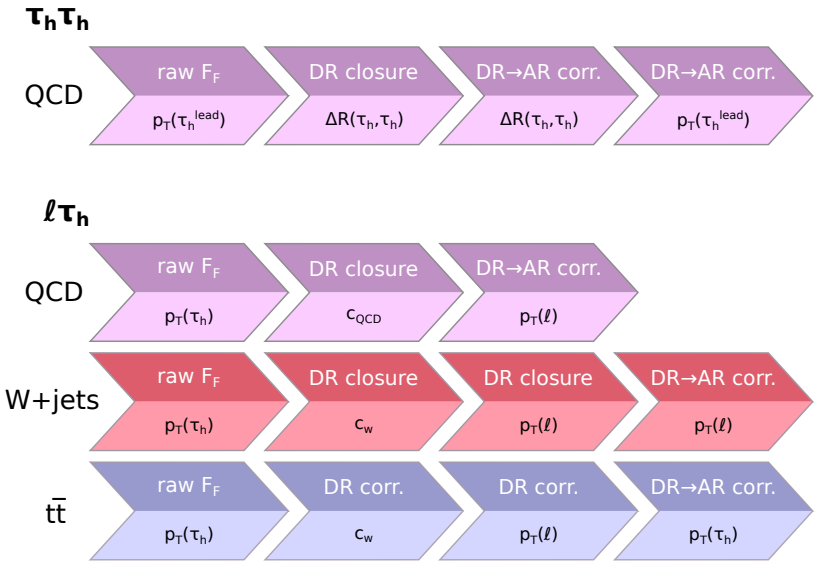

Figure 4.12: Illustration of the corrections to the raw F_{F}^{i} in the $\tau_{\mathrm{h}}\tau_{\mathrm{h}}$ and $\ell\tau_{\mathrm{h}}$ final states. The upper half of each arrow describes the type of correction applied and the lower half represents the functional dependence of the correction. The variable c_{QCD} (c_{w}) indicates the projection of the $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$ ($\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}} + \vec{p}_{\mathrm{T}}^{\,\ell}$ system) to the direction of the τ_{h} candidate relative to its transverse momentum.

the direction of the τ_{h} candidate relative to its transverse momentum in the $\ell\tau_{\mathrm{h}}$ final states. The corrections to the W + jets extrapolation factors are derived as a function of a similar variable as used for the QCD extrapolation factors but with $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$ replaced with the sum of $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$ and the momentum of the lepton. This addition approximately subtracts the genuine component of $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$ in the limit in which the W boson is produced at rest [135].

Secondly, differences between the DR_i, the extrapolation factors are measured in, and the AR, they are applied to, may lead to a bias of the background in the SR. To correct for this effect the measurement and application of the extrapolation factors is repeated in regions of the phase space orthogonal to the SR and DR_i. The difference between the predicted and observed jet $\rightarrow \tau_{\mathrm{h}}$ background in these regions is then used as correction to the extrapolation factors. These corrections are measured as function of ΔR and p_{T} of the leading τ_{h} candidate in the $\tau_{\mathrm{h}}\tau_{\mathrm{h}}$ final state and the p_{T} of the lepton in the $\ell\tau_{\mathrm{h}}$ final states. The different correction steps applied to the raw extrapolation factors to obtain the final F_{F}^{i} in the $\tau_{\mathrm{h}}\tau_{\mathrm{h}}$ and $\ell\tau_{\mathrm{h}}$ final states are illustrated in Figure 4.12.

4.4.4 QCD Multijet Background Estimation in the eμ Final State

The background from QCD multijet events in the eμ final state is estimated from data using a method similar to the F_{F}-method [15]. The AR is defined by the requirement of a selected eμ pair with same-sign charges. The necessary single extrapolation factor

F_{F} is measured in a DR defined by requiring the selected muon to be non-isolated, i. e., $0.2 < I_{\mathrm{rel}} < 0.5$. The measurement is performed as a function of the angular separation of the electron and muon in three categories based on the number of jets in the event. An extrapolation correction is measured in a region of non-isolated electrons as a function of the transverse momentum of the electron and the muon. The isolation requirement imposed on the electrons is $0.15 < {}^{\mathcal{I}^{\mathrm{e}}}/_{p_{\mathrm{T}}} < 0.5$. In events with at least one identified b jet, an additional extrapolation correction is applied as the method overestimates the background by a factor of roughly 1.4 in these events [30].

4.5 Corrections

Residual differences between the model and the data may arise from imperfections of the detector model and its response. These lead to differences in the reconstruction, efficiencies of the identification and isolation algorithms, as well as energy response and resolution of reconstructed objects. Before comparing the model with the data, corrections scaling these properties in the model to the observation in control regions are applied. The control regions are used to exclude the effects of the presence of signal events in the measurement of the corrections. A separate set of corrections is applied to simulation and the τ-embedded data sets as the tau lepton decays are simulated in an otherwise empty detector in the τ-embedded data sets [109].

These corrections can be divided in three classes. The first class corrects differences in the selection efficiency of the trigger, isolation, and identification algorithms applied in the selections of electrons, muons, and hadronically decaying tau leptons. In the case of b jets, corrections on the efficiency of the identification algorithm are applied. The second class comprises corrections to reconstructed kinematic quantities, in general the reconstructed energy of the objects. Both classes of corrections are measured in sideband regions to exclude the effects of the presence of signal events in the measurement of the corrections. The third class of corrections are not applied to correct for differences in reconstructed quantities of single objects but to correct for differences of properties of a full event. All corrections applied in the analysis are briefly described in the following.

Selection Efficiency Corrections

Selection efficiency corrections are applied to electrons, muons, and hadronically decaying tau leptons. They are applied for the trigger, identification, and isolation selections and are measured as conditional probabilities based on the previously applied corrections. The selection efficiency is thus given by

$$\epsilon = \epsilon(\mathrm{trg}|\mathrm{iso}) \cdot \epsilon(\mathrm{iso}|\mathrm{ID}) \cdot \epsilon(\mathrm{ID})\,, \tag{4.6}$$

where $\epsilon(\mathrm{ID})$ is the efficiency of the identification algorithm and $\epsilon(\mathrm{iso}|\mathrm{ID})$ and $\epsilon(\mathrm{trg}|\mathrm{iso})$ are the efficiencies of the isolation and trigger algorithms given the reconstructed object has passed the identification, and identification and isolation requirements, respectively.

In general, the selection efficiency corrections are measured using a tag-and-probe technique [136] as a function of the kinematic properties of the reconstructed objects. In simulation, the selection of genuine physics objects is directly possible. In data, however, a dedicated selection has to be defined to select genuine physics objects with a high probability while selecting only few background events. One possibility to achieve this are well understood resonances decaying into pairs of leptons like the Z boson and the J/ψ. The selection of a well identified tag object together with the boundary condition of the resonance allows a clean selection of the probe candidate for the measurement of the quantity in question. To measure comparable results, the same procedure is also applied to the simulation.

The corrections to the electron and muon identification, isolation, and trigger selection efficiencies used in the presented analysis have been measured in the context of [15] following the procedure described in [91]. The measurement is performed in $\mathrm{Z} \to \mathrm{ee}$ respectively $\mathrm{Z} \to \mu\mu$ events. The measurements of the identification, isolation, and trigger selection efficiencies comprise the same selection criteria for events and tag electrons or muons, respectively. Oppositely charged lepton pairs are selected where the tag electron or muon additionally is required to pass a transverse momentum requirement of $p_\mathrm{T} > 26$ (35) GeV in the 2016 (2017–2018) data-taking years and to be selected by a trigger that is chosen not to bias the measurement of the efficiency.

The selected events are split in a pass and fail region based on the fulfillment of the selection requirement imposed on the probe lepton. The selection efficiencies are measured from a combined fit of the ee or $\mu\mu$ invariant mass distribution in the pass and fail regions. They are obtained as the ratio of the number of signal events in the pass and the fail regions obtained from the fit. The signal is modeled by the sum of two Voigtian functions with the same width, while the background is modeled by the product of an error function with a falling exponential in case of the measurement of the identification efficiency. In the case of the remaining selection efficiencies a falling exponential function is used to model the background. These measurements confirm that the measured data-to-simulation scale factors of the electron selection efficiencies deviate not more than five percent from unity.

No separate isolation and identification requirements are imposed on hadronically decaying tau leptons but the DeepTau algorithm combines isolation and identification information to identify τ_h candidates. Thus, only corrections for the three DeepTau discriminants, D_jet, D_e, and D_μ and the trigger selection efficiencies are applied in this case. The corrections of the selection efficiency of D_jet in simulation have been measured in the context of [97] in the $\mu\tau_\mathrm{h}$ final state of $\tau\tau$ events using a tag-and-probe technique. The data-to-simulation scale factors are measured directly as relative normalization differences between simulated Z boson decays and data in a template fit of the distribution of the visible $\tau\tau$ mass $m_{\tau\tau}^\mathrm{vis}$ as function of p_T or the reconstructed decay mode of the τ_h candidate. To decrease the statistical uncertainties in the high-p_T region above 40 GeV, the p_T-dependent corrections are supplemented with an additional measurement of the

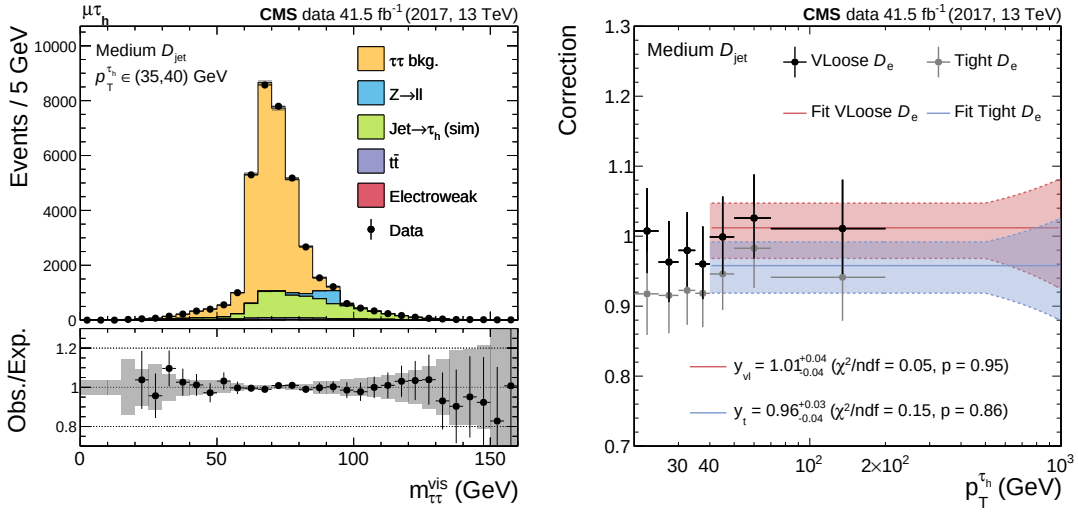

Figure 4.13: Illustration of the measurement of the corrections for the efficiency of the DeepTau D_{jet} discriminant in τ-embedded events. On the left-hand side, the distribution of $m_{\tau\tau}^{\text{vis}}$ after the fit to extract the correction is shown for the $p_{\text{T}}^{\tau_{\text{h}}}$ range from 35–40 GeV. The jet $\rightarrow \tau_{\text{h}}$ backgrounds except for QCD multijet production are estimated from simulation and shown stacked with the QCD multijet background in the plot. On the right-hand side, the measured corrections are shown for the Medium WP of the D_{jet} discriminant for both D_{e} WPs used in the presented analysis. The measured corrections in each p_{T} category are shown as black and gray markers. The corrections for $p_{\text{T}}^{\tau_{\text{h}}} > 40$ GeV obtained from the fit of a constant value to the three individual measurements for the highest $p_{\text{T}}^{\tau_{\text{h}}}$ categories are shown as solid colored lines. The uncertainties in these corrections applied in the analysis are indicated as colored filled areas.

corrections in highly virtual $\text{W}^* \rightarrow \tau\nu_\tau$ events [97], where the tau lepton decays hadronically. A constant value derived from a combined fit of the corrections measured with both methods is used in this region. The measurement of the identification efficiency as function of the decay mode is performed for τ_{h} candidates with $p_{\text{T}} > 40$ GeV. The corrections are measured separately for the one prong, the three prong, as well as the corresponding decay modes with additional neutral pions.

The measurement of the correction factors in the τ-embedded samples is performed mostly in the same way as in simulation. Only Z boson decays in the $\mu\tau_{\text{h}}$ final state are used in this case and the corrections above 40 GeV are derived by a fit of a constant value to the three measured corrections in this region. In contrast to the measurement of the corrections in simulation, no dedicated $\mu\mu$ control region is introduced and the corrections are directly obtained from the fit in the signal region. An exemplary distribution of $m_{\tau\tau}^{\text{vis}}$ after the fit to data performed to extract the correction, for $35 < p_{\text{T}}^{\tau_{\text{h}}} < 40$ GeV, is displayed in the left panels of Figure 4.13. The right panel shows the measured corrections for the Medium WP of D_{jet}. In the measurement, the contribution from W + jets and $\text{t}\bar{\text{t}}$ processes to the jet $\rightarrow \tau_{\text{h}}$ background is estimated from simulation. The contribution

from QCD multijet production is estimated from data in a sideband region requiring $\mu\tau_h$ pairs of same-sign charges and extrapolated to the signal region. During the measurement of the corrections in the τ-embedded samples and in simulation, a dependence of the corrections on the WP of D_e used to select the hadronically decaying tau leptons has been observed. Thus, a separate set of corrections has been derived for the Tight WP of D_e and applied in the $e\tau_h$ final state.

In the $\ell\tau_h$ final states, the correction factors measured as function of the p_T of the τ_h candidate are used. In the $\tau_h\tau_h$ final state, the different τ_h trigger selection efficiencies may alter the decay mode composition of selected τ_h candidates. Thus, the correction factors measured as function of the decay mode are applied in this final state.

As the track reconstruction for the decay products of the tau leptons in the τ-embedded data sets is performed in an empty detector, its efficiency is expected to be higher than in data. Thus an additional correction of 0.975 is applied for each charged track forming the τ_h candidate and 1.051 for neutral pion candidates [106]. This correction thus amounts to, e. g., $0.975 \cdot 1.051$ and $(0.975)^3$ for one prong decays with an additional neutral pion and three prong decays. In the τ-embedded data sets, no dedicated measurement of the identification efficiency corrections for tau leptons with high transverse momenta is performed as the number of available $\mu\tau_h$ events in this phase space region is too low. Therefore, the corrections measured for tau leptons with high transverse momenta in simulation scaled by the ratio of the efficiencies in simulation and τ-embedded events are used for transverse momenta above $p_T^{\tau_h} > 100\,\text{GeV}$.

The corrections on the misidentification rates of D_e and D_μ are measured in simulated $Z \to \ell\ell$ events, in which an electron or muon is misidentified as hadronically decaying tau lepton [96, 97]. While the method to extract the corrections is the same for both misidentification rates, the events are reconstructed in the $e\tau_h$ and $\mu\tau_h$ final state, respectively, .

The selected events are split in a pass and fail region based on the requirement imposed on D_ℓ. The correction factors are then obtained from a simultaneous binned maximum likelihood fit of the $m_{\tau\tau}^{\text{vis}}$ distribution of the selected $\ell\tau_h$ pair in the pass region and the normalization of all contributions in the fail region. Two parameters of interest (POI), are introduced in the maximum likelihood fit. The first POI is the misidentification rate, which represents the target of the measurement, while the second POI corresponds to the correction of the probability of the misidentified lepton to pass the D_{jet} discriminant. The second parameter leads to a correlated shift of the normalization in both the pass and fail regions. In the analysis, the product of both POIs is applied to the $Z \to \ell\ell$, $t\bar{t}$ and diboson processes where the lepton is misidentified as τ_h.

The measurement of the corrections to the misidentification rate of D_e is performed separately in the barrel and endcap regions of the ECAL. The corrections generally are larger than unity in the barrel region, smaller in the endcap regions, and range between

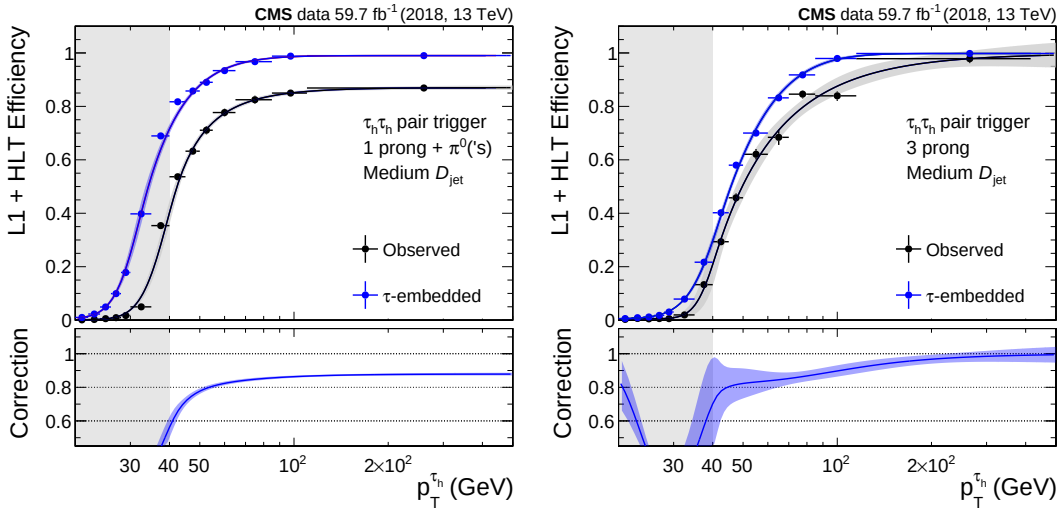

Figure 4.14: Efficiencies of a single leg of the $\tau_h\tau_h$ pair trigger as function of the offline p_T of the τ_h candidate. The observed efficiency is shown in black, the efficiency in τ-embedded data is shown in blue. The lines and the blue and black shaded areas indicate the fit result and the associated uncertainties. The gray shaded areas represent regions not used in the analysis. The lower panels show the corrections used in the analysis. The efficiencies are shown for one prong decays with additional neutral pions (left) and three prong decays (right) for the 2018 data-taking period.

values of 0.7 and 1.5 [97]. The correction factors for the misidentification rate of D_μ are measured as function of the pseudorapidity of the muon in five distinct regions. While the measurement yields comparable correction factors mostly consistent with unity for the WPs and pseudorapidity regions, the correction factor for the highest pseudorapidity region, $|\eta| > 1.7$ deviates significantly from unity [97]. The large deviation of this correction from unity originates from mismodeled track quality variables in these corners of the phase space of muons passing the discriminant albeit its large discrimination power [97].

Corrections to the efficiencies of triggers to select hadronically decaying tau leptons are performed using dedicated $\mu\tau_h$ pair monitoring triggers [96]. These monitoring triggers are constructed from the same muon selection criteria but different sets of τ_h selection criteria. The τ_h selection criteria are chosen such that they correspond to the selection criteria imposed on single τ_h candidates used in the $\mu\tau_h$, eτ_h, and $\tau_h\tau_h$ pair triggers which are employed in the analysis. Different methods are employed to measure the efficiencies in the simulated and τ-embedded samples as function of the transverse momentum and the reconstructed decay mode of the hadronically decaying tau lepton. The efficiencies and correction factors in the simulated samples are parameterized exploiting a non-parametric fit using Gaussian processes [137]. In the τ-embedded samples, the efficiencies are measured from a fit of the cumulative Crystal Ball distribution function to the binned data [96]. Examples of measured efficiencies to select a single τ_h candidate identified with the Medium WP of D_{jet} with the $\tau_h\tau_h$ pair trigger are shown in Figure 4.14.

Table 4.1: Summary of the different tag-and-probe methods used to measure the efficiency and misidentification rate corrections applied to the analysis. For each applied correction the event selection, the categories used in the measurement, the measurement method to derive the correction or the number of events in the pass and fail regions, the measured efficiency and the applied correction are listed. The estimated number of all selected signal events and the selected signal events in the pass region are indicated by N_{tot} and N_{p}, respectively. The index 'in' indicates the corresponding number prior to the maximum likelihood fit to extract the correction.

Quantity	Selection	Regions	Method	Efficiency ϵ	Correction
ℓ id, iso, trg	$Z \to \ell\ell$	Pass region; Fail region	Analytical fit	$N_{\text{p}}/N_{\text{tot}}$	$\epsilon_{\text{data}}/\epsilon_{\text{mod}}$
D_{jet} eff. (sim)	$Z \to \mu\tau_{\text{h}}$	Pass region; $\mu\mu$ CR	Template fit	-	$N_{\text{p}}/N_{\text{p,in}}$
D_{jet} eff. (emb)	$Z \to \mu\tau_{\text{h}}$	Pass region	Template fit	-	$N_{\text{p}}/N_{\text{p,in}}$
D_{ℓ} misid. rate	$Z \to \ell\ell(\ell\tau_{\text{h}})$	Pass region; Fail region	Template fit	-	$\frac{N_{\text{p}}N_{\text{tot,in}}}{N_{\text{tot}}N_{\text{p,in}}}$
τ_{h} trg eff.	$Z \to \mu\tau_{\text{h}}$	Pass region; Fail region	Event count	$N_{\text{p}}/N_{\text{tot}}$	$\epsilon_{\text{data}}/\epsilon_{\text{mod}}$

A summary of the different tag-and-probe methods used to measure the selection efficiency and misidentification rates for electron, muon and τ_{h} candidates is given in Table 4.1.

Differences in the selection efficiency and misidentification rate of the b jet identification algorithm are corrected on a jet-by-jet basis using the promotion-demotion technique [134]. In case the efficiency in data is lower than in simulation, a fraction of b jets selected as b-tagged is demoted to non b-tagged jets such that statistically the selection probability in simulation is adapted to the probability in data. Otherwise, a fraction of b jets is promoted to b-tagged jets in case they are not identified as such. A similar procedure is applied to c quark-, up, down, or strange quark-, or gluon-initiated jets. The simulation-to-data correction factors needed for this procedure are measured in a data set enriched in jets containing muons and a $\bar{\text{t}}\text{t}$ data set [101].

Energy Corrections

Corrections to the energy of reconstructed objects can be further subdivided in response and resolution corrections. They are often derived using similar selections to those used for the measurements of the selection efficiencies. For electrons corrections to the energy response and resolution are applied in the analysis, while for τ_{h} candidates only corrections to the energy response are applied. Energy resolution corrections for τ_{h} candidates as well as energy response and resolution corrections for muons are expected to be small and have only a negligible impact on the analysis. They are thus not applied.

After calibrating the electron energy as described in Section 3.2, residual differences of the response and resolution of the measured energies in simulation and data remain [91]. A shift of the electron energy response in data and an additional smearing of the energy resolution in simulation are applied to obtain a good agreement between data and simulation as the resolution in simulation is better than in data [91]. Both corrections are extracted from a $Z \to ee$ data set through a maximum likelihood fit of simulated distribution templates of the di-electron invariant mass to the data.

The response and resolution corrections are measured as a function of the electron pseudorapidity and the R_9 variable, which is defined as the energy deposited in the 3x3 ECAL crystals around the most energetic crystal of the supercluster divided by the supercluster energy [91]. It distinguishes between electrons with large and small amounts of radiated energy. The correction to the electron energy scale amounts to up to 1.5 % and is measured with an uncertainty smaller than 0.1 % (0.3 %) in the barrel (endcap) region [91]. A Gaussian energy spreading increasing the resolution by values between 0.1 and 1.5 % applied to simulation is sufficient to describe the data [91]. Corrections to the electron energy response in the τ-embedded data sets are derived from a maximum likelihood template fit of the ee invariant mass in e-embedded data [15]. The correction is measured separately in the ECAL barrel and endcap regions.

Three different corrections to the energy scale of hadronically decaying tau leptons are applied based on the origin of the reconstructed τ_h candidate. The measurement for genuine hadronically decaying tau leptons is performed with $\mu\tau_h$ events selected by imposing the same selection criteria as in the measurement of the selection efficiency correction for D_{jet} [97]. It follows the procedure described in [96]. The corrections are extracted from a maximum likelihood fit of simulated templates $m_{\tau\tau}^{vis}$ to data. The templates are generated from variations of the energy scale of the hadronically decaying tau lepton by up to 3 % around unity in steps of 0.2 percent points. The correction factors, measured separately for the four major reconstructed decay modes of hadronically decaying tau leptons, exhibit values smaller than unity but are compatible with unity within 1.5 standard deviations. The uncertainty in the measured values varies from 0.6 to 0.8 % depending on the reconstructed decay mode.

Corrections to the energy scale of electrons and muons misidentified as hadronically decaying tau leptons are measured in a similar way [15, 138]. A maximum likelihood fit of the di-lepton invariant mass with two POIs corresponding to the energy scale correction and the normalization of the $Z \to \ell\ell$ contribution is performed in $\ell\tau_h$ events enriched in leptons misidentified as τ_h candidates. The correction to the energy scale of the misidentified leptons is then obtained from a fit of a parabola to the likelihood ratios for different values of the energy scale where the normalization is fixed to the best fit value and different parameterizations for the parabolas to the left and right of the unity energy scale correction are allowed. The correction for electrons is measured separately for the barrel and endcap regions of the ECAL and for both one prong decay

modes of τ_h candidates, whereas the correction for muons is measured only for both one prong decay modes. The magnitude of the corrections for electrons misidentified as hadronically decaying tau leptons varies from unity up to 5 % [138]. The corrections increase the $e \to \tau_h$ energy response in the ECAL barrel region and decrease it in the ECAL endcap region. The correction for muons is measured to be smaller than 1 % and compatible with unity within the uncertainties. For both lepton types the measured corrections are consistent across the data-taking periods.

Event-Level Corrections

For simulated processes involving bosonic resonances, the p_T balance of the resonantly produced boson and the remaining hadronic activity in the event can be exploited to correct for residual differences in the response and resolution of the missing transverse momentum between data and simulation. The missing transverse momentum component arising from miscalibrations of the hadronic recoil is calculated using the known generator level transverse momentum of the boson. It is split into a component parallel and perpendicular to the $\vec{p}_T$ of the boson. Both components are then corrected separately using a mapping of the cumulative distribution functions from simulation to data. The cumulative distribution functions are measured in $Z \to \mu\mu$ events, in which no genuine $\vec{p}_T^{\,miss}$ is present, as a function of the number of hadronic jets in the event and the p_T of the Z boson at generator level. A more detailed description of the measurement is given in [106]. The correction is applied to the missing transverse momentum in the simulated $Z \to \ell\ell$, W + jets, $H(125) \to WW$, and $\phi \to \tau\tau$ data sets.

The simulations of the $Z \to \ell\ell$ and $t\bar{t}$ backgrounds do not describe the p_T spectra of the Z boson and top quarks [139] correctly. Thus, corrections reweighting the distributions in simulation to the observed distributions are applied to both backgrounds. The measurement of the corrections to the Z boson p_T distribution is performed as a function of the transverse momentum and the mass of the generated Z boson in a $Z \to \mu\mu$ data set [106]. The corrections applied to the $t\bar{t}$ background are derived as a function of the p_T of the top quarks and applied to the top quark and antiquark, respectively [140].

In the order of 10^{-3} events, large spurious missing transverse momenta may arise from insufficient cleaning of energy deposits in the detector in τ-embedded event samples [30]. The uncleaned energy deposits in the muon system may, e. g., be combined with energy deposits from the simulated tau leptons and increase the transverse momentum of the tau lepton or give rise to an increased $\vec{p}_T^{\,miss}$. To mitigate these effects, two modifications of the background prediction from the τ-embedding method are introduced [30]. The first modification involves a correction of the response and resolution of the missing transverse momentum not arising from neutrinos from the prediction in the τ-embedded data sets to the simulated distribution. The correction factor is derived as the ratio of the root mean square of the distribution in the simulated and τ-embedded data sets and applied separately to the x and y components of the missing transverse energy. In addition, a response correction of the genuine missing transverse energy is measured and applied to

the x and y components. The events with spurious $p_{\mathrm{T}}^{\mathrm{miss}}$ are only observed in final states involving hadronically decaying tau leptons. Thus, the measured corrections in the eμ final state are used to calibrate the corrections to the $p_{\mathrm{T}}^{\mathrm{miss}}$ distribution in data. The measured corrections range between 1–10 % [30].

The second modification introduced is the rejection of τ-embedded events where the transverse momentum of a reconstructed hadronically decaying tau lepton exceeds the visible p_{T} of the decay products of the tau lepton at generator level by a factor larger than 1.5. The effect on the normalization of the τ-embedded background from this selection criterion is removed through an increase of the normalization by the fraction of rejected events. This increase in the normalization amounts to 0.5–1 %.

Prior to data taking, the distribution of the number of additional proton-proton collisions in a single bunch crossing is not known precisely. As most of the simulated data sets are already produced beforehand, only the expected distribution can be simulated. The performance of the selection and reconstruction algorithms may vary as a function of the number of additional interactions. Thus, weights are associated to the simulation to match the distribution of the number of additional proton interactions in the data.

CHAPTER 5

Statistical Inference

In absence of observed signal events exclusion limits are derived on parameters of the signal hypothesis under consideration. In the model-independent interpretation, these parameters are the products of the cross section and branching fraction for the decay into tau leptons for the production of an additional heavy scalar boson ϕ via gluon fusion and in association with b quarks. In the exploration of the selected benchmark scenarios, the cross sections and branching fractions are not free parameters of the statistical model anymore. Instead, they are predicted by the investigated benchmark scenarios as a function of the mass of one of the additional Higgs bosons and $\tan\beta$. In the selected MSSM benchmark scenarios, the chosen mass parameter is m_{A}, whereas in the 2HDM benchmark scenarios it is m_{H}. The model-dependent exclusion contours are thus also derived as function of these free parameters.

The exclusion limits are derived as one-sided confidence intervals of the parameters in question from the distributions of the discriminating variable $m_{\mathrm{T}}^{\mathrm{tot}}$ in the signal and background categories described in Section 4.3. They are based on the profile likelihood ratio test statistic [141]. The method is summarized in this chapter together with the validation of the statistical model. For a more detailed description of the method the reader is referred to [142].

5.1 Constructing the Likelihood

The likelihood used for the statistical inference is constructed from the histogrammed distribution of the discriminating variable in all analysis categories. The observed number of events in each bin i of the histograms, n_i, is a random variable following a Poisson distribution

$$P(n_i|\mu s_i + b_i) = \frac{(\mu s_i + b_i)^{n_i}}{n_i!} \exp(-(\mu s_i + b_i)) \,. \tag{5.1}$$

In this formulation, a single signal process with an expected number of s_i events contributing to bin i is assumed. The variable b_i represents the expected number of events from the sum of the background processes. The free floating parameter μ acts as linear

signal strength modifier scaling the normalization of the signal process correlated across all bins. The likelihood used in the analysis is then obtained as the product

$$\mathcal{L}(n|\mu) = \prod_{i=1}^{N} P(n_i|\mu s_i + b_i)\,, \tag{5.2}$$

where N is the total number of bins entering the statistical inference.

The effects of uncertainties in the predicted distributions of the background and signal processes, e. g., originating from the uncertainties in the corrections described in Section 4.5 are incorporated in the form of nuisance parameters $\{\theta_j\}$. The functional dependence of the likelihood is thus given by

$$\mathcal{L}(n|\mu, \boldsymbol{\theta}) = \mathcal{L}(n|\mu \cdot s(\boldsymbol{\theta}) + b(\boldsymbol{\theta}))\,, \tag{5.3}$$

where $\boldsymbol{\theta}$ represent the vector of all nuisance parameters. Each introduced nuisance parameter modifies the likelihood in two different ways. Firstly, they alter the expected number of events from the signal or background processes, as indicated in Equation 5.3. Additionally, a constraint term is introduced in the likelihood for each nuisance parameter following a Gaussian distribution with a mean value of zero and unity variance. In particular, the complete likelihood used in the analysis is given by

$$\mathcal{L}(n|\mu, \boldsymbol{\theta}) = \prod_{i=1}^{N} P(n_i|\mu s_i(\boldsymbol{\theta}) + b_i(\boldsymbol{\theta})) \cdot \prod_{j=1} C(\tilde{\theta}_j|\theta_j)\,. \tag{5.4}$$

Here, $C(\tilde{\theta}_j|\theta_j)$ signifies the constraint term for each nuisance parameter and $\tilde{\theta}_j$ the nominal value of the nuisance parameter θ_j.

Three types of nuisance parameters with different effects on the signal and background distributions are distinguished. The first type of nuisance parameter alters only the normalization of a specific process. It corresponds to relative multiplicative uncertainties in the expected number of events and is modeled using the probability density function of the lognormal distribution [142]. Examples of nuisance parameters of this type are uncertainties in the cross section predictions for signal or background processes.

Nuisance parameters altering the shape of the distributions entering the statistical inference represent the second type of uncertainties. Possible sources of such nuisance parameters include corrections on the energy scale of reconstructed physics objects. These corrections alter the shape of all distributions derived from the object's energy. Additionally, the corrections may change the normalization of a given process [143]. To incorporate this type of nuisance parameter in the analysis, additional distributions of the discriminating variable are estimated for the up and down variations of each correction within one standard deviation.

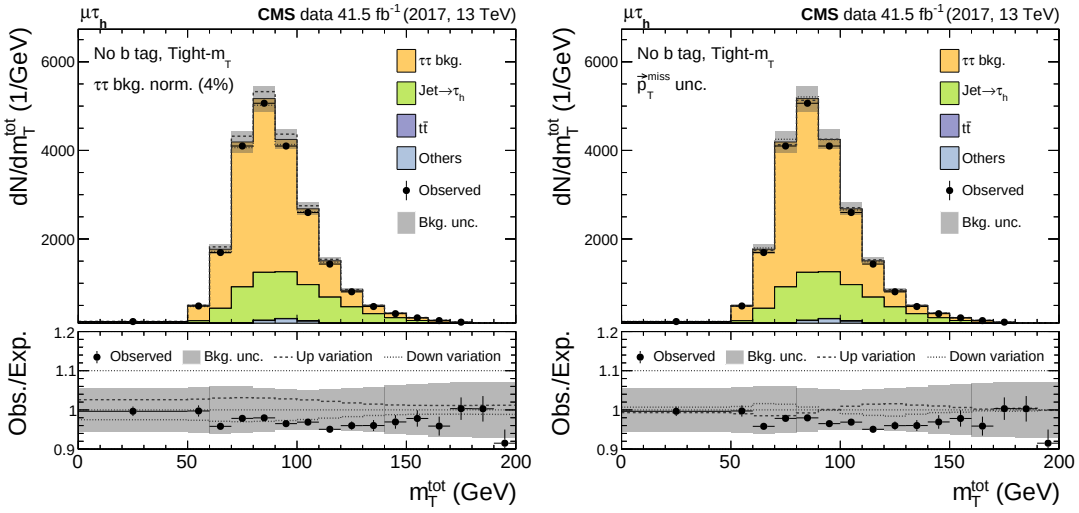

Figure 5.1: Illustration of the effect of a normalization and a shape-altering uncertainty each on the $m_{\mathrm{T}}^{\mathrm{tot}}$ distribution. The left panel shows the effect of the τ-embedding normalization uncertainty. In the right plot, the effect of the uncertainty in $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$ for the τ-embedded data set is illustrated. The nominal distributions are shown following the same conventions as in Figure 4.9. The effects of the uncertainties are indicated by the gray dashed lines corresponding to a variation of the uncertainty by one standard deviation up and down. The filled gray bands indicate the total background uncertainty. The distributions are shown prior to the fit to data.

Practically, the normalization and purely shape-altering effect of the variations are separated when constructing the likelihood but both effects are modeled using a common nuisance parameter. For the purely shape-altering component, the expected number of events in the bin is interpolated between the predictions from the up and down variations of the correction. Outside the variation range, the number of expected events is extrapolated linearly [143]. References to shape-altering uncertainties in the following always refer to the combined physical uncertainties.

The effect of a normalization and a shape-altering uncertainty on the distribution of $m_{\mathrm{T}}^{\mathrm{tot}}$ each are illustrated in Figure 5.1. While the effect of the nuisance parameter variations on the total background is displayed, they alter only the background estimated using the τ-embedding method. The left panels show the effect of a normalization uncertainty of 4 %. The right panels illustrate the effect of an uncertainty in the $p_{\mathrm{T}}^{\mathrm{miss}}$ distribution. Both uncertainties are described in more detail in the following. The normalization uncertainty also effectively alters the shape of the total background distribution because it only affects the $\tau\tau$ background and the background composition varies in different phase space regions.

The third type of nuisance parameters is introduced to account for the effect of the limited number of events entering the background and signal template distributions. To increase the sensitivity of the analysis, a fine binning of the histograms of the $m_{\mathrm{T}}^{\mathrm{tot}}$ distribution

is chosen. Thus, an additional uncertainty in the expected number of events arises from statistical fluctuations of the number of events from which the expectation in each bin is estimated. To model this additional uncertainty, the number of events may be treated as the realization of an independent Poisson distributed random variable for every process in each bin [144]. In the limit of large numbers, the parameters from each individual source are combined into a common Gaussian distributed nuisance parameter per bin [143]. Sparsely populated bins pose problems to the methods discussed above or compromise the asymptotic approximation used to extract the exclusion limits, as described in Section 5.4 [143]. Following the suggestion given in [143], bins from the highly granular sampled distributions of the discriminating variable are iteratively combined such that each bin contains at least 0.2 expected background events and the combined relative uncertainty for all backgrounds is below 90 %. The procedure starts from the bin with the lowest number of expected background events and combines bins such that the lowest number of bin combinations is required.

5.2 Uncertainty Model

In the following, the nuisance parameters that are part of the uncertainty model of the analysis and their sources are briefly discussed. Multiple nuisance parameters may be introduced for a single physical source of uncertainty. This allows independent variations of the uncertainty in different kinematic regions and the modeling of partial correlations across years or background estimation methods.

Normalization Uncertainties

An overview of the normalization uncertainties used in the analysis is given in Table 5.1. The different uncertainty sources and the magnitudes of the uncertainties are discussed in the following.

The predicted number of events from simulation are scaled to the cross section of each process and the integrated luminosity of the data set. The integrated luminosity is measured separately for each data-taking year [103–105]. The uncertainty in the measurement is propagated to the normalization of the simulated processes. The uncertainty is broken down into an uncorrelated, a fully correlated, and a part correlated across the 2017 and 2018 data-taking years. Normalization uncertainties in the cross sections of simulated background processes predicted from theoretical calculations are applied to the respective processes. The magnitude of these uncertainties ranges between 2 and 5 % depending on the process. The normalization uncertainty in the $\mathrm{t\bar{t}}$ cross section is modeled as unconstrained rate parameter allowing the constraint of the normalization from the control region in the eμ final state. The uncertainty is constrained to 2.2 % after the fit to the data.

Normalization uncertainties in the identification and selection efficiency corrections include uncertainties in the e, μ, τ_h, and b jet identification efficiencies and the selection efficiency of the eμ trigger. The uncertainties in the e and μ identification efficiency

Table 5.1: Overview of the normalization uncertainties used in the analysis. For each uncertainty, the processes it affects and its magnitude are listed. Uncertainties correlated across the data-taking years are indicated with ✓, while (✓) indicates partial correlation. The shorthand notation "rate par." denotes a freely floating nuisance parameter constrained from the control region.

Uncertainty source	Processes	Corr. across years	Magnitude
Integrated luminosity	sim	(✓)	1.2–2.5 %
Drell–Yan cross section	$Z \to \ell\ell$	✓	2 %
$t\bar{t}$ cross section	$t\bar{t}$	✓	rate par.
Diboson cross section	diboson	✓	5 %
W + jets cross section	W + jets	✓	4 %
jet $\to \tau_h$ (sim) norm.	jet $\to \tau_h$ (sim)	-	20 %
e identification	sim, emb	✓	2 %
μ identification	sim, emb	✓	2 %
τ_h identification ($D_{e(\mu)}$)	sim, emb	-	3 %, 6 %
eμ trigger	sim, emb	-	2 %
b-tagging efficiency	sim	-	< 11 %
Embedding norm.	emb	-	4 %
b tag yield	emb	-	2 %
e $\to \tau_h$ mis. rate	$Z \to \ell\ell$	-	5–23 %
$\mu \to \tau_h$ mis. rate	$Z \to \ell\ell$	-	< 6 %
eμ b tag closure (stat)	QCD	-	7 %
eμ b tag closure (syst)	QCD	✓	5 %
$h_{SM} \to \tau\tau$ BR	h_{SM}	✓	2 %
h_{SM} cross section	h_{SM}	✓	< 6 %
b quark ini. jet acceptance	bbϕ	✓	0.4–8 %
MSSM cross sections	ggϕ, bbϕ	✓	< 30 %
2HDM cross sections	ggϕ, bbϕ	✓	< 10 %

corrections amount to 2 % and are correlated across data-taking years. In the case of the uncertainties in the τ_h identification efficiency, the normalization represents only a part of the uncertainty, with the main part covered by a shape-altering uncertainty. The normalization part covers the uncertainty related to the use of different WPs of the D_μ and D_e discriminants in the measurement of the corrections and in the analysis in the $\mu\tau_h$ and $\tau_h\tau_h$ final states. The magnitude of the normalization uncertainty amounts to 3 % per selected τ_h candidate as recommended by the Tau subgroup of CMS. Additionally, an uncorrelated normalization uncertainty of 1 % per selected τ_h candidate is introduced in each final state involving hadronic tau lepton decays to account for the uncertainty in the selection efficiency of the D_μ and D_e discriminants for genuine τ_h candidates. The normalization uncertainties related to the τ_h identification efficiency are treated as uncorrelated across data-taking years.

Uncertainties in the measured efficiency corrections for the eμ trigger are included through a normalization uncertainty of 2 %. The uncertainty is treated as uncorrelated between the data-taking years and partially correlated between simulation and τ-embedded samples. Normalization uncertainties for the b jet identification efficiency are obtained from variations of the simulation-to-data correction factors in the promotion-demotion method, as described in Section 4.5. These variations mostly lead to normalization effects of the distributions in the No b tag and b tag categories and are thus implemented as normalization uncertainties.

The normalization of the background estimated from the τ-embedding method is taken directly from data after correcting for the muon selection efficiency in data. A 2 % uncertainty is assumed in this selection efficiency per muon. The final normalization uncertainty thus amounts to 4 % uncorrelated across data-taking years [109].

An additional normalization uncertainty of 2 % is applied to the τ-embedded data sets in the b tag categories to account for discrepancies in the distribution of the DeepJet discriminant observed in the μ-embedded control data set but not in the τ-embedded data sets used in the analysis [30].

Uncertainties in the measured corrections for the $e \to \tau_h$ and $\mu \to \tau_h$ misidentification rates are propagated to the normalization of the $Z \to \ell\ell$ background process. Although a shape-altering effect is possible from independent variations of the measured corrections in each pseudorapidity region, only the normalization of the m_T^{tot} distributions is expected to change from these variations. Therefore, the uncertainties in the measured corrections are translated to inclusive independent normalization uncertainties with magnitudes of 15–23 % and roughly 5 % originating from τ_h candidates reconstructed in the barrel and endcap regions for the $e \to \tau_h$ misidentification rate and less than 6 % in all pseudorapidity regions for the $\mu \to \tau_h$ misidentification rate. The uncertainties are treated uncorrelated across data-taking years as the uncertainty in the measurement is dominated by the statistical uncertainty.

Systematic and statistical uncertainties in the QCD multijet background estimation method in the eμ final state are propagated to its normalization for events with at least one identified b jet. The statistical component of the uncertainty amounts to 7 % and the systematic component to 5 % [30]. The systematic component is correlated across the data-taking years, whereas the statistical component is uncorrelated.

Uncertainties in the predictions of the production cross sections of the SM Higgs boson and branching fractions for the decay into tau leptons and W bosons are applied as normalization uncertainties to the simulated SM Higgs boson processes. The magnitudes of the uncertainties have been calculated in the scope of [39] and are taken from there. The uncertainties in the predicted branching fractions include three components, namely the theoretical uncertainty from missing higher-order corrections in the calculation as well as parametric uncertainties from variations of the quark masses and the strong coupling constant α_s within their uncertainties [145]. Two sources are distinguished for uncertainties in the production cross sections. These are a symmetrized theoretical uncertainty combined from missing higher-order and quark mass effects and a parametric uncertainty from variations of the PDF and the strong coupling constant α_s within their uncertainties [39].

Uncertainties in the acceptance of b jets from bbϕ production are applied as normalization uncertainties in the b tag and No b tag categories. Two uncertainty sources are distinguished. The first source covers uncertainties in the b jet acceptance from variations of the PDF, whereas the second comprises the linear combination of variations of the renormalization and factorization scales and the `hdamp` parameter in the `POWHEG` simulation. The magnitude of the first uncertainty source varies between 0.4 and 2.7 % [30]. In the No b tag categories, the uncertainty increases with increasing mass of the simulated Higgs boson. The uncertainty from the second source amounts to 0.6–8 % and rises for increasing Higgs boson mass in the No b tag categories as well [30].

In the model-dependent interpretations in MSSM benchmark scenarios, separate renormalization scale and PDF+α_s uncertainties in the predicted cross sections for ggϕ production and a combined uncertainty for bbϕ production are considered [63]. The uncertainties are provided together with the cross section predictions by the LHCHWG-3 [67] and range up to 30 %. In the 2HDM benchmark scenarios, the same uncertainty sources are considered for ggϕ production, and for bbϕ production only the PDF + α_s uncertainty is applied. The PDF+α_s and renormalization scale uncertainties are calculated as described in [59] and [50], respectively. The resulting uncertainties are smaller than 10 % over the full parameter space in these scenarios.

Shape-Altering Uncertainties

Uncertainties in the corrections of the trigger selection efficiencies are applied separately for each object that the trigger selects. Although the uncertainties in the efficiencies of the online selection criteria imposed on electrons, muons, and single τ_h candidates represent

normalization uncertainties, the combination of different triggers based on kinematic regions renders the effect on the $m_{\mathrm{T}}^{\mathrm{tot}}$ distributions to be shape-altering. Independent uncertainties are applied for the online selections of electrons and muons in single object triggers and triggers selecting $\ell\tau_{\mathrm{h}}$ pairs. The magnitude of the uncertainty applied in both cases amounts to 2 %. In the case of online selections of hadronically decaying tau leptons, the uncertainties in the non-parametric fit of the efficiencies in simulation and the Crystal-ball fit in the τ-embedded data sets are propagated to the $m_{\mathrm{T}}^{\mathrm{tot}}$ distributions. Independent variations of the corrections below and above $p_{\mathrm{T}}^{\tau_{\mathrm{h}}} = 100\,\mathrm{GeV}$ are allowed in the $\tau_{\mathrm{h}}\tau_{\mathrm{h}}$ final state. The uncertainties in the corrections of the online selection efficiencies are treated as partially correlated between simulation and τ-embedded data sets.

Uncertainties in the measured corrections of the τ_{h} identification efficiency are included in the analysis as shape-altering uncertainties. Independent uncertainties are used in three p_{T} regions below $p_{\mathrm{T}}^{\tau_{\mathrm{h}}} = 100\,\mathrm{GeV}$ in the $\ell\tau_{\mathrm{h}}$ final states and for the four distinguished decay modes in the $\tau_{\mathrm{h}}\tau_{\mathrm{h}}$ final state. Above $p_{\mathrm{T}}^{\tau_{\mathrm{h}}} = 100\,\mathrm{GeV}$ two independent uncertainties are used below and above a threshold of $p_{\mathrm{T}}^{\tau_{\mathrm{h}}} = 500\,\mathrm{GeV}$. These are treated as fully correlated between the $\ell\tau_{\mathrm{h}}$ and $\tau_{\mathrm{h}}\tau_{\mathrm{h}}$ final states. The uncertainties are obtained from the measurement of the identification corrections and treated as partially correlated between simulation and τ-embedded data sets.

Variations of the systematic uncertainties in the measured response corrections for electrons are used to model the uncertainty in the $m_{\mathrm{T}}^{\mathrm{tot}}$ distribution. A combined systematic uncertainty from variations of the distribution of the R_9 variable, the general electron selections, and p_{T} requirements is obtained by adding the contributions in quadrature [91]. The combined uncertainties range from 0.05–0.1 % in the ECAL barrel to 0.1–0.3 % in the ECAL endcap region depending on the R_9 category. For electrons with high transverse momenta an additional uncertainty is derived that amounts to roughly 1 % in the p_{T} region with $120 < p_{\mathrm{T}} < 300\,\mathrm{GeV}$ and is inflated by a factor of 2 in the ECAL barrel and a factor of 3 in the ECAL endcap region [91]. In the τ-embedded data sets the uncertainty in the energy response amounts to 0.5 % in the ECAL barrel and 1.25 % in the ECAL endcap region [15]. The variations in the barrel and endcap regions are treated as fully correlated in the analysis. The uncertainties in the electron energy response are treated as uncorrelated between simulation and τ-embedded data sets and correlated across data-taking years.

Uncertainties in the energy response of reconstructed τ_{h} candidates are obtained from the measurements for the three sources of τ_{h} candidates. In the case of genuine τ_{h} candidates, independent uncertainties are applied for the four distinguished decay modes, whereas for the τ_{h} candidates originating from misidentified electrons and muons only uncertainties in the measured values for one prong decays are applied. The uncertainties are treated as uncorrelated across data-taking years and partially correlated between simulation and τ-embedded data sets for genuine τ_{h} candidates. The uncertainties applied to misidentified leptons are only applied to the distributions of the $\mathrm{Z} \to \ell\ell$ process.

Variations of the energy scale of jets originating from quarks and gluons are applied to account for the uncertainty in their energy calibration. Multiple sources of uncertainty are defined to model the correlations of the uncertainties across p_{T} and η regions [93]. A reduced set of eleven uncertainty sources is obtained from grouping multiple sources based on detector regions in which the jets are reconstructed and to model correlations across years. The uncertainty in the measurement of the energy resolution of jets is propagated to their energy as well, following the same procedure as for the actual correction. It is treated as uncorrelated across years. The variations of the energy scale and resolutions of jets are propagated to the missing transverse energy and thus directly affect the $m_{\mathrm{T}}^{\mathrm{tot}}$ distributions of the simulated processes.

Different sets of additional variations of the missing transverse momentum are applied based on the presence of bosonic resonances in the simulated events. In simulated data sets without bosonic resonances the transverse momenta of PF candidates not associated with reconstructed leptons, photons, jets, or hadronically decaying tau leptons are varied within their measured momentum resolution. The differences between the varied and nominal momenta are then propagated to $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$ [98]. In events containing bosonic resonances, variations of the $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$ response and resolution within the uncertainties of the measured corrections based on the hadronic recoil in the event are applied. Both sets of variations of $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$ are treated as uncorrelated across data-taking years.

In the τ-embedded data sets, variations of $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$ are produced from the application of reduced and increased values of the corrections to the spurious missing transverse momentum. These variations range from 14 to 64 % of the full correction in the 2018 and 2017 data-taking years to account for the systematic uncertainty in the measurement [30].To cover statistical uncertainties in the measurement, an additional variation of 25 % is applied independently for the $\mathrm{e}\tau_{\mathrm{h}}$, $\mu\tau_{\mathrm{h}}$, and $\tau_{\mathrm{h}}\tau_{\mathrm{h}}$ final states. The variations are applied independently in the three data-taking years.

Uncertainties in the reweighting procedures applied to the p_{T} and mass distributions of Z bosons from the simulated $\mathrm{Z} \to \ell\ell$ process and the p_{T} distributions of top quarks from the $\mathrm{t\bar{t}}$ process are accounted for by applying the corrections twice or not using them at all. Differences between the predicted and observed p_{T} and mass distributions of the Z boson have also been noticed in the μ-embedded control data set [30]. Thus, an additional uncertainty is added to the predictions from the τ-embedded data sets that allows the fit to correct for these differences.

Additional uncertainties in the predictions from the τ-embedded data sets include an uncertainty in the expected fraction of $\mathrm{t\bar{t}}$ events in decays in two tau leptons contained in the embedded data sets, and an uncertainty in the track reconstruction efficiency of τ_{h} candidates in the embedded data sets [109]. The uncertainty in the expected fraction of $\mathrm{t\bar{t}}$ events is estimated from variations of 10 % of the $\mathrm{t\bar{t}}$ background in decays to two tau leptons from simulation. Variations of the track reconstruction efficiencies of 0.8 % and 1.3 % are applied to charged tracks and neutral pion candidates in hadronic tau lepton

decays [106].

During data taking in 2016 and 2017, a shift of the time response of the ECAL trigger cells was present at the L1 trigger in the forward regions [82]. In events with large energy deposits in the ECAL in these regions, the energy deposits may have been associated with the previous bunch crossing, which has then been selected by the trigger. As a positive response of the L1 trigger is not allowed for consecutive bunch crossings, the actual event from which the energy deposit originated was not selected. This effect, referred to as prefiring, is not modeled in simulation and can not be measured using the tag-and-probe method described in Section 4.5. The probability of an event to experience this prefiring effect has been measured by the CMS Collaboration and is used in the analysis to correct for the effect in simulation. Uncertainties in the measured probabilities are propagated to the final discriminant as correlated uncertainty in the 2016 and 2017 data-taking years.

Multiple uncertainties in the QCD multijet background estimation method in the eμ final state are considered. These include uncertainties in the three parameters of the polynomial fit for the extrapolation factors in each N_{jet} category in which the extrapolation factors are measured [15]. These uncertainties are treated as uncorrelated across the data-taking years. The uncertainty in the additional extrapolation factor from events with non-isolated muons to events with isolated muons is taken into account by variations of the $m_{\text{T}}^{\text{tot}}$ distribution created from applying the correction twice or not applying it at all. This uncertainty is treated as correlated across the data-taking years.

Furthermore, uncertainties in the expected rates of simulated backgrounds from four sources are propagated to distributions of the QCD multijet background. The propagation is performed through variations of these backgrounds during their subtraction in the AR. The four uncertainty sources comprise uncertainties in the differences of the probability to reconstruct an electron with the wrong charge, to misidentify a muon as an electron or a jet as a muon or an electron between data and simulation. These uncertainties are also applied directly to the distributions of the simulated backgrounds.

Multiple sources of uncertainties are considered in the background predictions from the F_{F}-method. Uncertainty sources common to all final states involving and each considered source of the jet $\rightarrow \tau_{\text{h}}$ backgrounds are statistical uncertainties in the parameters of the fits of the F_{F}^{i} and of the fits of the corrections to them. Furthermore, systematic uncertainties are defined from variations in which the corrections to the F_{F}^{i} are not applied or applied twice. Residual differences of the estimated and observed jet $\rightarrow \tau_{\text{h}}$ backgrounds in distributions of variables not used in the measurement in the DR_{i} are used to add further sources of uncertainty related to $F_{\text{F}}^{\text{QCD}}$ to the model. Variations of the background fractions of the W + jets and $\text{t}\bar{\text{t}}$ processes represent another source of uncertainties.

In the $\ell\tau_{\text{h}}$ final states, additional uncertainties arising from correlated variations of all statistical uncertainties and variations of the predicted backgrounds in the phase space

region governed by the $\ell\tau_{\mathrm{h}}$ trigger are included in the statistical model. Further uncertainties common to all final states include the uncertainty in the choice of the function used to model the F_{F}^{i} and the subtraction of simulated backgrounds in the AR. The uncertainty in the subtraction of backgrounds is modeled from variations of the normalization of the subtracted backgrounds in the AR by 10 %. All uncertainties in the background prediction from the F_{F}-method are treated as independent across data-taking years and investigated final states.

5.3 Signal Model Construction

This section discusses the construction of the signal model from the simulated signal data sets for both the model-independent and model-dependent interpretations of the data. As described in Section 4.4, predictions for additional neutral Higgs bosons have been simulated separately for the gluon fusion and b quark associated production modes in a mass range from 60 to 3500 GeV.

The simulation of the signal process includes fixed order calculations of squared matrix elements in NLO QCD implemented in the `POWHEG` event generator and resummation of leading-logarithmic (LL) terms in the parton shower [129]. The necessity for this resummation arises as terms of the form $\ln\left(p_{\mathrm{T}}^{\phi}/m_{\phi}\right)$ emerge in the fixed-order calculation of p_{T} distributions of the neutral Higgs bosons and render them divergent for small transverse momenta [146]. The resummation, however, is only valid for small transverse momenta and the fixed-order calculation is more reliable for transverse momenta comparable to the mass of the Higgs boson and above. Thus, the introduction of an arbitrary momentum scale is necessary to switch between the calculations in the different regimes. Physical results should be independent from the arbitrary choice of this scale but the calculations are not as the series expansion is truncated at NLO and LL precision, respectively. A sub-optimal choice of the scale may impact the accuracy of the result negatively [146].

A special treatment of the Higgs boson production in gluon fusion is advantageous as the different masses of the quarks running in the loop introduce multiple scales in the calculations [146–148]. Three separate scales μ_i are defined for the top quark, μ_t, bottom quark, μ_b, and top-bottom interference, μ_{int}, contributions to the total differential cross section and their contributions are calculated separately [146]. The total differential cross section reads as

$$\frac{\mathrm{d}\sigma}{\mathrm{d}p_{\mathrm{T}}} = \left.\frac{\mathrm{d}\sigma_{\mathrm{t}}}{\mathrm{d}p_{\mathrm{T}}}\right|_{\mu_{\mathrm{t}}} + \left.\frac{\mathrm{d}\sigma_{\mathrm{b}}}{\mathrm{d}p_{\mathrm{T}}}\right|_{\mu_{\mathrm{b}}} + \left.\frac{\mathrm{d}\sigma_{\mathrm{int}}}{\mathrm{d}p_{\mathrm{T}}}\right|_{\mu_{\mathrm{int}}} . \tag{5.5}$$

The differential cross section of the top-bottom interference contribution is calculated by subtracting the top and bottom quark contributions from the total cross section via

$$\left.\frac{\mathrm{d}\sigma_{\mathrm{int}}}{\mathrm{d}p_{\mathrm{T}}}\right|_{\mu_{\mathrm{int}}} = \left.\frac{\mathrm{d}\sigma}{\mathrm{d}p_{\mathrm{T}}}\right|_{\mu_{\mathrm{int}}} - \left.\frac{\mathrm{d}\sigma_{\mathrm{t}}}{\mathrm{d}p_{\mathrm{T}}}\right|_{\mu_{\mathrm{int}}} - \left.\frac{\mathrm{d}\sigma_{\mathrm{b}}}{\mathrm{d}p_{\mathrm{T}}}\right|_{\mu_{\mathrm{int}}} . \tag{5.6}$$

The three differential cross sections on the right-hand side of the equation are evaluated at the interference scale μ_{int}.

In the MSSM and 2HDM, the couplings of quarks to the three neutral Higgs bosons are modified with respect to their SM values by the respective scaling of the Yukawa couplings Y_{t} and Y_{b}. In the cross section calculations split by internal quarks in the loop contributions, the modified couplings factorize from the remainder of the calculation. The derivation of the matching scales for the three contributions to the total differential cross section is independent from the scaling of the couplings as well [146]. This allows writing Equation 5.5 as

$$\tilde{\sigma} = |Y_{\text{t}}(\beta,\alpha)|^2\,\tilde{\sigma}_{\text{t}} + |Y_{\text{b}}(\beta,\alpha)|^2\,\tilde{\sigma}_{\text{b}} + |Y_{\text{t}}(\beta,\alpha)Y_{\text{b}}^{*}(\beta,\alpha)|\,\tilde{\sigma}_{\text{int}}\,, \tag{5.7}$$

where the short-hand notation $\tilde{\sigma}_i = \frac{\mathrm{d}\sigma_i}{\mathrm{d}p_{\text{T}}}\Big|_{\mu_i}$ has been introduced and β and α are the angles describing the ratio of the VEVs of the Higgs doublets and the mixing of the CP-even Higgs bosons, as discussed in Section 2.3.1.

To describe the transverse momentum distribution of the neutral Higgs bosons, it is thus sufficient to simulate the distributions at a reference parameter point in terms of the Yukawa couplings for each mass point, denoted as $(\beta_{\text{ref}}, \alpha_{\text{ref}})$. The result can then be translated to every other parameter point by means of rescaling the different contributions by the relative difference of the Yukawa couplings for both bottom and top quarks. The differential cross section is thus given by

$$\tilde{\sigma} = \frac{|Y_{\text{t}}(\beta,\alpha)|^2}{|Y_{\text{t}}(\beta_{\text{ref}},\alpha_{\text{ref}})|^2}\tilde{\sigma}_{\text{t}} + \frac{|Y_{\text{b}}(\beta,\alpha)|^2}{|Y_{\text{b}}(\beta_{\text{ref}},\alpha_{\text{ref}})|^2}\tilde{\sigma}_{\text{b}} + \frac{|Y_{\text{t}}(\beta,\alpha)Y_{\text{b}}^{*}(\beta,\alpha)|}{|Y_{\text{t}}(\beta_{\text{ref}},\alpha_{\text{ref}})Y_{\text{b}}^{*}(\beta_{\text{ref}},\alpha_{\text{ref}})|}\tilde{\sigma}_{\text{int}}\,. \tag{5.8}$$

In the model-dependent interpretation, the Yukawa couplings necessary for the rescaling are extracted directly from the model-predictions provided by the LHC Higgs Working Group 3 (LHCHWG-3) [63, 67]. In the model-independent interpretation, the Yukawa couplings expected for h_{SM} at equal mass are used.

The composition of the transverse momentum spectrum of the CP-odd Higgs boson is shown in Figure 5.2. The p_{T} spectra are shown for $\tan\beta = 5,\ 15,\ 30$, and 50 for $m_{\text{A}} = 700\,\text{GeV}$. With increasing values of $\tan\beta$, the contribution of the bottom quark loop increases relative to the contribution from the top quark loop and the p_{T} spectrum becomes softer. Two shape-altering uncertainties related to the simulation of the transverse momentum distributions are considered in the analysis [30]. The first uncertainty is obtained from variations of the matching scale, whereas the second is derived from variations of the factorization and renormalization scales used in the simulation. As the matching scales take different values for the loop contributions, the uncertainty is treated as uncorrelated across the three contributions.

In the model-independent interpretation of the data, a search for an additional neutral scalar resonance ϕ produced via gluon fusion or in association with b quarks above the

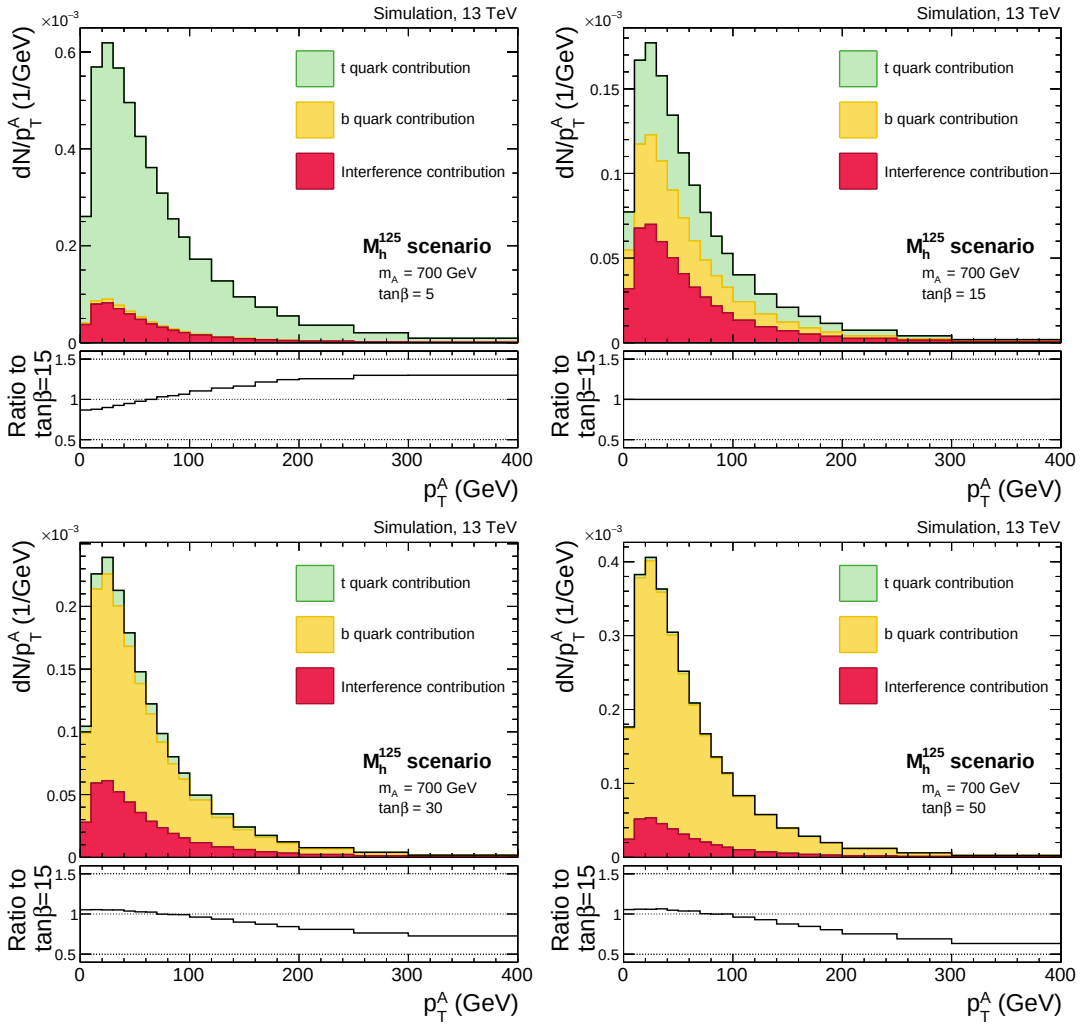

Figure 5.2: Composition of the Higgs p_{T} spectrum of the CP-odd Higgs boson for different parameter points of the M_{h}^{125} benchmark scenario. The parameter points are chosen to show the composition of the p_{T} spectrum as function of $\tan\beta$ for a fixed value of m_{A}. The softer p_{T} spectrum with increasing values of $\tan\beta$ is illustrated in the lower panels, where the probability density for the considered parameter point is shown normalized to the probability density at $\tan\beta = 15$. The Y_{t}, Y_{b} and the cross section predictions are taken from [67] and the histograms at the reference point from [30].

background predicted by the SM including $m_{\mathrm{H}(125)}$ is performed. The predictions of the distributions of the final discriminant for both production modes are taken directly from the available corresponding simulated data sets. Exclusion limits are derived in terms of the product of production cross section and branching fraction for the decay into tau leptons $\sigma\mathcal{B}(\phi \to \tau\tau)$ separately for both production modes. Consequently, two signal strength modifiers $\mu_{\mathrm{gg}\phi}$ and $\mu_{\mathrm{bb}\phi}$, which independently scale the normalization for both production modes, are introduced in the likelihood from Equation 5.1. The signal strength modifiers correspond to $\sigma_{\mathrm{gg}\phi/\mathrm{bb}\phi}\,\mathcal{B}(\phi \to \tau\tau)$ measured in pb. More than one signal source can be introduced in the likelihood of Equation 5.1 by the replacement

$$\mu s_i \to \sum_k \mu_k s_{ki}\,, \tag{5.9}$$

where k runs over all considered signal sources. During the calculation of exclusion limits on one of the production modes, the signal strength parameter of the other is considered as additional unconstrained nuisance parameter and profiled.

In the model-dependent interpretation of the data, the complete spectrum of neutral Higgs bosons predicted by 2HDMs is exploited to derive exclusion limits in specific benchmark models. The signal model thus consists of three resonances in which one of the neutral scalar Higgs bosons is identified as H(125). For a large part of the parameter space of the models, the decoupling limit is reached. The additional neutral Higgs bosons become degenerate in mass and the mass spectrum effectively reduces to two resonances. The cross sections for both production modes and branching fractions for the decay into tau leptons are scaled independently for the three neutral Higgs bosons to the values predicted by the considered benchmark scenario. In the case of the observed Higgs boson, the vector boson fusion and vector boson associated production modes are considered as well.

The predictions of the masses of the neutral Higgs bosons, their production cross sections, and branching ratios in the MSSM benchmark scenarios are provided by the LHCHWG-3 in the form of `ROOT` files [67]. In the case of the 2HDM benchmark scenarios, the model predictions are calculated as described in Section 2.3.1. Lastly, the individual distributions for all neutral Higgs bosons and production modes are combined to construct the full signal model. The total transverse mass distribution of the full signal model in the model-dependent interpretation is shown for two example benchmark points in Figure 5.3. The shown distributions serve only as illustration of the Higgs spectrum in the MSSM and are not used in the same way in the statistical inference. Instead, the predictions of the properties of h are modified as described in Section 5.4.

As opposed to the model-independent interpretation, the masses of the three neutral Higgs bosons can not be freely chosen to correspond to the mass points for which the signal processes are simulated but are continuous functions of the two free parameters of the benchmark model. The necessity to interpolate the distribution of the final discriminant thus arises since the distributions are only available for the discrete set of simulated mass points. The interpolation is performed using a horizontal morphing technique based on

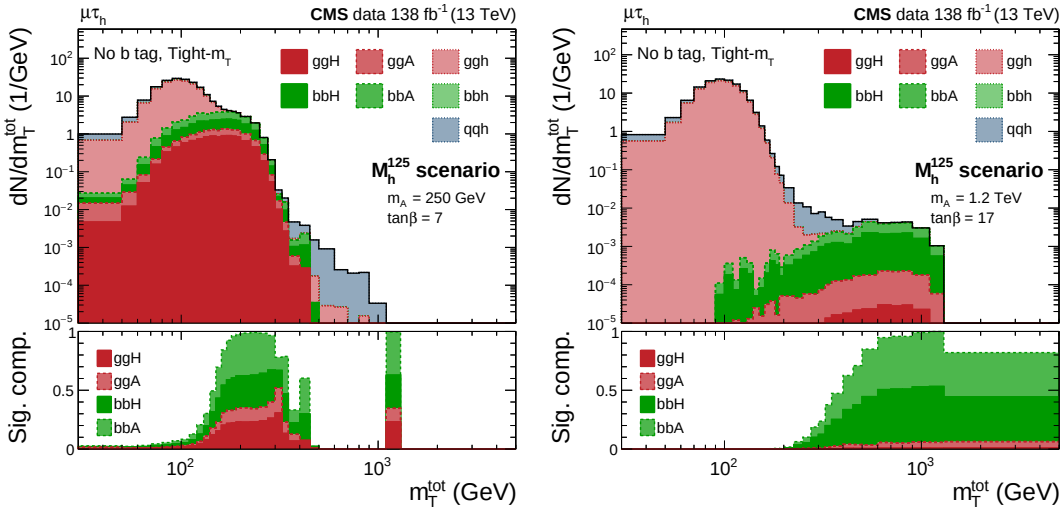

Figure 5.3: Distribution of the total transverse mass of the neutral MSSM Higgs boson spectrum at two parameter points in the M_{h}^{125} benchmark scenario. The distributions are shown for illustrative purposes and do not enter the statistical inference as shown in the plot. The upper panels show the contribution of each neutral Higgs boson to the $m_{\mathrm{T}}^{\mathrm{tot}}$ distribution in the ggϕ and bbϕ production modes. The lower panels show the relative contribution of each heavy neutral Higgs boson to the total prediction from all signal processes. The contribution from the light neutral Higgs boson (ggh, bbh, qqh) are not explicitly shown in the lower panel and are thus indicated by the white area.

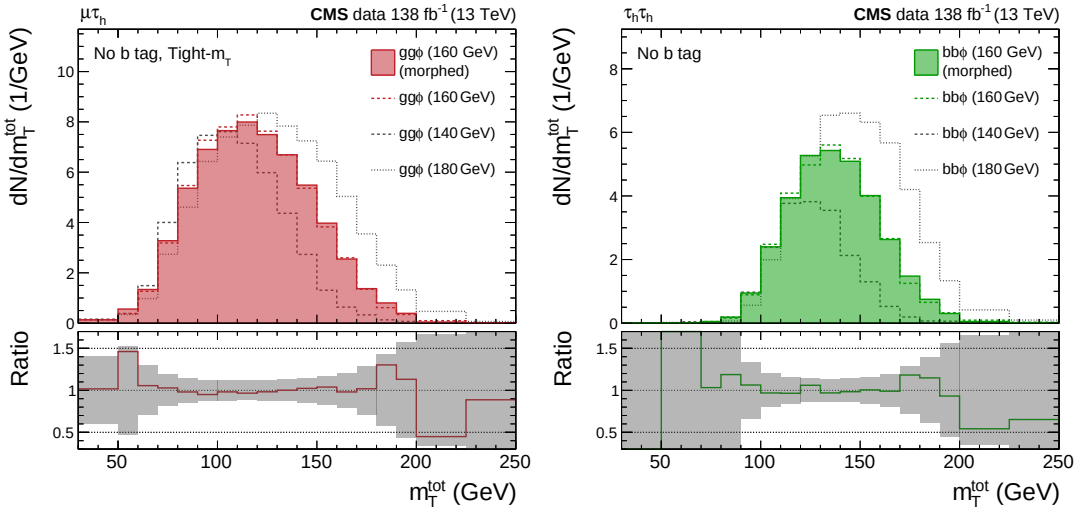

Figure 5.4: Illustration of the horizontal morphing technique used in the model-dependent interpretation. A comparison of the $m_{\mathrm{T}}^{\mathrm{tot}}$ distributions for $m_{\phi} = 160\,\mathrm{GeV}$ obtained from the neighboring simulated mass points of 140 and 180 GeV and directly from simulation is displayed in the upper panels. The morphed distributions are shown as filled areas. The distributions for ggϕ and bbϕ production are shown in red and in green, respectively. The distribution obtained directly from simulation is shown as colored dashed line. The distributions from the neighboring mass points are shown as gray dashed lines. The lower panels show the ratio of the prediction from the morphed distribution relative to the distribution obtained directly from simulation. The statistical uncertainties in the prediction obtained from simulation are indicated by the gray shaded areas. The normalization of all predictions corresponds to a cross section of 1 pb.

the cumulative distribution functions of the $m_{\mathrm{T}}^{\mathrm{tot}}$ distributions of neighboring simulated mass points [149].

In Figure 5.4, the technique is illustrated in a comparison of the distribution of the total transverse mass obtained using the horizontal morphing technique and taken directly from the simulated data set at $m_{\phi} = 160\,\mathrm{GeV}$. The comparison is shown for ggϕ production in the No b tag, Tight-m_{T} category in the $\mu\tau_{\mathrm{h}}$ final state in the left panel and for bbϕ production in the No b tag category in the $\tau_{\mathrm{h}}\tau_{\mathrm{h}}$ final state in the right panel. In the central region of the distributions, where the bulk of the events is located, the morphed and simulated distributions agree well, whereas some deviations can be observed in the tails of the distributions. These deviations occur as the method does not take the change in the bin width, present in the outer regions of the shown mass range, into account.

5.4 Derivation of Exclusion Limits

The derivation of exclusion limits on hypothesized physics models corresponds to the estimation of one-sided confidence intervals on μ. In this section, first, the method used to derive the exclusion limits is described. It is followed by a discussion of the differences in its application to the model-independent and model-dependent interpretations.

Description of the General Procedure

The test statistic used to estimate the confidence intervals is the profile likelihood ratio

$$\tilde{q}_\mu = -2 \ln \frac{\mathcal{L}(n|\mu, \hat{\hat{\boldsymbol{\theta}}}(\mu))}{\mathcal{L}(n|\hat{\mu}, \hat{\boldsymbol{\theta}}(\hat{\mu}))}\,,\ 0 \leq \hat{\mu} \leq \mu\,, \tag{5.10}$$

based on the likelihood as described in Section 5.1 [141, 150]. The numerator in the definition of $\tilde{q}_\mu$ is the profile likelihood function obtained from maximizing the likelihood of Equation 5.4 with respect to the nuisance parameters for a fixed value μ. In this notation, $\hat{\hat{\boldsymbol{\theta}}}(\mu)$ corresponds to the vector of nuisance parameters that maximize the likelihood for given μ, and $\hat{\mu}$ and $\hat{\boldsymbol{\theta}}$ represent the maximum likelihood estimates of μ and $\boldsymbol{\theta}$. In the definition of $\tilde{q}_\mu$, the assumption of a single signal strength modifier scaling the signal process is entering again. In the derivation of the exclusion limits in the model-independent interpretation, possible further signal strength modifiers are profiled like the other nuisance parameters during the fits performed for the calculation of $\tilde{q}_\mu$.

In case $\hat{\mu}$ is found to be smaller than zero, it is clipped to zero in the calculation of $\tilde{q}_\mu$. This yields

$$\tilde{q}_\mu = -2 \ln \frac{\mathcal{L}(n|\mu, \hat{\hat{\boldsymbol{\theta}}}(\mu))}{\mathcal{L}(n|0, \hat{\boldsymbol{\theta}}(0))}\,,\ \hat{\mu} < 0 \tag{5.11}$$

as value of $\tilde{q}_\mu$ in this case. Since the objective of the statistical inference is the upper limit on μ, $\tilde{q}_\mu$ is set to zero in the case that $\hat{\mu}$ exceeds μ [141, 150].

From the observed value of the test statistic, $\tilde{q}_\mu^{\text{obs}}$, and the distribution of $\tilde{q}_\mu$ under the null hypothesis, $f(\tilde{q}_\mu|\mu)$, the p-value p_μ, quantifying the statistical compatibility of the null hypothesis with the observed data, is calculated. The p-value is defined as

$$p_\mu = \int_{\tilde{q}_\mu^{\text{obs}}}^{\infty} f(\tilde{q}_\mu|\mu)\, \mathrm{d}\tilde{q}_\mu \tag{5.12}$$

and corresponds to the probability to observe $\tilde{q}_\mu > \tilde{q}_\mu^{\text{obs}}$ given the null hypothesis. When deriving exclusion limits, the null hypothesis corresponds to the presence of the signal process with $\mu > 0$ and is thus also referred to as signal-plus-background (s+b) hypothesis. The exclusion limit on μ at the confidence level (CL) $1 - \alpha$, chosen prior to calculating $\tilde{q}_\mu^{\text{obs}}$, is then obtained by solving the equation

$$p_\mu = \alpha \tag{5.13}$$

for μ. For the results presented in Chapter 6, the CL is chosen to be 95 %.

In general, the distribution of an arbitrary test statistic t under the null hypothesis H_0, $f(t|H_0)$, might not be known analytically. In this case, one has to rely on ensemble tests to sample $f(t|H_0)$, which may easily become computationally expensive. In the asymptotic limit of large numbers, however, $f(\tilde{q}_\mu|\mu)$ is known analytically [141, 150]. To estimate the variance of $\hat{\mu}$, entering the calculation of $f(\tilde{q}_\mu|\mu)$, the Asimov data set is used, which is a single representative data set in which all model parameters take their expectation values [141, 150]. In practice, it is constructed from the background predictions, when deriving exclusion limits.

The p-value p_μ can then be calculated from the known cumulative distribution function of $\tilde{q}_\mu$, $F(\tilde{q}_\mu|\mu)$, and is given by

$$p_\mu = 1 - F(\tilde{q}_\mu^{\text{obs}}|\mu) = \begin{cases} 1 - \Phi(\sqrt{\tilde{q}_\mu^{\text{obs}}}), & \tilde{q}_\mu^{\text{obs}} \leq \tilde{q}_{\text{A}} \\ 1 - \Phi\left(\frac{\tilde{q}_\mu^{\text{obs}} + \tilde{q}_{\text{A}}}{2\sqrt{\tilde{q}_\mu^{\text{obs}}}}\right), & \tilde{q}_\mu^{\text{obs}} > \tilde{q}_{\text{A}} \end{cases}, \tag{5.14}$$

where $\tilde{q}_{\text{A}}$ refers to the value of $\tilde{q}_\mu$ for the Asimov data set and Φ represents the cumulative distribution function of the normal distribution [141, 150].

In the case that the rate of the considered signal process is small, two issues of the frequentist calculations of exclusion limits as described above arise. Firstly, the expected sensitivity of an experiment increases with the expected number of background events, even though the expected number of signal events stays the same [151]. Lastly, signal might be excluded in the case that the observed number of events fluctuates below the background expectation [152]. To mitigate these issues, the CL_s criterion [151, 152] is used to derive the exclusion limits. Before solving Equation 5.13 for μ, the p-value is divided by the probability to observe $\tilde{q}_\mu > \tilde{q}_\mu^{\text{obs}}$ given the background-only (b-only) hypothesis,

$$\text{CL}_\text{s} = \frac{p_\mu}{1 - p_0}, \tag{5.15}$$

where p_0 is given by $p_0 = F(\tilde{q}_\mu^{\text{obs}}|0)$. In the limit of high sensitivity, the denominator approaches unity and the frequentist calculation is retained. For low sensitivity, the CL_s value is increased with respect to p_μ leading to more conservative exclusion limits.

In order to use the CL_s criterion, $f(\tilde{q}_\mu|0)$ has to be known in addition to $f(\tilde{q}_\mu|\mu)$. [141, 150] In the limit of large numbers, it is known analytically and follows a distribution similar to $f(\tilde{q}_\mu|\mu)$. The value of p_0 is then given by

$$p_0 = \begin{cases} \Phi(\sqrt{\tilde{q}_\mu^{\text{obs}}} - \sqrt{\tilde{q}_{\text{A}}}), & \tilde{q}_\mu^{\text{obs}} \leq \tilde{q}_{\text{A}} \\ \Phi\left(\frac{\tilde{q}_\mu^{\text{obs}} - \tilde{q}_{\text{A}}}{2\sqrt{\tilde{q}_\mu^{\text{obs}}}}\right), & \tilde{q}_\mu^{\text{obs}} > \tilde{q}_{\text{A}} \end{cases}. \tag{5.16}$$

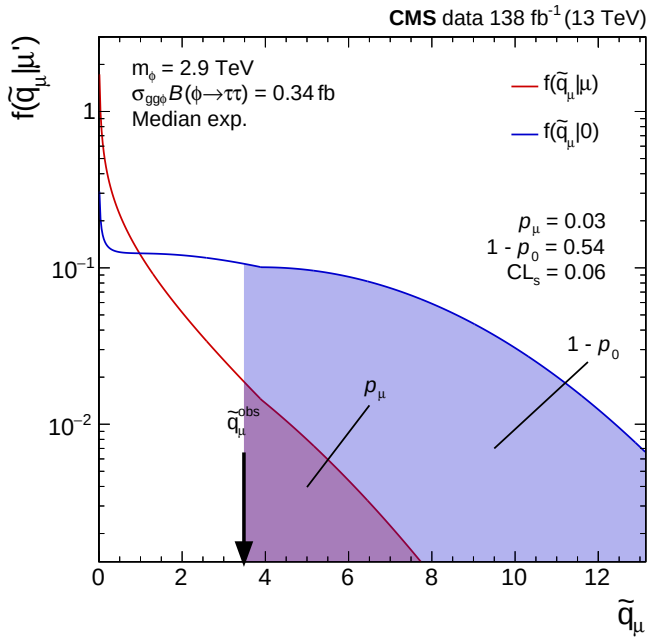

Figure 5.5: Illustration of the calculation of p_μ and $\mathrm{CL_s}$ using the asymptotic approximation. The distributions of $\tilde{q}_\mu$ given the $s+b$ and b-only hypotheses are shown as red and blue lines, respectively. The values of p_μ and $1-p_0$, which enter the calculation of $\mathrm{CL_s}$ in the numerator and denominator, are indicated by filled red and blue areas and their values are listed in the figure. The observed value of the test statistic, $\tilde{q}_\mu^{\mathrm{obs}}$, is indicated by a black arrow. The plot is shown for $m_\phi = 2.9\,\mathrm{TeV}$ and $\sigma_{\mathrm{gg}\phi}\,\mathcal{B}(\phi \to \tau\tau) = 0.34\,\mathrm{fb}$, which corresponds to the median expected limit at the 95 % CL obtained using the $\mathrm{CL_s}$ criterion.

The calculation of p_μ and $\mathrm{CL_s}$ for $m_\phi = 2.9\,\mathrm{TeV}$ and $\mu = \sigma_{\mathrm{gg}\phi}\,\mathcal{B}(\phi \to \tau\tau) = 0.34\,\mathrm{fb}$, which corresponds to the median expected limit in absence of signal, is illustrated in Figure 5.5. In Figure 5.6, the same illustration is shown for a set of μ values corresponding to the 2.5, 16, 50, 84, and 97.5 % expected, and the observed exclusion limits in ggϕ production for $m_\phi = 2.9\,\mathrm{TeV}$. The increasing separation of the distributions based on the $s+b$ and b-only hypothesis with increasing number of expected signal events is clearly visible.

In the high-mass tails of the $m_{\mathrm{T}}^{\mathrm{tot}}$ distributions, the expected number of background events is small. To validate that the asymptotic approximation is still valid for high-mass signals, a comparison of the exclusion limits in the model-independent interpretation obtained from the asymptotic formulae for $f(\tilde{q}_\mu|\mu)$ and $f(\tilde{q}_\mu|0)$ and from ensemble tests has been performed for a set of five mass points ranging from 2.3 to 3.5 TeV in ggϕ and bbϕ production each.

To derive the exclusion limits following both approaches, $\mathrm{CL_s}$ is calculated at six grid points corresponding to the median, 68, and 95 % expected, and observed exclusion limits in the asymptotic approximation. The ensembles entering the calculation of $\mathrm{CL_s}$ at each

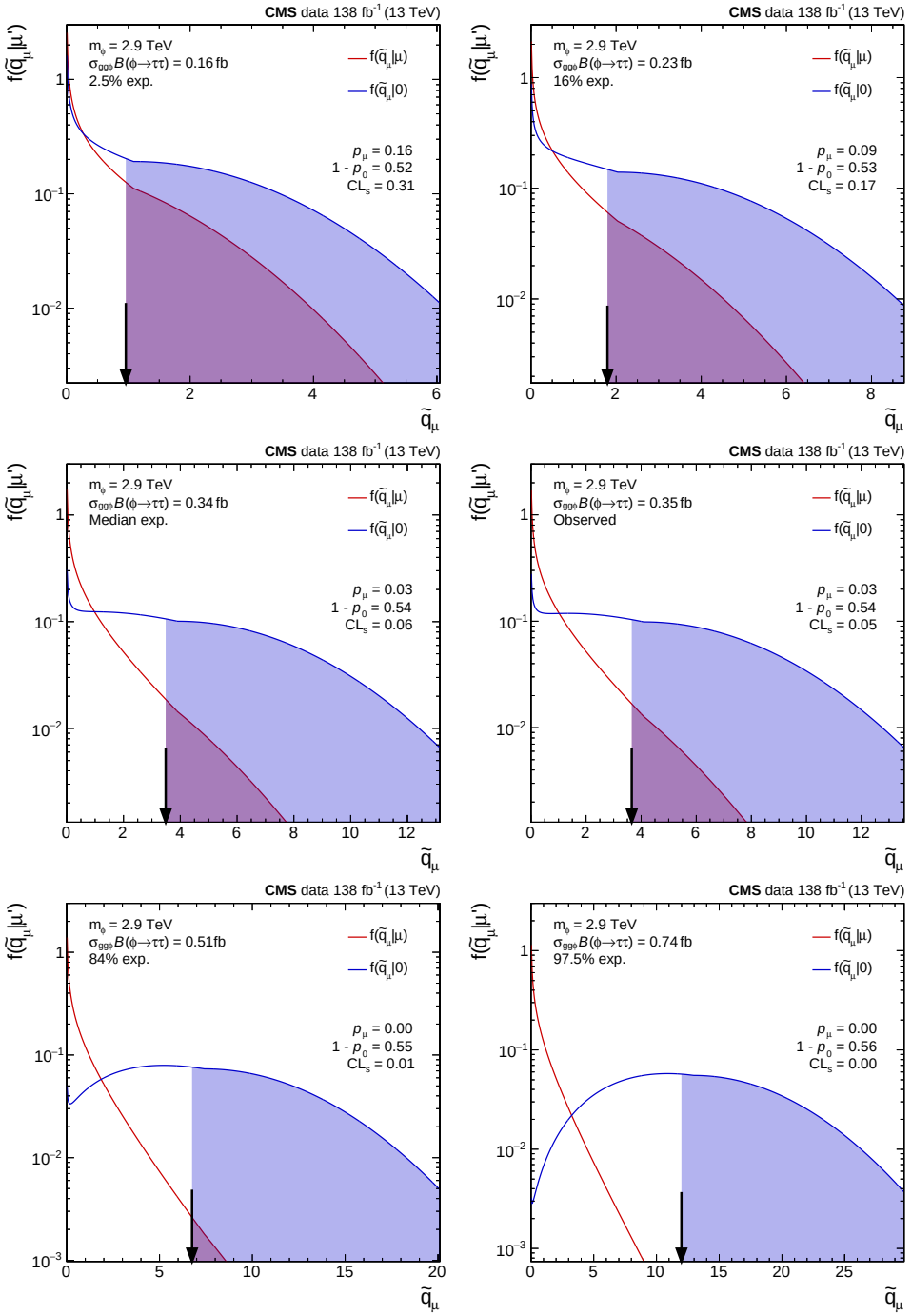

Figure 5.6: Illustration of $\mathrm{CL_s}$ for different values of $\sigma_{\mathrm{gg}\phi}\,\mathcal{B}(\phi \to \tau\tau)$. The distributions of the test statistic $\tilde{q}_\mu$ given the signal-plus-background and background-only hypotheses are shown for the mass point of $m_\phi = 2.9\,\mathrm{TeV}$. The shown values of $\sigma_{\mathrm{gg}\phi}\,\mathcal{B}(\phi \to \tau\tau)$ correspond to the 2.5, 16, 50, 84, and 97.5 % quantiles of the expected exclusion limit from left to right and top to bottom interspersed with the values of the observed exclusion limit in the middle right plot. The observed value of the test statistic $\tilde{q}_\mu^{\mathrm{obs}}$, indicated by a black arrow, changes across the different plots as the tested signal-plus-background hypothesis changes.

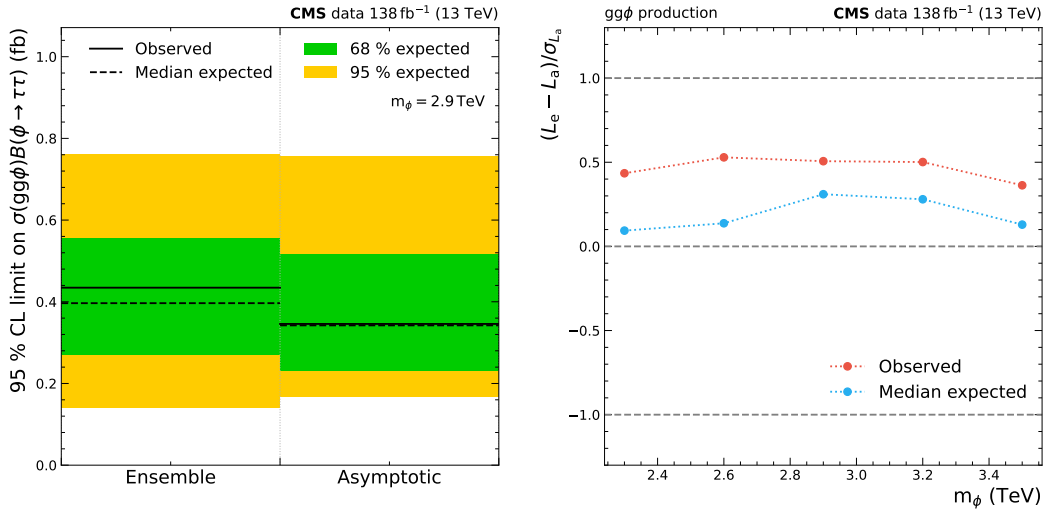

Figure 5.7: Comparison of the exclusion limits at the 95 % CL obtained from ensemble tests and the asymptotic approximation. In the left panel, a comparison of the observed and expected upper limits on ggϕ production at $m_\phi = 2.9\,\text{TeV}$ is shown. The observed and median expected limits are shown as solid and dashed black lines, respectively. The 68 and 95 % central confidence intervals for the expected exclusion limits are indicated by green and yellow bands. In the right panel, the difference between the exclusion limits on ggϕ production based on the ensemble tests (L_e) and the asymptotic approximation (L_a) is shown for the observed (red markers) and median expected (blue markers) limits. For a better assessment, the differences are normalized to the 68 % expected confidence interval in the asymptotic approximation σ_{L_a}. The 68 % central confidence interval for the expected limit and the median expected limit in the asymptotic approximation are indicated by gray dashed lines.

grid point comprise 5000 toy data sets generated from the $s + b$ and b-only hypotheses each. A linear interpolation between the grid points is performed to extract the exclusion limits at the 95 % CL.

A comparison of the exclusion limits on ggϕ production for $m_\phi = 2.9\,\text{TeV}$ is shown in Figure 5.7 (left). The right panel of Figure 5.7 shows the difference between the exclusion limits on ggϕ production based on ensemble tests and the asymptotic approximation. For a better assessment, the differences are normalized to the 68 % expected confidence interval derived in the asymptotic approximation. For both production modes, the calculation based on ensemble tests yields more conservative exclusion limits. The differences between the two approaches are within 60 % of the 68 % expected confidence interval. As the exclusion limits in the two approaches are compatible within the 68 % expected limits and the calculation of the toy based results is computationally expensive, the results presented in the following are based on the asymptotic approximation.

Adaptations for the Interpretation in MSSM and 2HDM Benchmark Scenarios

In the model-independent interpretation, the exclusion limits are derived by solving

$$\mathrm{CL_s}(\mu) = \alpha \tag{5.17}$$

for μ numerically for every tested mass point. In the model-dependent interpretations, $\mathrm{CL_s}$ is evaluated for a fixed value of $\mu = 1$, corresponding to the model predicted by the benchmark scenario. The calculation of $\mathrm{CL_s}$ is performed in a 2D grid as function of the free parameters of each corresponding scenario. The exclusion limits are then obtained by finding the contour where Equation 5.17 is fulfilled.

In principle, the signal model for $\mu = 1$ is constructed as described in Section 5.3. The use of the asymptotic approximation and the uncertainty in the prediction of m_h, however, require two modifications to the predictions for h. The asymptotic formulae for the distributions of the test statistic under the given hypotheses are only valid for nested models, in which the $s + b$ hypothesis is a superset of the b-only hypothesis and transforms into it for vanishing signal strength [153]. As H(125) is part of both the b-only and $s+b$ hypotheses, the likelihood is reparameterized to make this nesting more obvious.

The SM prediction for the observed Higgs boson is treated as part of the background model and the potential difference in normalization between the BSM and SM prediction is scaled with μ [30]. The functional form of the likelihood in this case is given by

$$\mathcal{L}(n|\mu, \boldsymbol{\theta}) = \mathcal{L}(n|\mu((h - h_\mathrm{SM}) + H + A) + h_\mathrm{SM} + b, \boldsymbol{\theta})\,. \tag{5.18}$$

For clarity, the SM prediction h_SM for H(125) is written explicitly and not subsumed in the background prediction b. The dependence of the predictions for all neutral Higgs bosons, h, H, A, and h_SM, and the background on the nuisance parameters is implied in the equation. The signal templates for the difference in normalization of h and h_SM are obtained from the simulated samples for SM Higgs boson production. A physically valid model is only obtained from the nested model when μ takes values of zero or unity. When deriving exclusion limits, the background-only hypothesis is expected to be realized and thus the maximum likelihood estimate is expected to be close to zero.

Another modification in the derivation of exclusion limits in the model-dependent interpretation concerns the predictions for h in the MSSM benchmark scenarios. The cross section and branching ratio predictions in these scenarios are functions of the mass of the considered Higgs boson. The theoretical calculations of the mass predictions are truncated in a series expansion and an uncertainty is assigned to the prediction to take the effect of missing higher-order calculations into account. This uncertainty is taken to be $m_\mathrm{h} \pm 3\,\mathrm{GeV}$ in the case of H(125) for all parameter points with m_h being the predicted mass for the given parameter point [70]. Experimental analyses, however, are sensitive to the changes of the cross section and branching ratio predictions within this mass interval [63].

As the theoretical uncertainty has no statistical interpretation [70], there is no preferred value for the predicted mass of the observed Higgs boson within the given interval. To turn the analysis insensitive to the mass dependence of the predictions of the cross sections and branching fractions for h in the MSSM benchmark scenarios within the uncertainty in m_{h}, they are rescaled to the predictions for $m_{\mathrm{H}(125)} = 125.38\,\mathrm{GeV}$. The rescaling exploits the assumption that their scaling with the mass of the Higgs boson is the same as expected for h_{SM} [63]. The cross sections and branching fractions for $m_{\mathrm{H}(125)}$ are then given by

$$\sigma_{\mathrm{MSSM}}(m_{\mathrm{H}(125)}) = \frac{\sigma_{\mathrm{SM}}(m_{\mathrm{H}(125)})}{\sigma_{\mathrm{SM}}(m_{\mathrm{h}})}\sigma_{\mathrm{MSSM}}(m_{\mathrm{h}})\,, \tag{5.19}$$

$$\mathcal{B}_{\mathrm{MSSM}}(m_{\mathrm{H}(125)}) = \frac{\mathcal{B}_{\mathrm{SM}}(m_{\mathrm{H}(125)})}{\mathcal{B}_{\mathrm{SM}}(m_{\mathrm{h}})}\mathcal{B}_{\mathrm{MSSM}}(m_{\mathrm{h}})\,. \tag{5.20}$$

Here, σ_{MSSM} and $\mathcal{B}_{\mathrm{MSSM}}$ refer to the cross section and branching fraction predictions in the corresponding MSSM benchmark scenarios and σ_{SM} and $\mathcal{B}_{\mathrm{SM}}$ the respective quantities as predicted in the SM.

5.5 Validation of the Statistical Model

Prior to the extraction of exclusion limits in both interpretations, the statistical model is validated. Two different methods are pursued for this purpose. Firstly, goodness-of-fit (GoF) tests are performed with distributions of the discriminating variable as input. These tests are based on a profile likelihood ratio test statistic t involving the saturated likelihood model in the denominator [154]. The value of the test statistic is calculated as

$$t = -2\ln\frac{\mathcal{L}(n|0,\hat{\boldsymbol{\theta}}_0)}{\mathcal{L}_{\mathrm{sat}}}\,, \tag{5.21}$$

where $\mathcal{L}_{\mathrm{sat}}$ is given by

$$\mathcal{L}_{\mathrm{sat}} = \mathcal{L}(n|\mu s + b = n, \hat{\boldsymbol{\theta}}_n)\,. \tag{5.22}$$

Here, $\hat{\boldsymbol{\theta}}_0$ indicates the maximum likelihood estimate of the nuisance parameters for a signal strength of zero and $\hat{\boldsymbol{\theta}}_n$ for the model where the expected numbers of events are taken to be the observed ones. During the GoF tests, the signal strength μ is fixed to zero since the tests should not depend on a choice of a specific signal model. The p-values and critical regions of the GoF tests are estimated using ensemble tests. The tests are performed separately for each category, the combination of all categories per final state for each data-taking year, the combination of all categories per data-taking year and the combination of all categories, to verify the description of the data in each category and the correlations of the nuisance parameters across data-taking years and final states. In every case, a fit to the data is performed for the investigated categories and the $\mathrm{t\bar{t}}$ control regions from the respective data-taking years, which are included to constrain the free floating nuisance parameters controlling the $\mathrm{t\bar{t}}$ background. The test statistic

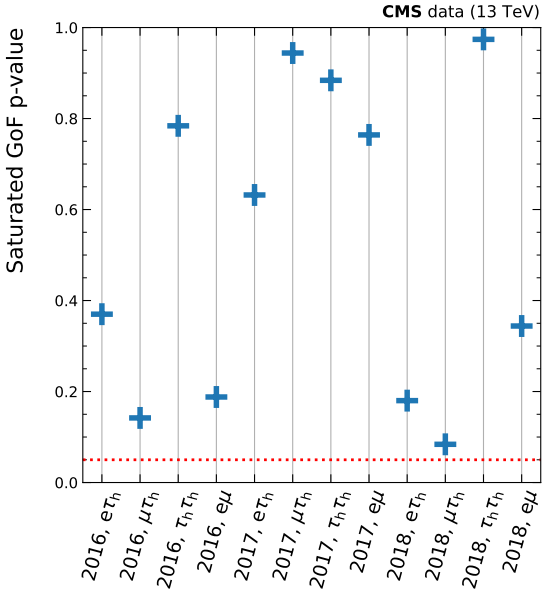

Figure 5.8: P-values of GoF tests performed for a fit of the combination of all categories per final state for each data-taking year. The p-value of each test is indicated by a blue cross. The critical region of the tests defined via $p < 0.05$ are indicated by a horizontal red line. The p-values are derived from ensemble tests consisting of 500 toy data sets each.

is then evaluated on the distributions of the investigated categories only. No significant deviations have been observed. From the 51 GoF tests carried out for the individual categories, only a single test rejects the null hypothesis at the 95 % CL. None of the performed GoF tests of combinations of categories rejects the null hypothesis. Figure 5.8 exemplary shows the p-values of the GoF tests performed for a fit of the combination of all categories per final state for each data-taking year.

Furthermore, differences between the best fit and initial values of the nuisance parameters relative to their standard deviations prior to the fit to the data and their impacts on the best fit value of $\sigma_{gg\phi}\,\mathcal{B}(\phi \to \tau\tau)$ are extracted from maximum likelihood fits to obtain further insights into the statistical model. For practical reasons, the fits are performed for a small subset of the investigated masses. The extracted differences between the best fit and initial values of the nuisance parameters quantify how much the background model has to be adapted to describe the data. Many nuisance parameters with large differences would hint to an insufficient modeling of the background.

Figure 5.9 shows the distribution of the differences for all nuisance parameters in the model for a fit of $\sigma_{gg\phi}\,\mathcal{B}(\phi \to \tau\tau)$ at a mass of $m_\phi = 250\,\text{GeV}$. The distributions are shown separately for the nuisance parameters modeling the effect of the limited number of events from which the background and signal predictions are formed and all other nuisance parameters. A normal distribution with zero mean and unity variance is overlaid

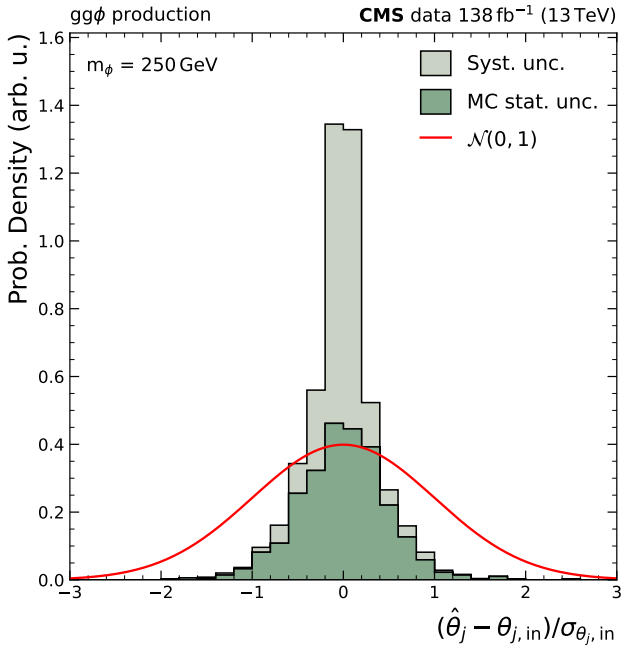

Figure 5.9: Distribution of the differences between the best fit and initial values of all nuisance parameters normalized to their initial standard deviation. The differences are shown for a maximum likelihood fit of $\sigma_{\mathrm{gg}\phi}\,\mathcal{B}(\phi \to \tau\tau)$ for $m_\phi = 250\,\mathrm{GeV}$. The nuisance parameters modeling the uncertainties arising from the limited number of events used to form the background and signal predictions (MC stat. unc.) are shown separately from the remaining nuisance parameters (syst. unc.). A normal distribution with zero mean and unity standard deviation normalized to the number of nuisance parameters in the model is overlaid as red line.

for comparison. A single value of the differences of the best fit and initial values of the nuisance parameters larger than two initial standard deviations are observed and the differences are in general smaller than expected from the overlaid normal distribution. The distributions for other mass points and bbϕ production show similar behavior.

The impacts of the nuisance parameters on the best fit value of $\sigma_{gg\phi}\,\mathcal{B}(\phi \to \tau\tau)$ are obtained by fixing the investigated nuisance parameter to the boundaries of its 68 % confidence interval after the fit to the data before repeating the fit. All other nuisance parameters are profiled again during this second fit. The impacts provide information on the nuisance parameters with the largest effect on the measured value of $\sigma_{gg\phi}\,\mathcal{B}(\phi \to \tau\tau)$ and the correlation of the signal with the nuisance parameter in question. For $m_\phi = 1.2\,\text{TeV}$, the nuisance parameters with the largest impact on $\sigma_{gg\phi}\,\mathcal{B}(\phi \to \tau\tau)$ belong to the nuisance parameters modeling the uncertainties arising from the limited number of events used to form the background and signal predictions in the region, where the majority of signal events is located. For $m_\phi = 250\,\text{GeV}$, uncertainties in the background predictions from the F_F-method exhibit the largest impacts on the measured strength. The largest impact from a single nuisance parameter amounts to roughly 6 % (14 %) of the best fit value for $\sigma_{gg\phi}\,\mathcal{B}(\phi \to \tau\tau)$ for $m_\phi = 1.2\,\text{TeV}$ ($m_\phi = 250\,\text{GeV}$).

The studies outlined in this section show that the statistical model detailed above is able to describe the observed data appropriately. Thus, the statistical model can be used to derive the exclusion limits in both the model-independent and model-dependent interpretations shown in the following chapter.

CHAPTER 6

Results

6.1 Model-Independent Interpretation

The observed and expected 95 % CL upper exclusion limits on $\sigma\,\mathcal{B}(\phi \to \tau\tau)$ for the model-independent ggϕ and bbϕ production are displayed in Figure 6.1. In both cases, the observed limit for the investigated mass points is indicated by the black markers and linearly interpolated in between. The median expected exclusion limits in the absence of signal are indicated by the black dashed line. The green and yellow bands indicate the 68 and 95 % central confidence intervals for the expected limits.

The exclusion limits span five orders of magnitude over the investigated mass range. At $m_\phi = 60\,\text{GeV}$, the observed exclusion limits on $\sigma\,\mathcal{B}(\phi \to \tau\tau)$ are at 13 pb and 23 pb for ggϕ and bbϕ production. With increasing values of m_ϕ, the distribution of m_T^tot shifts to larger values. In this region, the background distribution drops off exponentially increasing the sensitivity of the search in turn. The lowest exclusion limit with values of 0.28 fb and 0.33 fb for ggϕ and bbϕ production is achieved for the highest probed mass of $m_\phi = 3.5\,\text{TeV}$.

At high masses the sensitivity of the search is driven by the $\tau_\text{h}\tau_\text{h}$, $\mu\tau_\text{h}$, and eτ_h final states as the branching fractions of the $\tau\tau$ system are larger in these final states than in the eμ final state. The $\tau_\text{h}\tau_\text{h}$ final state is the most sensitive final state over a large mass range since the background contributions are lower the $\tau\tau$ branching fractions higher in this final state. At low masses, the contribution of the eμ final state to the sensitivity of the analysis is in the same order of magnitude as for the other final states. In this case, the sensitivity of the $\tau_\text{h}\tau_\text{h}$ final state is limited by the high trigger thresholds imposed on the transverse momenta of the τ_h candidates.

The median exclusion limits in absence of signal split by final state and data-taking year are shown in Figure B.1 and Figure B.2 in Appendix B.1, respectively. Each data-taking year contributes roughly the same sensitivity to the analysis. With respect to the previous CMS analysis of 2016 data [31], the sensitivity for each data-taking period alone has increased. This is also illustrated in Figure B.2. The increase in sensitivity originates

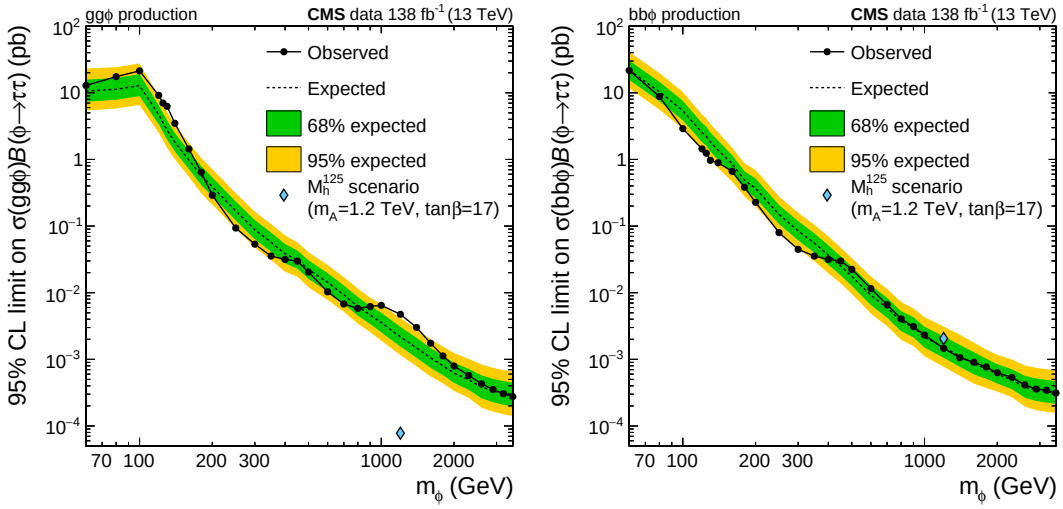

Figure 6.1: Observed and expected 95 % CL upper limits on $\sigma\,\mathcal{B}(\phi \to \tau\tau)$ for ggϕ and bbϕ production. The limits for ggϕ production are shown in the left panel and the limits for bbϕ production in the right panel. The observed limit is indicated with black markers. The solid black line represents the linear extrapolation between the masses for which the limits are derived. The dashed black line indicates the median expected limit in the absence of signal. The green and yellow bands indicate the 68 % and 95 % central confidence intervals for the expected limits. The blue marker indicates the prediction for $\sigma_{\mathrm{A}}\,\mathcal{B}(\mathrm{A} \to \tau\tau) + \sigma_{\mathrm{H}}\,\mathcal{B}(\mathrm{H} \to \tau\tau)$ in the M_{h}^{125} scenario for $m_{\mathrm{A}} = 1.2\,\mathrm{TeV}$ and $\tan\beta = 17$.

from the use of the DeepTau discriminants to identify τ_h candidates and the single τ_h triggers in addition to the previously used triggers.

Two local excesses over the background prediction are observed in ggϕ production. These are located at $m_\phi = 130\,\text{GeV}$ and $m_\phi = 1.2\,\text{TeV}$ and exhibit local significances of 2.6 standard deviations (s.d.) and 2.8 s.d., respectively. These local significances correspond to global significances of 1.7 s.d. and 2.0 s.d.. The global significances are derived from an ensemble test [30] described in more detail in Appendix B.3. The local p-values quantifying the compatibility of the SM predictions including H(125) with the data as a function of m_ϕ are shown in Appendix B.2. The best fit value of $\sigma_{gg\phi}\,\mathcal{B}(\phi \to \tau\tau)$ is $5.4^{+2.0}_{-1.7}\,\text{pb}$ at $m_\phi = 130\,\text{GeV}$ and $3.2^{+1.2}_{-0.9}\,\text{fb}$ at $m_\phi = 1.2\,\text{TeV}$.

For both excesses the consistency of the best fit value of $\sigma_{gg\phi}\,\mathcal{B}(\phi \to \tau\tau)$ across final states and data-taking years is verified at the 95 % CL. The p-values amount to 64 % and 75 % for the excess at $m_\phi = 130\,\text{GeV}$ and 63 % and 11 % at $m_\phi = 1.2\,\text{TeV}$ for the compatibility of the results across data-taking years and final states, respectively. Comparisons of the fit results for $\sigma_{gg\phi}\,\mathcal{B}(\phi \to \tau\tau)$ across final states and data-taking periods to the combined best fit result at $m_\phi = 130\,\text{GeV}$ and $m_\phi = 1.2\,\text{TeV}$ are shown in Appendix B.4. A similar search [155] performed on the data set collected by the ATLAS experiment during LHC Run 2 observes an underfluctuation for the mass hypothesis at $m_\phi = 1.2\,\text{TeV}$ for a comparable expected exclusion limit. For bbϕ production, no excess larger than 1 s.d. above the background prediction is observed but in an m_ϕ range from 100 to 300 GeV an underfluctuation is present.

A second representation of the results in the model-independent interpretation is provided in the form of scans of the profile likelihood function in terms of the two parameters of interest $\sigma_{gg\phi}\,\mathcal{B}(\phi \to \tau\tau)$ and $\sigma_{bb\phi}\,\mathcal{B}(\phi \to \tau\tau)$. The 68 and 95 % CL confidence intervals are extracted using the test statistic [156]

$$q(\mu_g, \mu_b) = -2 \ln \frac{\mathcal{L}(n|\mu_g, \mu_b, \hat{\hat{\boldsymbol{\theta}}}(\mu_g, \mu_b))}{\mathcal{L}(n|\hat{\mu}_g, \hat{\mu}_b, \hat{\boldsymbol{\theta}}(\hat{\mu}_g, \hat{\mu}_b))}\,, \tag{6.1}$$

where the shorthand notations μ_g and μ_b have been introduced for $\sigma_{gg\phi}\,\mathcal{B}(\phi \to \tau\tau)$ and $\sigma_{bb\phi}\,\mathcal{B}(\phi \to \tau\tau)$, respectively. In the asymptotic approximation, $q(\mu_g, \mu_b)$ is expected to follow a χ^2 distribution with two degrees of freedom. The 68 and 95 % confidence intervals are derived from $q(\mu_g, \mu_b) = 2.28$ and $q(\mu_g, \mu_b) = 5.99$ corresponding to the respective quantiles of the χ^2 distribution. These confidence intervals and the point corresponding to the best fit value are shown for two example mass points in Figure 6.2. The contours for all other mass points are shown in Appendix B.5. The points corresponding to the best fit values are shown as black crosses. During the scan and in the derivation of the best fit points, the parameters of interest are restricted to positive values. The scans provide information on the sensitivity of the analysis and the correlation of the POIs.

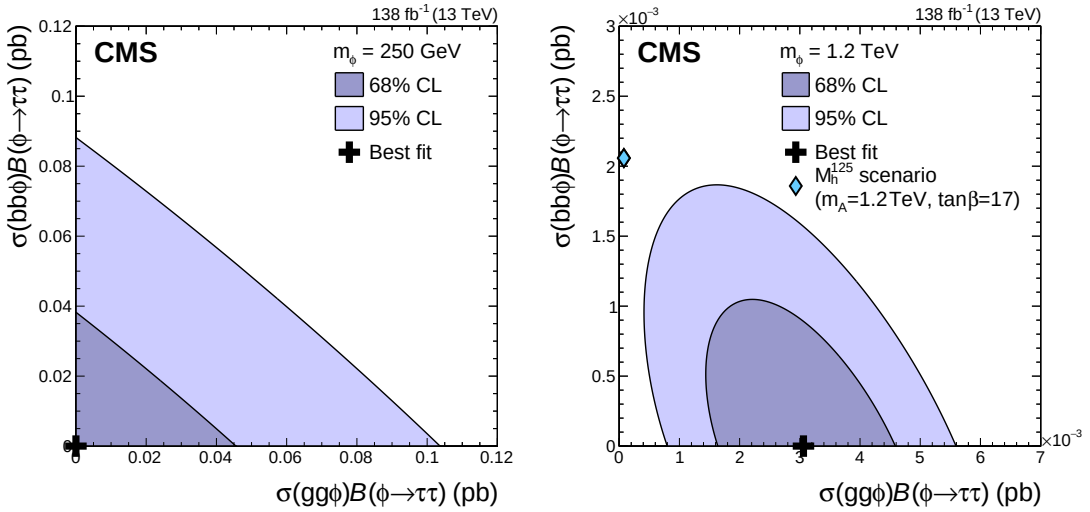

Figure 6.2: Scans of the profile likelihood function in terms of $\sigma_{\mathrm{gg}\phi}\,\mathcal{B}(\phi \to \tau\tau)$ and $\sigma_{\mathrm{bb}\phi}\,\mathcal{B}(\phi \to \tau\tau)$. The 68 and 95 % CL confidence intervals are shown in dark and light purple, respectively. The best fit points are indicated by black crosses. The contours are shown for $m_\phi = 250\,\mathrm{GeV}$ on the left-hand side and $m_\phi = 1.2\,\mathrm{TeV}$ on the right-hand side. In the right panel, the blue marker indicates the prediction for $\sigma_\mathrm{A}\,\mathcal{B}(\mathrm{A} \to \tau\tau) + \sigma_\mathrm{H}\,\mathcal{B}(\mathrm{H} \to \tau\tau)$ in the M_h^{125} scenario for $m_\mathrm{A} = 1.2\,\mathrm{TeV}$ and $\tan\beta = 17$. The left figure is published in [30] and the right figure is a modified version of a figure published in [30].

6.2 Interpretation in MSSM and 2HDM Benchmark Models

In the model-dependent interpretation of the data, exclusion contours are derived within specific MSSM and 2HDM benchmark scenarios. In case of the MSSM benchmark scenarios, the predictions of all required masses, cross sections, and branching fractions are obtained from the calculations provided by the LHCHWG-3 [63, 67]. The predictions for the properties of the Higgs sector in the 2HDM benchmark scenarios are calculated as described in Section 2.3.1.

6.2.1 Interpretation in Selected MSSM Benchmark Scenarios

The 95 % CL observed and expected exclusion contours in the M_h^{125} and $M_{\mathrm{h,EFT}}^{125}$ MSSM benchmark scenarios are presented in Figure 6.3. They are shown as a function of the two free parameters of the benchmark scenarios, m_A and $\tan\beta$. The properties of the two displayed benchmark scenarios are discussed in Section 2.3.2. Exclusion contours in further MSSM benchmark scenarios proposed in [63] are given in Appendix C. The median expected exclusion contour in absence of signal is shown as dashed black line. The 68 and 95 % expected exclusion contours are shown as dark and light gray bands. The excluded parameter space is indicated by a blue filled area. Regions of the parameter space for which m_h does not coincide with $m_{\mathrm{H}(125)}$ within a theoretical uncertainty of

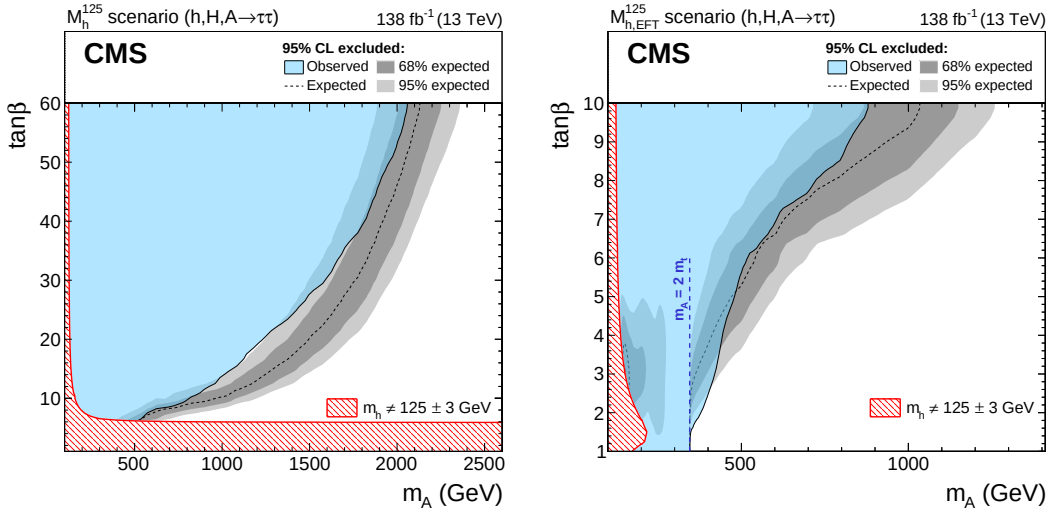

Figure 6.3: Observed and expected 95 % CL exclusion contours in the M_{h}^{125} and $M_{\mathrm{h,\,EFT}}^{125}$ MSSM benchmark scenarios. The M_{h}^{125} scenario is shown on the left and the $M_{\mathrm{h,\,EFT}}^{125}$ scenario on the right. In both cases, the median expected contour and the central 68 and 95 % intervals in absence of signal are indicated by a black dashed line and by dark and light gray bands, respectively. The observed contour is shown as black line and the excluded parameter space indicated by a light blue filled area. Regions of the parameter space for which m_{h} does not coincide with $m_{\mathrm{H(125)}}$ within a theoretical uncertainty of ± 3 GeV are indicated by a red hatched area. In the case of the $M_{\mathrm{h,\,EFT}}^{125}$ scenario, the threshold for the decay of A into two on-shell top quarks at $m_{\mathrm{A}} = 2m_{\mathrm{t}}$ is indicated as dashed blue line. Figures published in [30].

$\pm 3\,\text{GeV}$ are indicated by red hatched areas. In both shown benchmark scenarios, this corresponds to $m_\text{h} < 122\,\text{GeV}$.

The course of the exclusion contours is determined by the properties of the MSSM in the corresponding scenario and the kinematic properties of the τ leptons in the events. With rising m_A also m_H increases and the cross section for the production of the corresponding Higgs bosons decreases. As the cross sections fall off faster than the sensitivity of the analysis increases for large Higgs boson masses, a threshold value is reached above which no further mass point can be excluded. In addition, decays of the heavy Higgs bosons into SUSY particles are kinematically allowed above $m_\text{A} \approx 2\,\text{TeV}$ which reduces the branching fraction of the heavy Higgs bosons to tau leptons. The decoupling limit is reached for a large fraction of the allowed parameter space in the M_h^{125} and $M_\text{h,EFT}^{125}$ benchmark scenarios. Thus, the cross section for $\text{bb}\phi$ and $\text{gg}\phi$ production as well as the branching fractions for decays of the heavy Higgs bosons to tau leptons increase with increasing values of $\tan\beta$. This leads to a larger range of excluded m_A values for large values of $\tan\beta$.

In the M_h^{125} benchmark scenario, m_A values below $500\,\text{GeV}$ are excluded over the full $\tan\beta$ range allowed by the constraint on the prediction of m_h. The range of excluded values of $\tan\beta$ decreases with increasing values of m_A. Above $m_\text{A} \approx 2050\,\text{GeV}$, no further parameter points can be excluded. In general, the observed exclusion contour is looser than the median expected and in the range from $m_\text{A} \approx 1.1\,\text{TeV}$ to $m_\text{A} \approx 1.3\,\text{TeV}$, the observed exclusion limit lies outside the 95 % expected limit. This deviation corresponds to the excess already seen in the exclusion limits for $\text{gg}\phi$ production in the model-independent interpretation at $m_\phi = 1.2\,\text{TeV}$. The excess is located in the No b tag categories, whereas in the b tag categories no excess is observed. For the $\tan\beta$ values in between the median expected and observed exclusions, the additional Higgs bosons in the M_h^{125} benchmark scenario are dominantly produced via $\text{bb}\phi$ production. In particular, for $\tan\beta = 17$, $\text{bb}\phi$ production accounts for more than 90 % of the expected signal from the additional Higgs bosons in the No b tag categories as illustrated in Figure 5.3. Thus, the observed excess is not compatible with an MSSM Higgs boson signal. This incompatibility is also illustrated in Figure 6.1 and Figure 6.2, where the M_h^{125} scenario prediction for $m_\text{A} = 1.2\,\text{TeV}$ and $\tan\beta = 17$ is shown in comparison to the obtained results.

The exclusion contours in the $M_\text{h,EFT}^{125}$ benchmark scenario cover a more restricted parameter space than the ones in the M_h^{125} benchmark scenario. The excluded parameter space ranges from an exclusion of the full parameter space for m_A values below $345\,\text{GeV}$ to values around $m_\text{A} \approx 900\,\text{GeV}$ where values of $\tan\beta \geq 10$ are excluded. In this region, the onset of the excess, as discussed already for the M_h^{125} scenario, is also visible. At $m_\text{A} = 345\,\text{GeV}$, the observed and all expected contours coincide for $\tan\beta < 2$. The excluded values of $\tan\beta$ rise steeply above this threshold. This observation originates from the leading role of the A boson produced via gluon fusion in this region, with cross section and branching fraction predictions exceeding the corresponding quantities

of the H boson. Above the kinematic threshold for $A \to tt$ at $m_A = 2m_t = 345\,\text{GeV}$, the branching fraction for $A \to \tau\tau$ drops instantly. Consequently, a further exclusion of the parameter space, which is fully excluded below the threshold, is not possible for low values of $\tan\beta$.

Another consequence of the gluon fusion dominated sensitivity in the low m_A-low $\tan\beta$ region, is an island of reduced sensitivity for $2.5 < \tan\beta < 4$. For low values of m_A, the product of the cross section and branching fraction for A boson production decreases with increasing values of $\tan\beta$ before it increases again above $\tan\beta \approx 6$. The $bb\phi$ production mode exhibits a sizable contribution to the sensitivity starting at intermediate values of $\tan\beta$, leading to a minimum of the total prediction of the product of cross section and branching ratio for the A and H bosons in the region, where the island is observed. Additionally, the transition from the high-mass No b tag categories to the categories based on the NN discriminant at $m_{\tau\tau} = 250\,\text{GeV}$ reduces the sensitivity to the additional Higgs bosons in the mass range from 200 to 250 GeV. A crossing of the observed and median expected exclusion contours is present at $m_A \approx 500\,\text{GeV}$. Below this value, the observed limit is more restrictive than the expected and less restrictive above. In particular, for m_A values between 350 and 400 GeV, the observed limit lies below the 95 % expected limit.

For values of $\tan\beta \lesssim 6$, the radiative corrections to m_h in the M_h^{125} scenario are not large enough to raise the tree-level prediction to values above 122 GeV. As discussed already in Section 2.3.2, the $M_{h,\text{EFT}}^{125}$ employs similar SUSY parameters but opens up the low-$\tan\beta$ region for investigation again through adjustments of M_S to larger values up to 10^{16} GeV as function of m_A and $\tan\beta$ [65]. The $M_{h,\text{EFT}}^{125}$ scenario can consequently be viewed as low-$\tan\beta$ extension of the M_h^{125} scenario [63]. In contrast to the M_h^{125} scenario, the large logarithms in the calculation of m_h arising from large values of M_S are resummed in a 2HDM effective field theory. A combination of the results in both benchmark scenarios in a single figure is shown in Figure 6.4. It includes the exclusion contours in the M_h^{125} benchmark scenario down to values of $\tan\beta = 10$, below which the exclusion contours from the $M_{h,\text{EFT}}^{125}$ scenario are shown. Even though the exclusion contours take on similar values in the transition region, a small offset is visible. It originates from the predictions of the production cross sections for the heavy neutral Higgs bosons, which are 10–30 % and 20 % larger in the $M_{h,\text{EFT}}^{125}$ scenario in gluon fusion and b quark associated production at $\tan\beta = 10$. The increased cross section predictions in the $M_{h,\text{EFT}}^{125}$ scenario are a direct consequence of the smaller reduction of the effective b quark Yukawa coupling introduced via the Δ_b terms [157, 158] with respect to the Yukawa couplings in a general 2HDM.

6.2.2 Interpretation in Selected 2HDM scenarios

The 95 % CL observed and expected exclusion contours in the 2HDM benchmark scenarios introduced in Section 2.3.1 are shown in Figure 6.5. For each benchmark scenario, they are presented for Type I (left column) and Type II (right column) coupling structures in the corresponding Yukawa sectors. The contours are derived as a function of the two parameters, which are conventionally scanned in the different benchmark scenarios.

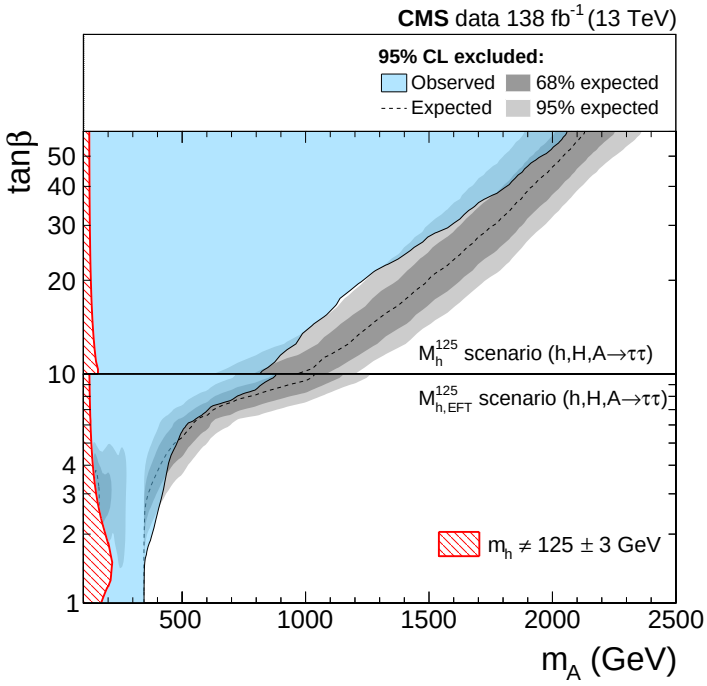

Figure 6.4: Combined view of the observed and expected 95 % CL exclusion contours in the M_{h}^{125} and $M_{\mathrm{h,\,EFT}}^{125}$ MSSM benchmark scenarios. The contours for the M_{h}^{125} and $M_{\mathrm{h,\,EFT}}^{125}$ scenarios are shown in the middle and lower panel, respectively. In both cases, the median expected contour and the central 68 and 95 % intervals in absence of signal are indicated by a black dashed line and by dark and light gray bands, respectively. The observed contour is shown as black line and the excluded parameter space indicated by a light blue filled area. Regions of the parameter space for which the predicted mass of the observed Higgs boson is outside the theoretical uncertainty of $\pm 3\,\mathrm{GeV}$ around the value of $m_{\mathrm{h}} = 125\,\mathrm{GeV}$ are indicated by red hatched areas.

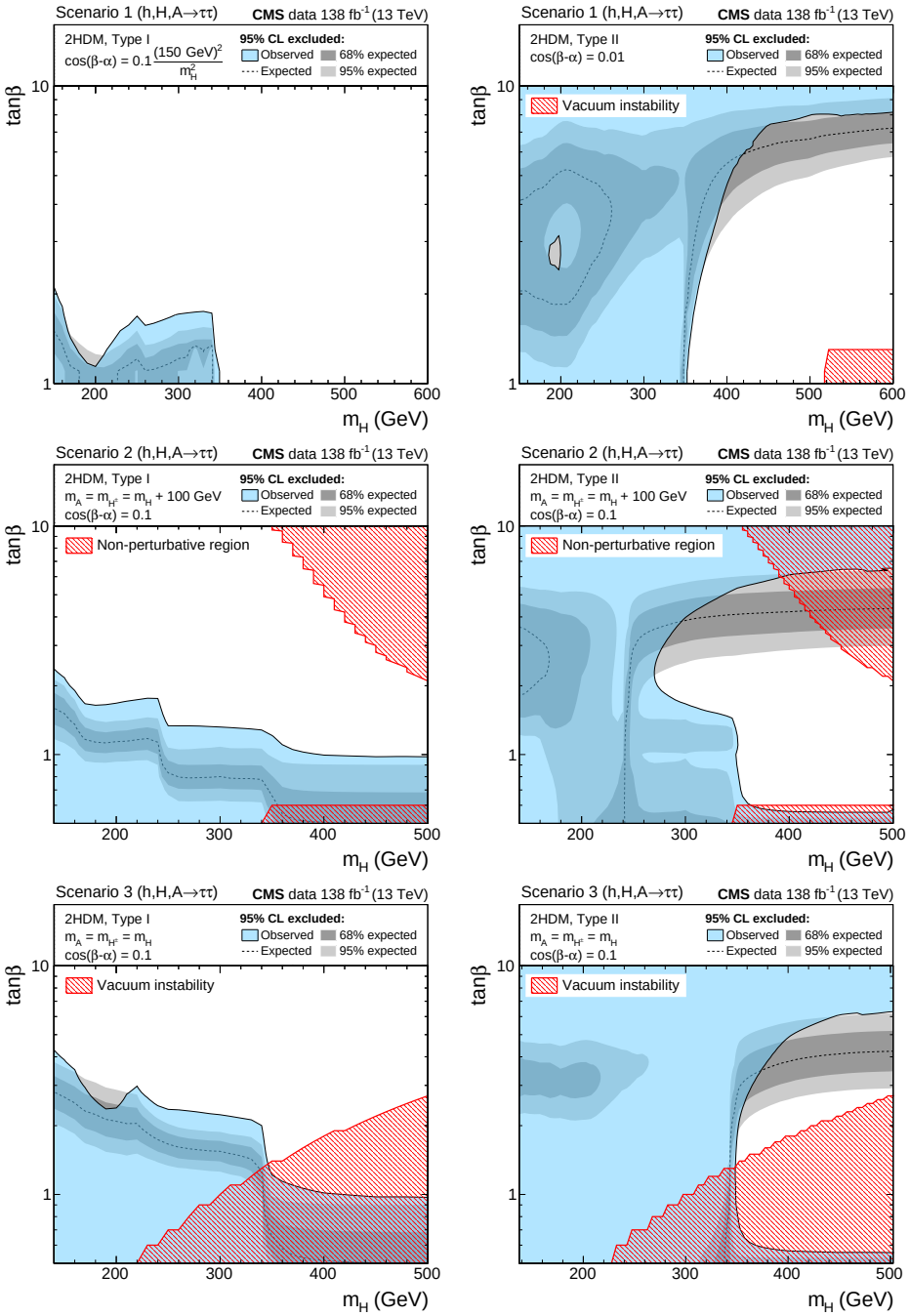

Figure 6.5: Observed and expected 95 % CL exclusion contours in selected 2HDM benchmark scenarios. Scenarios exhibiting Type I (Type II) couplings in the corresponding Yukawa sector are shown in the left (right) column. The median expected contour and the central 68 and 95 % intervals in absence of signal are indicated by a black dashed line and by dark and light gray bands, respectively. The observed contour is is shown as a black line and the excluded parameter space indicated by a light blue filled area. Regions of the parameter space for which the Higgs potential does not fulfill perturbativity or vacuum stability constraints are indicated by red hatched areas. The steps at the edges of these areas originate from the granularity of the parameter grid for which the predictions are calculated.

These are chosen to be m_{H} and $\tan\beta$. The observed and expected contours are shown following the same conventions as in the case of the MSSM benchmark scenarios. Those regions of the parameter space already excluded from constraints on perturbativity or vacuum stability of the Higgs potential are indicated by red hatched areas.

Both the M_{h}^{125} and $M_{\mathrm{h,EFT}}^{125}$ MSSM benchmark scenarios effectively correspond to 2HDMs of Type II with couplings constrained from SUSY as the phenomenology of the models for lower masses of the additional Higgs bosons is not directly influenced by SUSY particles. Consequently, the exclusion contours in the investigated 2HDMs with Type II Yukawa couplings exhibit similar features to the exclusion limits in the $M_{\mathrm{h,EFT}}^{125}$ benchmark scenario, which predicts $m_{\mathrm{h}} = 125 \pm 3\,\mathrm{GeV}$ for the considered $\tan\beta$ values. As such the kinematic threshold for the decay of the heavy neutral Higgs bosons into top quark pairs at $m_{\mathrm{H/A}} = 2m_{\mathrm{t}} = 345\,\mathrm{GeV}$ still constitutes an important feature of the exclusion contours in all considered 2HDM benchmark scenarios. Below this threshold, the full considered parameter space can be excluded. Above, the exclusion limits rise sharply until they reach values of $\tan\beta \approx 8$ in Scenario 1 and $\tan\beta \approx 6$ in Scenarios 2 and 3 and do not rise much further for increasing values of m_{H}. The median expected limit shows a behavior similar to the observed limit but reaches its plateau value already at $\tan\beta \approx 4$ in Scenarios 2 and 3. The observed exclusion contour in Scenario 2 shows two milder thresholds located where first m_{A} and then m_{H} reach the threshold value.

Two properties of the benchmark scenarios lead to the shape of the exclusion contours. Firstly, the $\tan\beta$-enhancement of the Higgs boson couplings to down-type fermions raises the branching fraction to tau leptons again with increasing values of $\tan\beta$. Secondly, the predicted cross section for ggh production and the branching fraction for h decays into tau leptons decreases with increasing values of $\tan\beta$. This effect is larger in Scenarios 2 and 3 because $\cos(\beta - \alpha)$ takes a higher value with respect to Scenario 1. It also allows the explanation of the larger difference between the observed and median expected exclusion contours through the measured value of 0.82 ± 0.11 for the inclusive signal strength of H(125) in the categories used for $m_{\tau\tau} < 250\,\mathrm{GeV}$ [15]. As in the case of the $M_{\mathrm{h,EFT}}^{125}$ scenario, the island of low sensitivity for low values of m_{H} respectively m_{A} and $2 \lesssim \tan\beta \lesssim 4$ is present to different extents in the three 2HDM scenarios.

In all three considered benchmark scenarios, a smaller parameter space is excluded for the manifestations with Type I Yukawa couplings. This lower sensitivity relates to the enhancement of the heavy Higgs boson couplings to down-type fermions for larger values of $\tan\beta$ in Type II models, which is not present in models of Type I. In the latter, not only the branching fraction for decays into tau leptons but also the $\mathrm{gg}\phi$ and $\mathrm{bb}\phi$ production cross sections decrease with increasing values of $\tan\beta$. Therefore only low values of $\tan\beta$ may be excluded in these models.

The highest sensitivity to Type I manifestations is achieved in Scenario 3, where $\tan\beta$ values up to $\tan\beta \approx 4$ are excluded for low values of m_{H} with smaller $\tan\beta$ ranges excluded with increasing m_{H}. In the other two scenarios, exclusions up to $\tan\beta \approx 2$ are

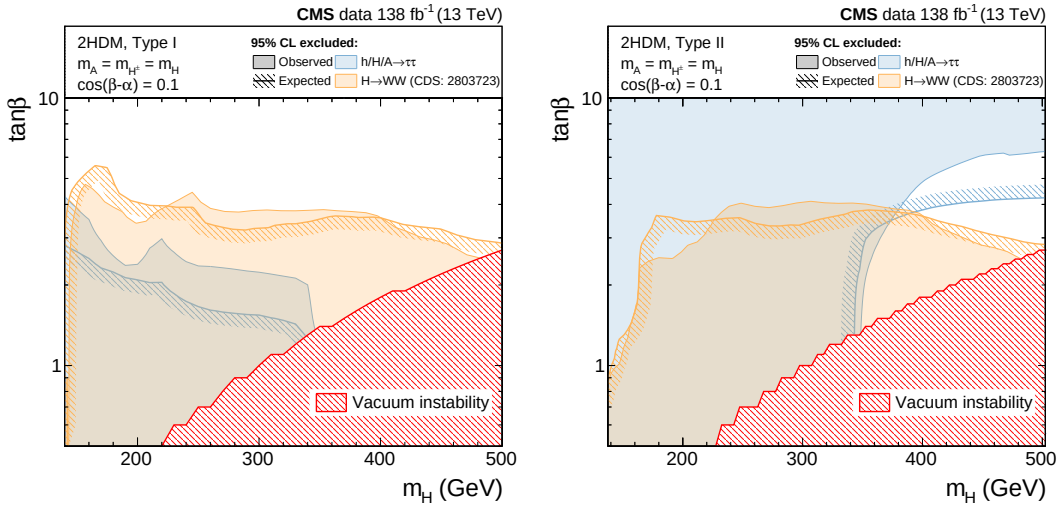

Figure 6.6: Comparison of the observed and expected 95 % CL exclusion contours in the 2HDM benchmark Scenario 3 from Higgs boson decays to tau leptons or W bosons. The exclusion contours for the Type I (Type II) manifestation of the scenario are shown in the left (right) panel. In both cases the observed and median expected contours are shown as filled areas and solid lines with hatches indicating the excluded regions, respectively. The red hatched area corresponds to the area of the parameter space in which the Higgs potential does not fulfill the vacuum stability constraints. The exclusion contours from the search for additional neutral Higgs bosons in decays to W bosons are taken from [61].

possible for low values of m_{H}. The kinematic thresholds for $\mathrm{H/A} \to \mathrm{tt}$ decays are also visible in the form of steps, at which the observed and expected exclusion contours drop to lower values of $\tan\beta$. The effect of the kinematic threshold on the exclusion limits is most prominently observed for Scenarios 3, in which both additional neutral Higgs bosons exhibit the same mass. Above the threshold, the exclusion contours in Scenarios 2 and 3 become comparable and $\tan\beta$ values below $\tan\beta \approx 1$ are excluded over the full considered mass range in both scenarios. In Scenario 1, no further exclusion is observed above the threshold. For all scenarios, the observed exclusion contour rules out a larger parameter space than the median expected contour.

A comparison of the exclusion contours in Scenario 3 to the exclusion contours obtained from the published search for additional neutral Higgs bosons in the decay to W bosons [61] is shown in Figure 6.6. The search for additional neutral Higgs bosons in decays to W bosons excludes parameter points up to values of $\tan\beta = 5$ for low m_{H} and most of the parameter space below $\tan\beta = 3$ in the shown mass range. The exclusion contours show almost the same dependence of the exclusion region on the parameters for both types of considered Yukawa couplings. In both searches, the parameter space corresponding to $\tan\beta$ values larger than $\tan\beta \approx 4$ can not be excluded. For the Type II manifestation, the $\tau\tau$ final state provides complementary information with respect to the search in decays to W bosons. It excludes the parameter space starting from larger values

of $\tan\beta$, while the exclusion contours from decays to W bosons restrict the parameter space for low values of $\tan\beta$. Above $m_{\mathrm{H}} = 380\,\mathrm{GeV}$ a parameter space with intermediate values of $\tan\beta$ opens up, which is not excluded by either search.

CHAPTER 7

Summary

Despite its considerable success in describing the strong and electroweak interactions, the standard model of particle physics (SM) is generally believed to be a low-energy effective theory of a more extended theory. Many of the proposed extensions to the SM also include an extended Higgs sector which manifests itself in additional scalar bosons and modified couplings of the observed Higgs boson (H(125)) with respect to the SM expectation. One of the mostly discussed extensions of the SM is the minimal supersymmetric SM (MSSM) predicting two additional neutral and two charged Higgs bosons. Decays of the neutral Higgs bosons into pairs of tau leptons provide the principal experimental probe for the extended Higgs sector of the MSSM.

The results presented in this thesis demonstrate the experimental sensitivity of the analysis of $\tau\tau$ final states in two interpretations of the complete data set collected by the CMS experiment during the LHC Run 2 data-taking period. The analyzed data set corresponds to an integrated luminosity of $138\,\text{fb}^{-1}$. In a first model-independent interpretation a search for further scalar resonances ϕ in addition to H(125) is performed. The model-independent search is performed separately for the gluon fusion ($\text{gg}\phi$) and b quark associated production ($\text{bb}\phi$) modes of ϕ. In the second interpretation, the exclusion limits are translated to exclusion contours in selected MSSM and 2HDM benchmark scenarios.

The observed exclusion limits on $\sigma\mathcal{B}(\phi \to \tau\tau)$ at the 95 % confidence level reach from values in the order of 10 pb at the lowest masses to approximately 0.3 fb at the highest masses for both production modes. Expressed in terms of the MSSM benchmark scenarios, values of m_A below 350 GeV are excluded in both considered benchmark scenarios over the full $\tan\beta$ range. For large values of $\tan\beta$, the exclusion limits on m_A reach up to values above 2 TeV. In the considered 2HDM scenarios, values of $\tan\beta$ above $\tan\beta = 8$ are excluded in all scenarios with Type II Yukawa couplings, while for the scenarios with Type I Yukawa couplings only a small fraction of the parameter space at low values of $\tan\beta$ can be excluded.

Two local excesses located at $m_\phi = 130\,\text{GeV}$ and $m_\phi = 1.2\,\text{TeV}$ are observed in $\text{gg}\phi$ production, while in $\text{bb}\phi$ production no excess above the background prediction is ob-

served. The local significances of the observed excesses correspond to 2.7 s.d. and 2.8 s.d. at $m_\phi = 130\,\text{GeV}$ and $m_\phi = 1.2\,\text{TeV}$, respectively. The parameter space of the MSSM benchmark scenarios corresponding to the resonance at $m_\phi = 130\,\text{GeV}$ is excluded in all considered benchmark scenarios. For $m_\phi = 1.2\,\text{TeV}$, the observed excess is not compatible with predictions from the investigated MSSM benchmark scenarios since the additional neutral Higgs bosons are predicted to be predominantly produced via $\text{bb}\phi$ production.

The results of the presented analysis will likely represent the most stringent direct exclusions on the MSSM parameter space from data collected with the CMS experiment until the end of the LHC Run 3 data-taking period in 2025.

APPENDIX A

Postfit Distributions

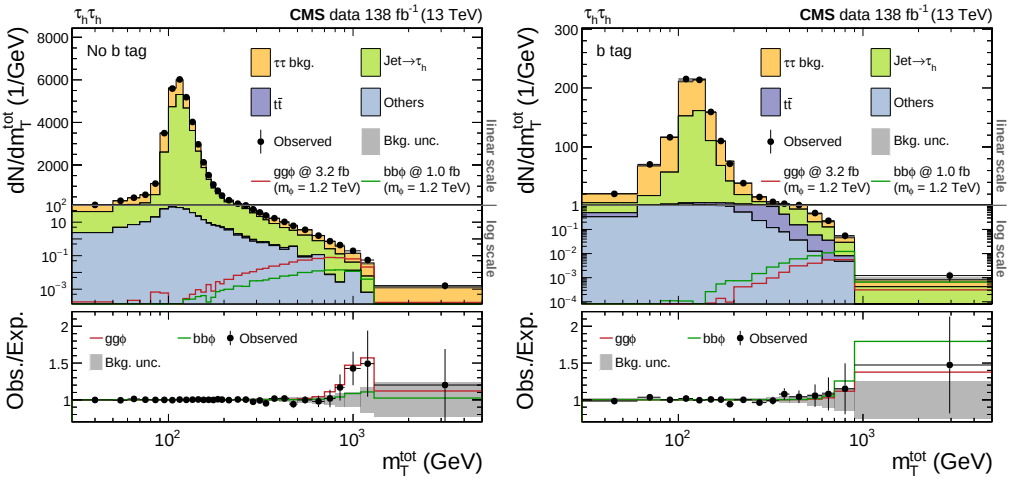

Figure A.1: Distributions of the total transverse mass $m_{\mathrm{T}}^{\mathrm{tot}}$ in the No b tag (left) and b tag (right) categories in the $\tau_{\mathrm{h}}\tau_{\mathrm{h}}$ final state. The distributions are shown combined over the three data-taking years after the fit of the background model to the data. The figures are modified versions of figures published in [30].

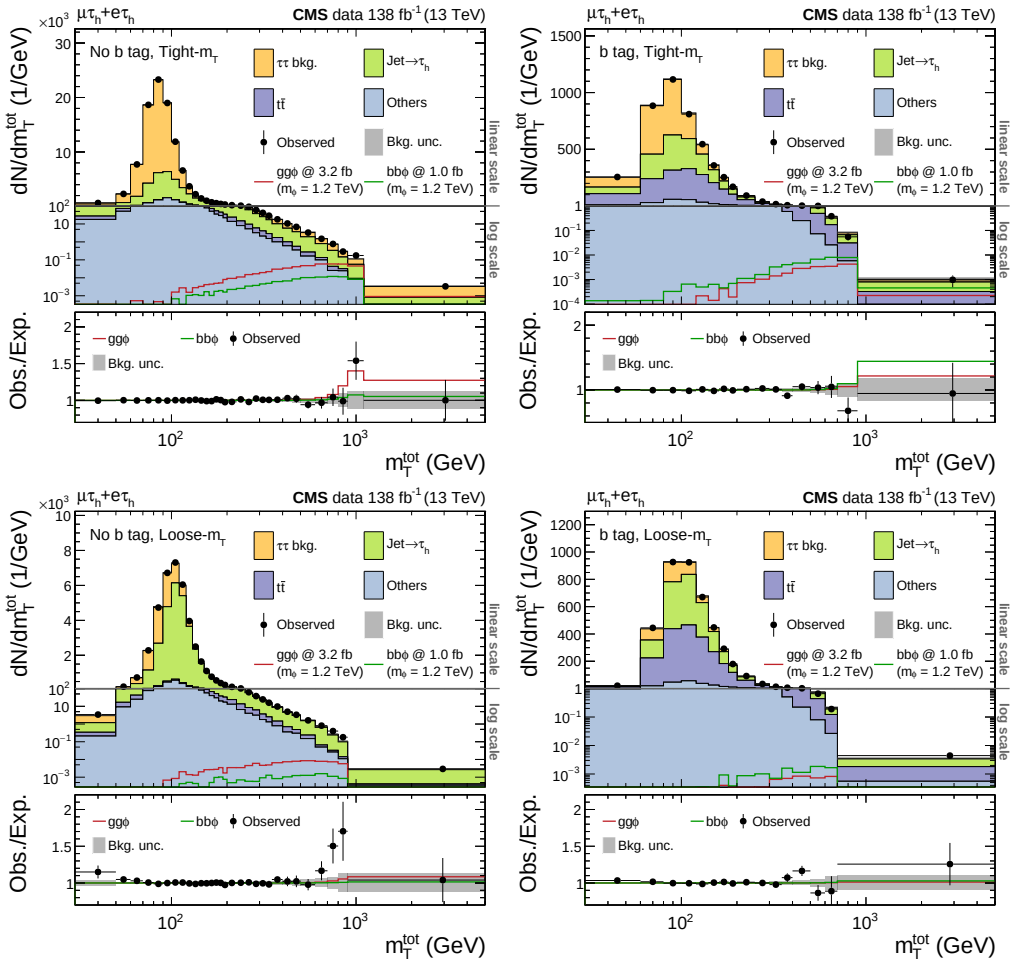

Figure A.2: Distributions of the total transverse mass $m_{\mathrm{T}}^{\mathrm{tot}}$ in the $\mu\tau_{\mathrm{h}}$ and $\mathrm{e}\tau_{\mathrm{h}}$ final states. The distributions are shown for the No b tag,Tight-m_{T} (upper left), b tag,Tight-m_{T} (upper right), No b tag,Loose-m_{T} (lower left), and b tag Loose-m_{T} (lower right) categories. They are shown combined over the three data-taking years after the fit of the background model to the data. The figures are modified versions of figures published in [30].

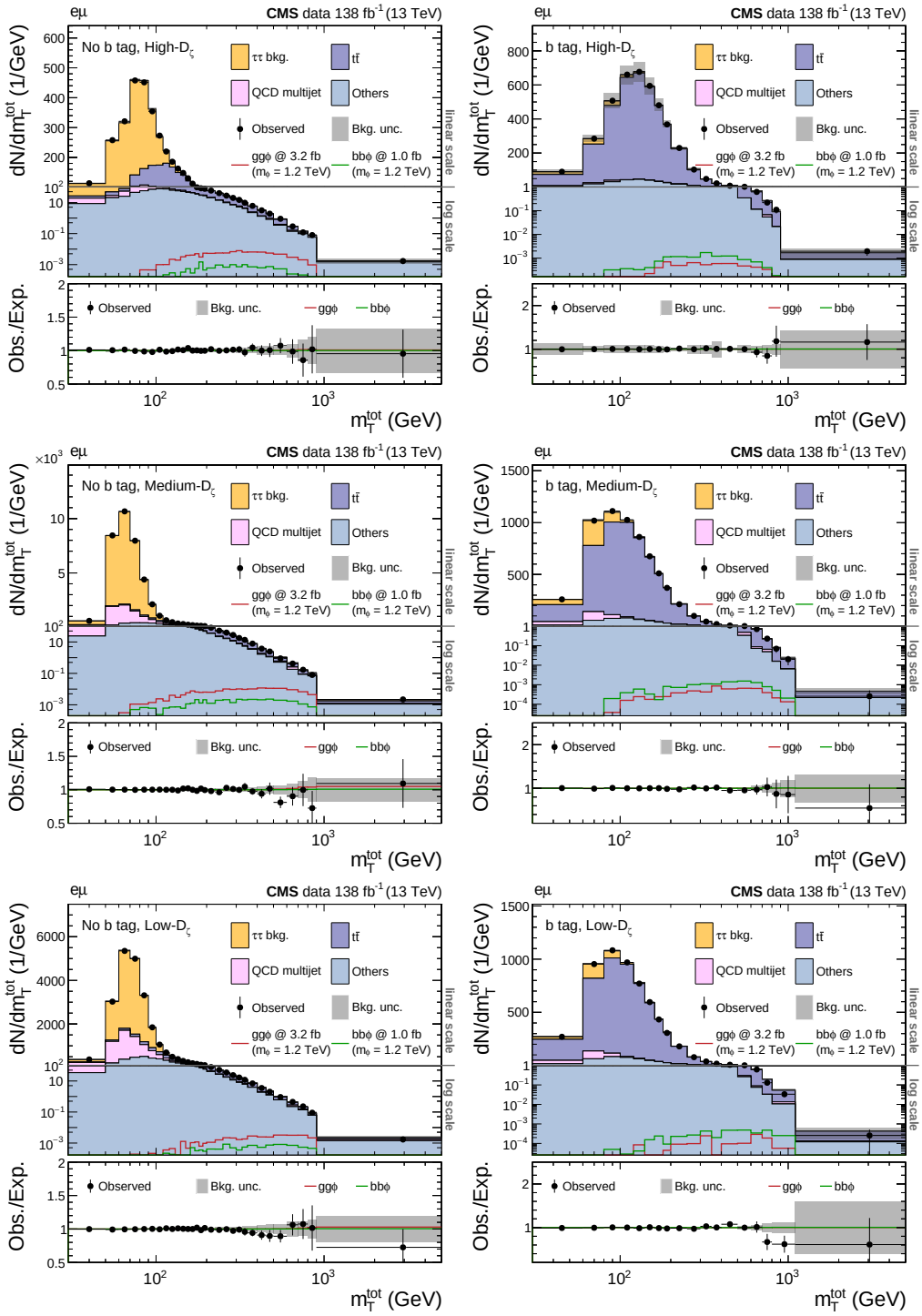

Figure A.3: Distributions of the total transverse mass $m_{\mathrm{T}}^{\mathrm{tot}}$ in the eμ final state. The distributions are shown for the global No b tag and b tag categories in the left and right columns, respectively, and for the Low-D_ζ, Medium-D_ζ, and High-D_ζ categories from top to bottom. They are shown combined over the three data-taking years after the fit of the background model to the data. The figures are modified versions of figures published in [30].

APPENDIX B

Additional Material on Model-Independent Exclusion Limits

B.1 Breakdown of the Exclusion Limits

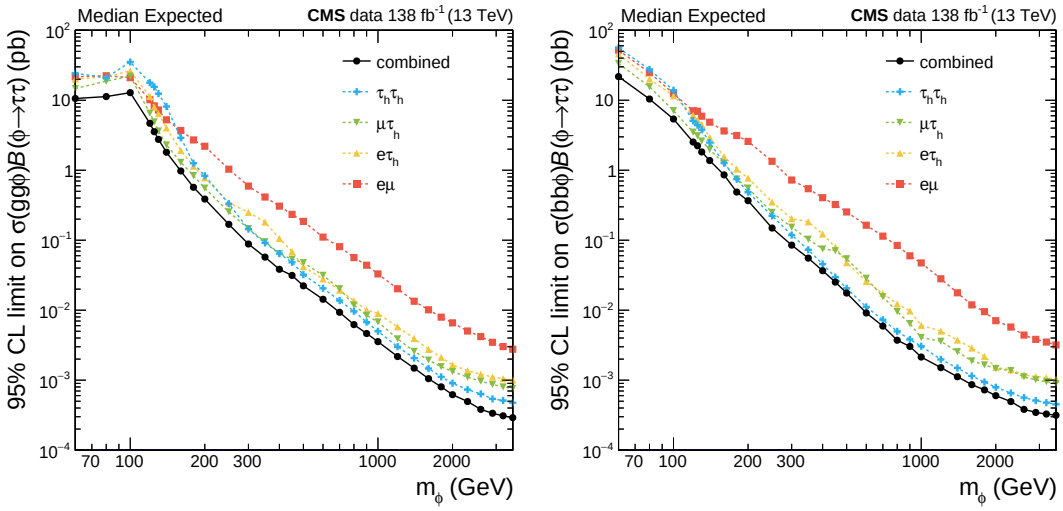

Figure B.1: Breakdown of the median expected 95 % CL upper limits on $\sigma\,\mathcal{B}(\phi \to \tau\tau)$ in absence of signal for ggϕ and bbϕ production in the analyzed final states. The limits on ggϕ production are shown in the left panel; the limits on bbϕ production are shown in the right panel.

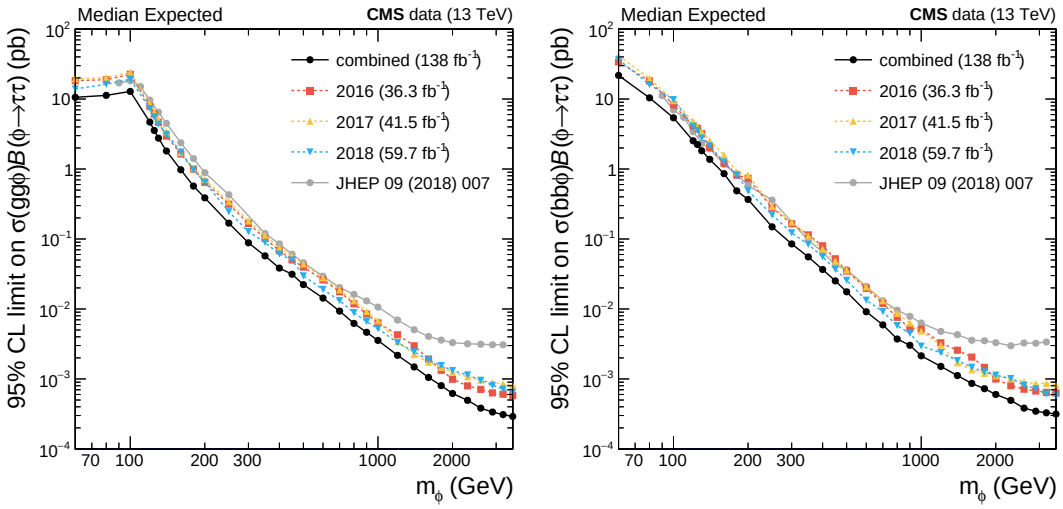

Figure B.2: Breakdown of the median expected 95 % CL upper limits on $\sigma\,\mathcal{B}(\phi \to \tau\tau)$ in absence of signal for ggϕ and bbϕ production in the analyzed data-taking years. The limits on ggϕ production are shown in the left panel; the limits on bbϕ production are shown in the right panel. The median expected upper exclusion limits from the previously published analysis of the data taken with the CMS experiment in 2016 [31] are included for comparison and indicated by gray markers.

B.2 Local p-values

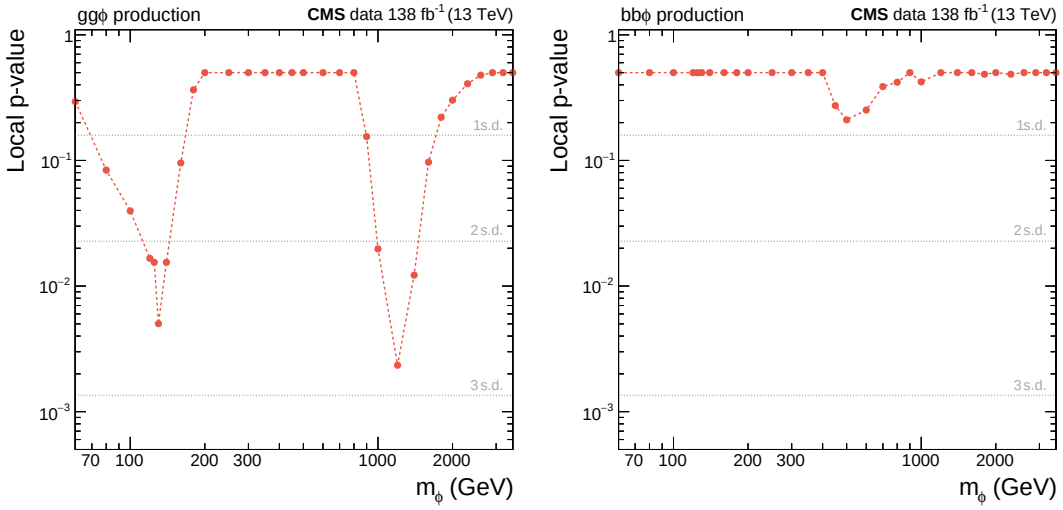

Figure B.3: Local p-values for the compatibility of the SM prediction including H(125) with the data as a function of m_ϕ. The local p-values are shown separately for ggϕ (left) and bbϕ (right) production. Local p-values corresponding to significances of 1, 2, and 3 s.d. are indicated by gray dotted lines.

B.3 Global p-value Calculation

The global p-values corresponding to the two observed local excesses for ggϕ production of the additional resonance at $m_\phi = 130\,\text{GeV}$ and $m_\phi = 1.2\,\text{TeV}$ are derived from an ensemble test [30]. For this ensemble test, 8000 background-only toy data sets are generated from the distributions of m_T^tot in all categories after a fit to the data. The maximum local significance across all considered mass hypotheses is then calculated for each generated toy data set. Finally, the global p-value is calculated as the ratio of the number of toy data sets for which the maximum local significance is larger than the observed local significance over the total number of generated toy data sets and translated to the global significance. The calculation of the global p-value is illustrated in Figure B.4. The obtained global significances are

$$\sigma^{\text{glob.}}_{130\,\text{GeV}} = 1.72^{+0.03}_{-0.02}\,\text{s.d.} \tag{B.1}$$

$$\sigma^{\text{glob.}}_{1.2\,\text{TeV}} = 2.05^{+0.03}_{-0.03}\,\text{s.d.}\,. \tag{B.2}$$

They correspond to the probability to observe a more significant local excess for ggϕ production of the additional resonance for any of the probed mass hypotheses than the respective observed excess.

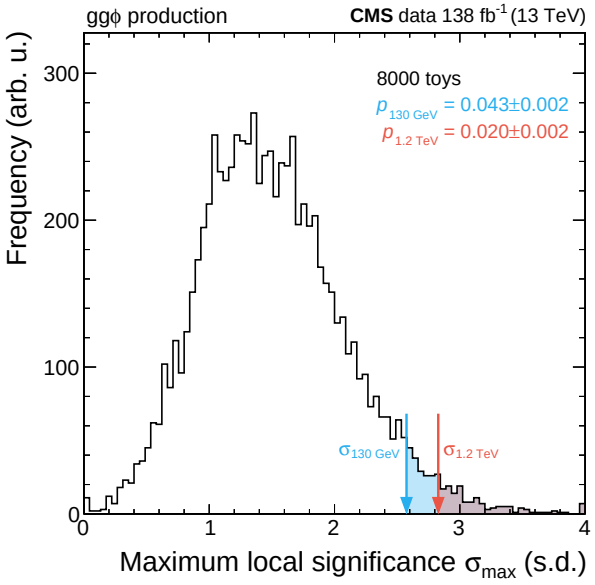

Figure B.4: Illustration of the calculation of the global p-value for the observed excesses. The distribution of the maximum local significance for each generated toy data set is indicated by a black line. Values of the maximum local significance exceeding the shown range are collected in the last shown bin. The local significances of the observed excesses at $m_\phi = 130\,\text{GeV}$ and $m_\phi = 1.2\,\text{TeV}$ are indicated by blue and red arrows, respectively. The areas of the distribution to which the global p-values correspond are indicated by blue and red filled areas for the excesses at $m_\phi = 130\,\text{GeV}$ and $m_\phi = 1.2\,\text{TeV}$.

B.4 Channel Compatibility of the Excesses

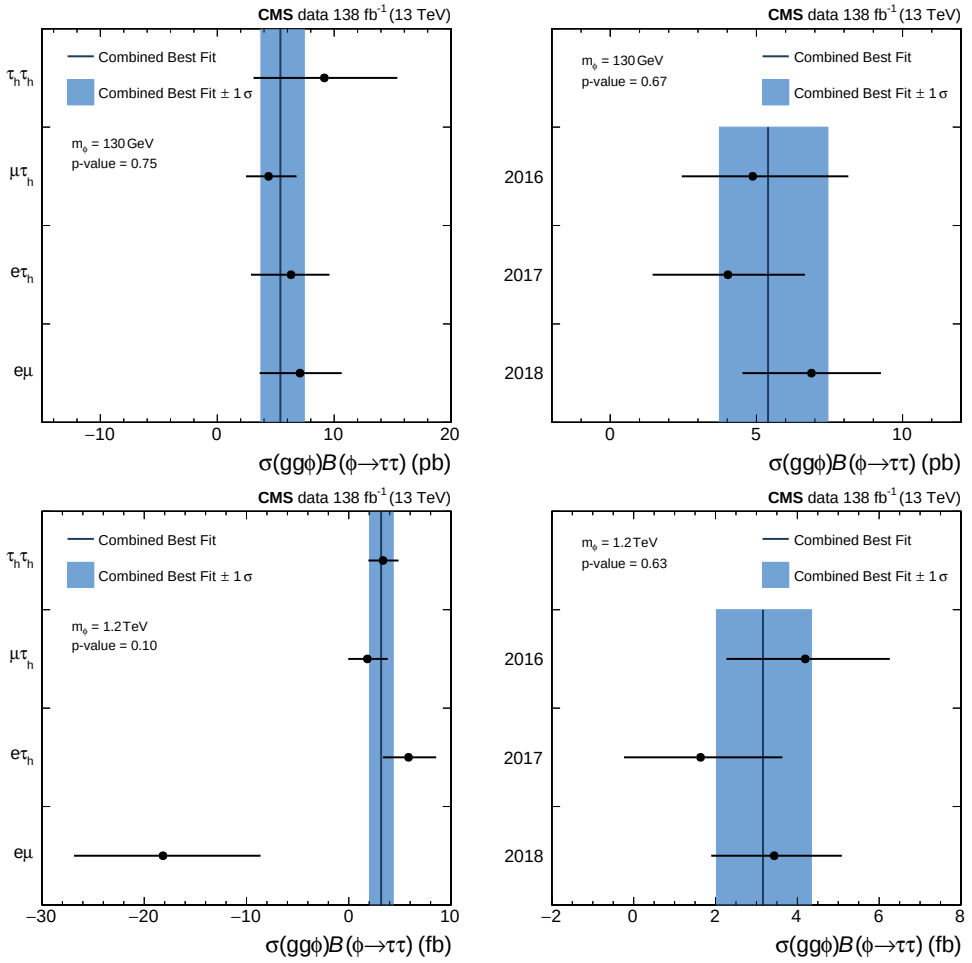

Figure B.5: Comparison between the best fit result of the combined fit of $\sigma_{gg\phi}\,\mathcal{B}(\phi \to \tau\tau)$ and the results of a fit with individual POIs for every final state or data-taking year. The comparisons for individual POIs for every final state (data-taking year) are shown in the left (right) panels. The comparisons are shown for $m_\phi = 130\,\text{GeV}$ and $m_\phi = 1.2\,\text{TeV}$ in the upper and lower panels, respectively. The p-values indicating the compatibility of the combined fit to the fit with individual POIs, given in the panels, are obtained from ensemble tests comprising 950 toys each. The figures in the lower row are modified versions of figures published in [30].

B.5 Scans of the Profile Likelihood Function

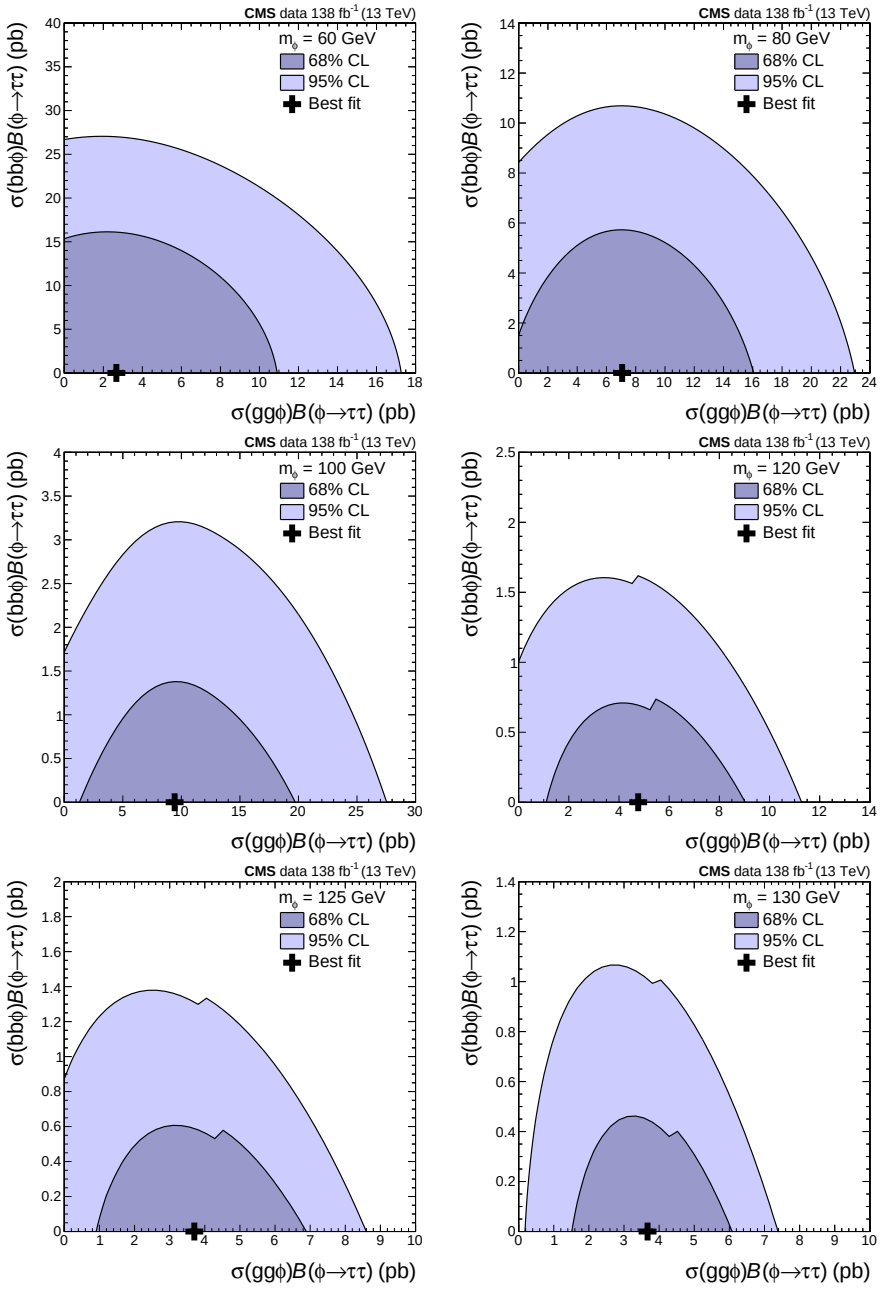

Figure B.6: Scans of the profile likelihood function in terms of $\sigma_{\mathrm{gg}\phi}\,\mathcal{B}(\phi \to \tau\tau)$ and $\sigma_{\mathrm{bb}\phi}\,\mathcal{B}(\phi \to \tau\tau)$. The 68 and 95 % CL contours are shown in dark and light purple, respectively. The best fit points are indicated by black crosses. The contours are shown for $m_\phi = 60,\ 80,\ 100,\ 120,\ 125,$ and 130 GeV from left to right and top to bottom.

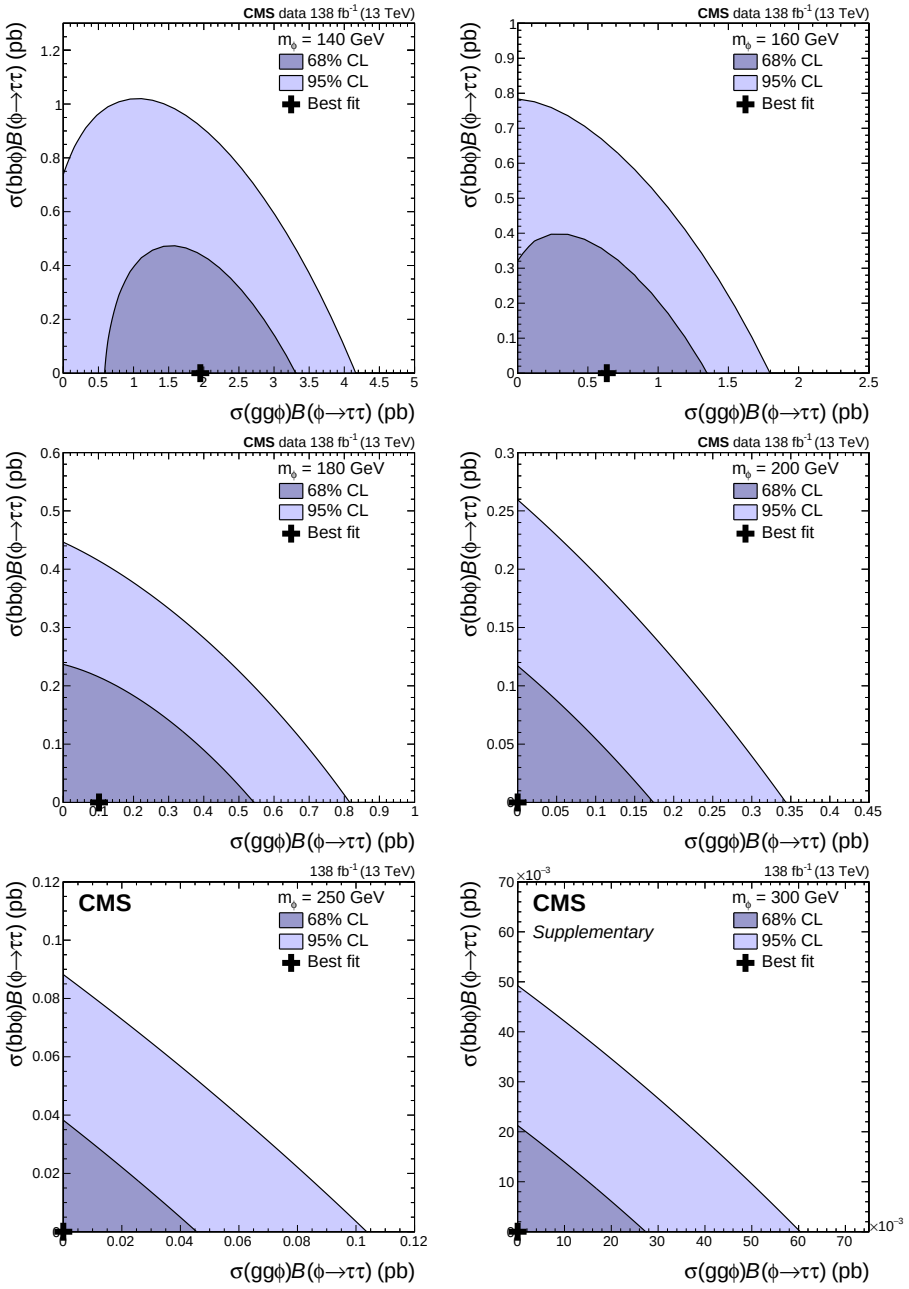

Figure B.7: Scans of the profile likelihood function in terms of $\sigma_{\mathrm{gg}\phi}\,\mathcal{B}(\phi \to \tau\tau)$ and $\sigma_{\mathrm{bb}\phi}\,\mathcal{B}(\phi \to \tau\tau)$. The 68 and 95 % CL contours are shown in dark and light purple, respectively. The best fit points are indicated by black crosses. The contours are shown for $m_\phi = 140,\ 160,\ 180,\ 200,\ 250,$ and 300 GeV from left to right and top to bottom. Lower row published in [30].

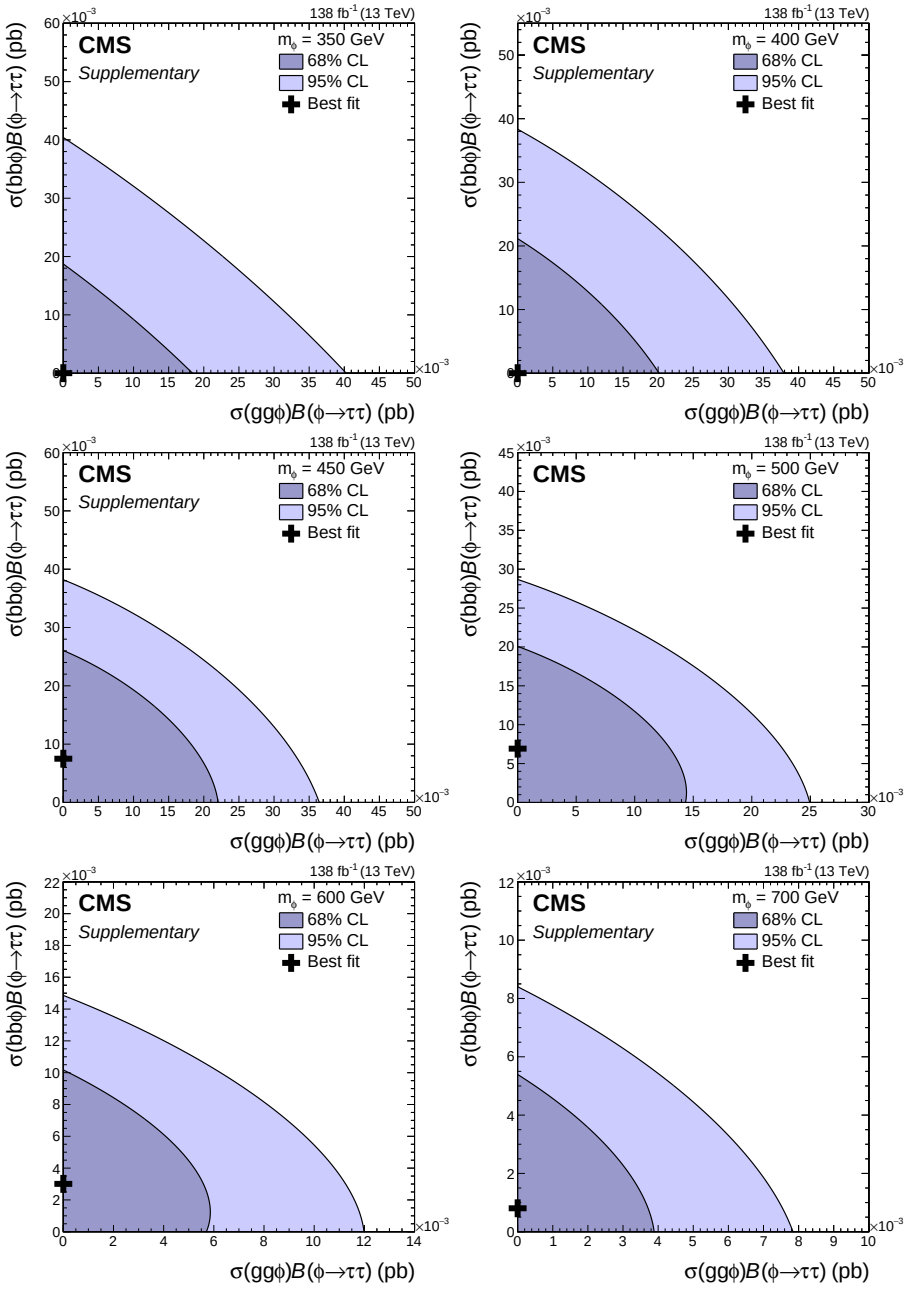

Figure B.8: Scans of the profile likelihood function in terms of $\sigma_{\mathrm{gg}\phi}\,\mathcal{B}(\phi \to \tau\tau)$ and $\sigma_{\mathrm{bb}\phi}\,\mathcal{B}(\phi \to \tau\tau)$. The 68 and 95 % CL contours are shown in dark and light purple, respectively. The best fit points are indicated by black crosses. The contours are shown for $m_\phi = 350,\ 400,\ 450,\ 500,\ 600,$ and 700 GeV from left to right and top to bottom. Figures published in [30].

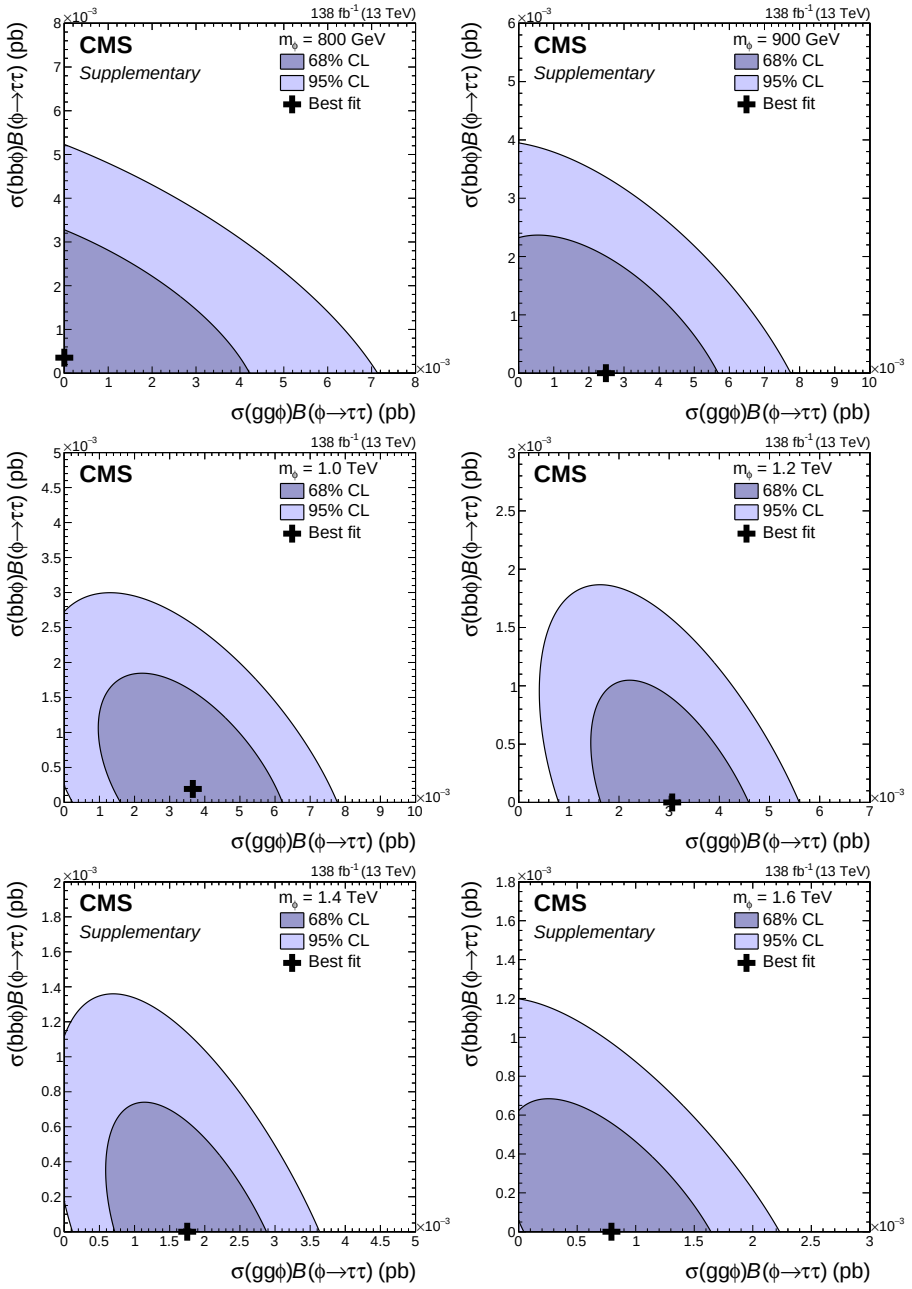

Figure B.9: Scans of the profile likelihood function in terms of $\sigma_{\mathrm{gg}\phi}\,\mathcal{B}(\phi \to \tau\tau)$ and $\sigma_{\mathrm{bb}\phi}\,\mathcal{B}(\phi \to \tau\tau)$. The 68 and 95 % CL contours are shown in dark and light purple, respectively. The best fit points are indicated by black crosses. The contours are shown for $m_\phi = 800,\ 900,\ 1000,\ 1200,\ 1400,$ and 1600 GeV from left to right and top to bottom. Figures published in [30].

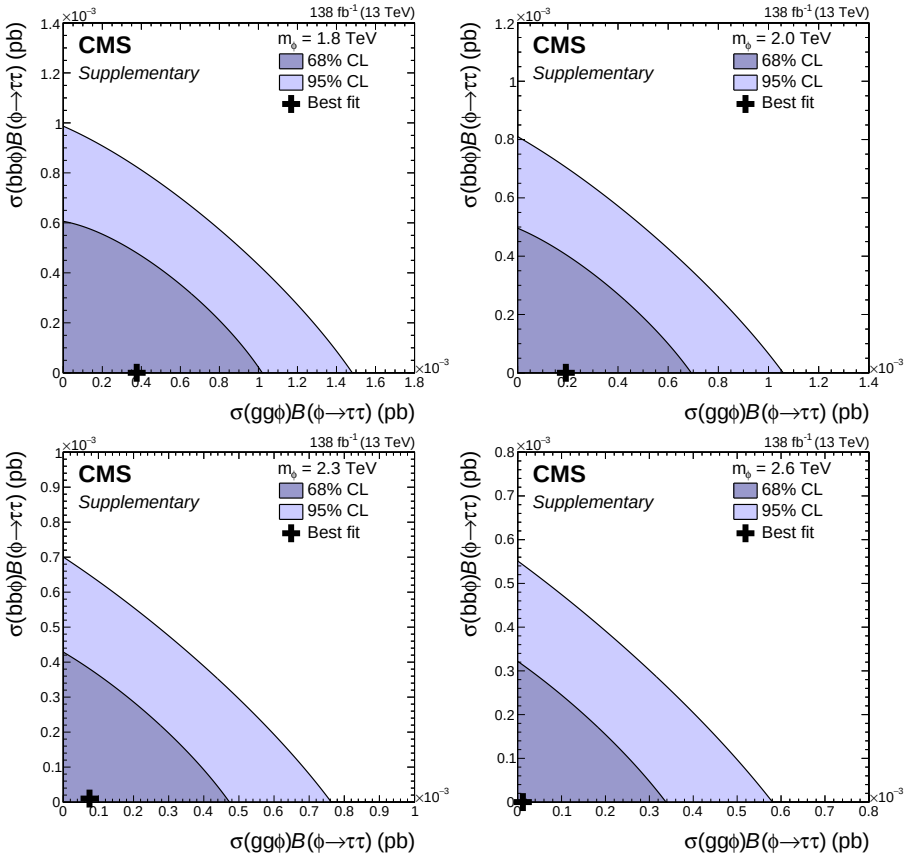

Figure B.10: Scans of the profile likelihood function in terms of $\sigma_{\mathrm{gg}\phi}\,\mathcal{B}(\phi \to \tau\tau)$ and $\sigma_{\mathrm{bb}\phi}\,\mathcal{B}(\phi \to \tau\tau)$. The 68 and 95 % CL contours are shown in dark and light purple, respectively. The best fit points are indicated by black crosses. The contours are shown for $m_\phi = 1.8,\ 2.0,\ 2.3,$ and 2.6 TeV from left to right and top to bottom. Figures published in [30].

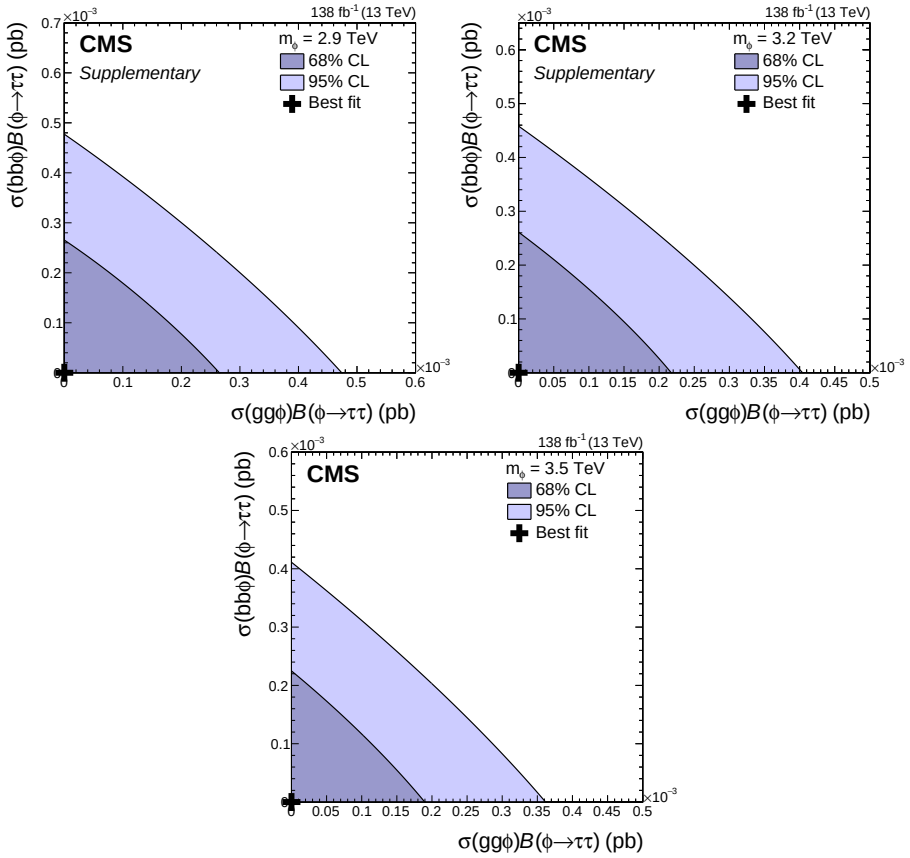

Figure B.11: Scans of the profile likelihood function in terms of $\sigma_{\mathrm{gg}\phi}\,\mathcal{B}(\phi \to \tau\tau)$ and $\sigma_{\mathrm{bb}\phi}\,\mathcal{B}(\phi \to \tau\tau)$. The 68 and 95 % CL contours are shown in dark and light purple, respectively. The best fit points are indicated by black crosses. The contours are shown for $m_\phi = 2.9,\ 3.2,$ and $3.5\,\mathrm{TeV}$ from left to right and top to bottom. Figures published in [30].

APPENDIX C

Exclusion Contours in Additional MSSM Benchmark Scenarios

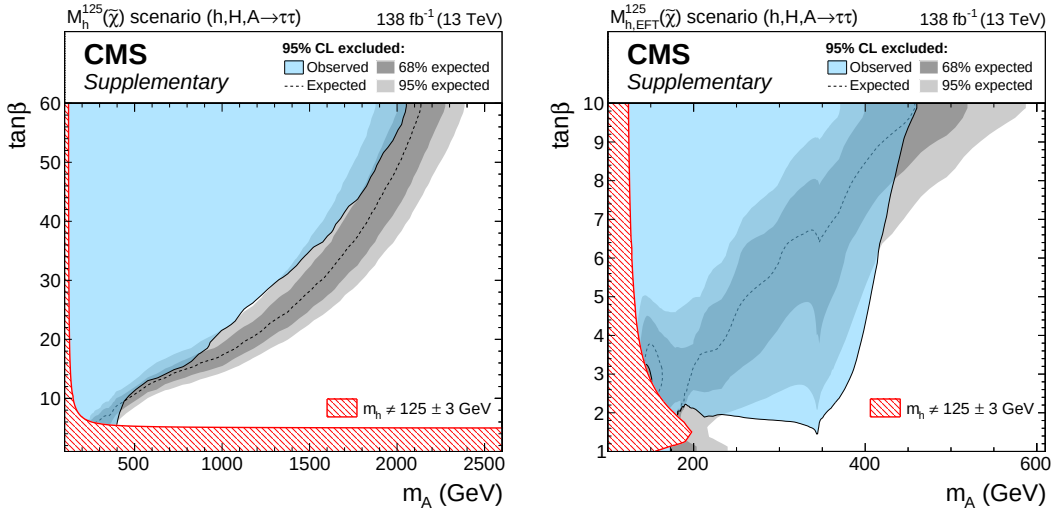

Figure C.1: Observed and expected 95 % CL exclusion contours in additional MSSM benchmark scenarios. The exclusion contours are shown for the $M_{\mathrm{h}}^{125}(\tilde{\chi})$ and $M_{\mathrm{h,EFT}}^{125}(\tilde{\chi})$ scenarios, as discussed in [63], on the left and right, respectively. The median expected contour and the central 68 and 95 % intervals in absence of signal are indicated by a black dashed line and by dark and light gray bands, respectively. The observed contour is shown as black line and the excluded parameter space indicated by a light blue filled area. Regions of the parameter space for which m_{h} does not coincide with $m_{\mathrm{H(125)}}$ within a theoretical uncertainty of ±3 GeV are indicated by red hatched areas. Figures published in [30].

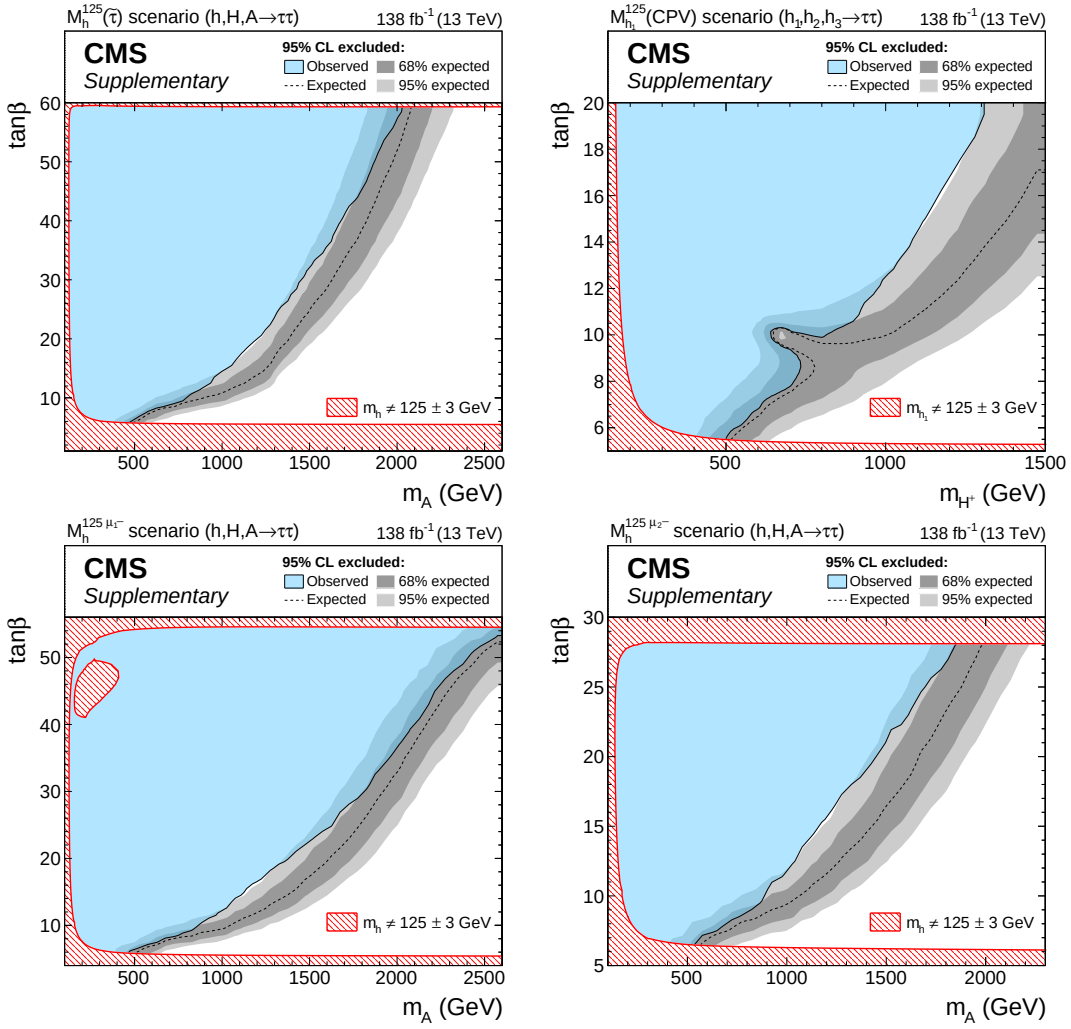

Figure C.2: Observed and expected 95 % CL exclusion contours in additional MSSM benchmark scenarios. The exclusion contours are shown for the $M_{\mathrm{h}}^{125}(\tilde{\tau})$, $M_{\mathrm{h_1}}^{125}$(CPV), $M_{\mathrm{h}}^{125\,\mu_1-}$, and $M_{\mathrm{h}}^{125\,\mu_2-}$ scenarios, as discussed in [63], from left to right and top to bottom. The median expected contour and the central 68 and 95 % intervals in absence of signal are indicated by a black dashed line and by dark and light gray bands, respectively. The observed contour is shown as black line and the excluded parameter space indicated by a light blue filled area. Regions of the parameter space for which m_{h} does not coincide with $m_{\mathrm{H}(125)}$ within a theoretical uncertainty of ±3 GeV are indicated by red hatched areas. Figures published in [30].

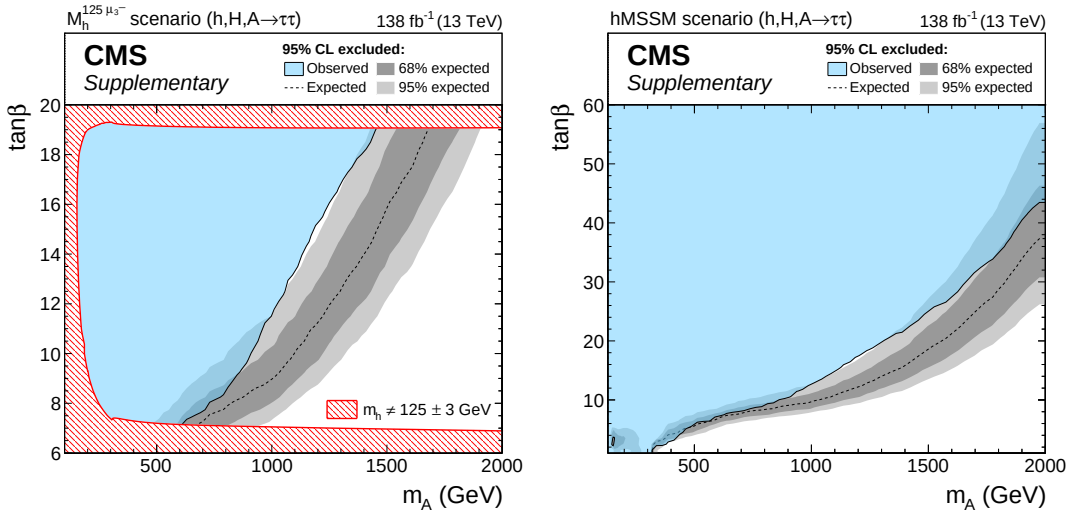

Figure C.3: Observed and expected 95 % CL exclusion contours in additional MSSM benchmark scenarios. The exclusion contours are shown for the $M_{\mathrm{h}}^{125\,\mu_3-}$ and hMSSM scenarios, as discussed in [63], on the left and right, respectively. The median expected contour and the central 68 and 95 % intervals in absence of signal are indicated by a black dashed line and by dark and light gray bands, respectively. The observed contour is shown as black line and the excluded parameter space indicated by a light blue filled area. Regions of the parameter space for which m_{h} does not coincide with $m_{\mathrm{H(125)}}$ within a theoretical uncertainty of ± 3 GeV are indicated by red hatched areas. Figures published in [30].

List of Figures

List of Tables

Bibliography

[1] The CMS Collaboration. „Observation of a New Boson at a Mass of 125 GeV with the CMS Experiment at the LHC“. *Phys. Lett. B* 716.1 (2012), pp. 30–61. DOI: 10.1016/j.physletb.2012.08.021.

[2] The 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* 716.1 (2012), pp. 1–29. DOI: 10.1016/j.physletb.2012.08.020.

[3] Glashow, S. L. „Partial-Symmetries of Weak Interactions“. *Nucl. Phys.* 22.4 (1961), pp. 579–588. DOI: 10.1016/0029-5582(61)90469-2.

[4] Weinberg, S. „A Model of Leptons“. *Phys. Rev. Lett.* 19.21 (1967), pp. 1264–1266. DOI: 10.1103/physrevlett.19.1264.

[5] Englert, F. and Brout, R. „Broken Symmetry and the Mass of Gauge Vector Mesons“. *Phys. Rev. Lett.* 13.9 (1964), pp. 321–323. DOI: 10.1103/physrevlett.13.321.

[6] Higgs, P. W. „Broken symmetries, massless particles and gauge fields“. *Phys. Lett.* 12.2 (1964), pp. 132–133. DOI: 10.1016/0031-9163(64)91136-9.

[7] Higgs, P. W. „Broken Symmetries and the Masses of Gauge Bosons“. *Phys. Rev. Lett.* 13 (1964), pp. 508–509. DOI: 10.1103/physrevlett.13.508.

[8] Higgs, P. W. „Spontaneous Symmetry Breakdown without Massless Bosons“. *Phys. Rev.* 145.4 (1966), pp. 1156–1163. DOI: 10.1103/PhysRev.145.1156.

[9] The CMS Collaboration. „Study of the Mass and Spin-Parity of the Higgs Boson Candidate via Its Decays to Z Boson Pairs“. *Phys. Rev. Lett.* 110.8, 081803 (2013). DOI: 10.1103/physrevlett.110.081803.

[10] The ATLAS 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“. *J. High Energy Phys.* 2016.08, 045 (2016).
DOI: 10.1007/JHEP08(2016)045.

[11] The CMS Collaboration. „Observation of $t\bar{t}H$ Production“. *Phys. Rev. Lett.* 120.23, 231801 (2018).
DOI: 10.1103/physrevlett.120.231801.

[12] The CMS Collaboration. „Measurements of properties of the Higgs boson decaying to a W boson pair in pp collisions at $\sqrt{s} = 13$ TeV“. *Phys. Lett. B* 791 (2019), pp. 96–129.
DOI: 10.1016/j.physletb.2018.12.073.

[13] The CMS Collaboration. „Combined measurements of Higgs boson couplings in proton-proton collisions at $\sqrt{s} = 13$ TeV“. *Eur. Phys. J. C* 79.5, 421 (2019).
DOI: 10.1140/epjc/s10052-019-6909-y.

[14] The CMS Collaboration. „A measurement of the Higgs boson mass in the diphoton decay channel“. *Phys. Lett. B* 805, 135425 (2020).
DOI: 10.1016/j.physletb.2020.135425.

[15] The CMS Collaboration. „Measurements of Higgs boson production in the decay channel with a pair of τ leptons in proton-proton collisions at $\sqrt{s} = 13$ TeV“ (2022). Accepted by: *Eur. Phys. J. C*.
DOI: 10.48550/ARXIV.2204.12957.

[16] The CMS Collaboration. „Analysis of the CP structure of the Yukawa coupling between the Higgs boson and τ leptons in proton-proton collisions at $\sqrt{s} = 13$ TeV“. *J. High Energy Phys.* 2022.06, 012 (2022).
DOI: 10.1007/jhep06(2022)012.

[17] The CMS Collaboration. „A portrait of the Higgs boson by the CMS experiment ten years after the discovery“. *Nature* 607.7917 (2022), pp. 60–68.
DOI: 10.1038/s41586-022-04892-x.

[18] The CMS Collaboration. „Observation of the diphoton decay of the Higgs boson and measurement of its properties“. *Eur. Phys. J. C* 74.10, 3076 (2014).
DOI: 10.1140/epjc/s10052-014-3076-z.

[19] The CMS Collaboration. „Measurement of the properties of a Higgs boson in the four-lepton final state“. *Phys. Rev. D* 89, 092007 (2014).
DOI: 10.1103/PhysRevD.89.092007.

[20] The CMS Collaboration. „Observation of the Higgs boson decay to a pair of τ leptons with the CMS detector“. *Phys. Lett. B* 779 (2018), pp. 283–316.
DOI: 10.1016/j.physletb.2018.02.004.

[21] The CMS Collaboration. „Observation of Higgs Boson Decay to Bottom Quarks“. *Phys. Rev. Lett.* 121.12, 121801 (2018).
DOI: 10.1103/physrevlett.121.121801.

[22] The CMS Collaboration. „Measurements of properties of the Higgs boson decaying to a W boson pair in pp collisions at $\sqrt{s} = 13\,\text{TeV}$“. *Phys. Lett. B* 791 (2019), pp. 96–129.
DOI: 10.1016/j.physletb.2018.12.073.

[23] Wess, J. and Zumino, B. „Supergauge transformations in four dimensions“. *Nucl. Phys. B* 70.1 (1974), pp. 39–50.
DOI: 10.1016/0550-3213(74)90355-1.

[24] Golfand, Y. A. and Likhtman, E. P. „Extension of the Algebra of Poincare Group Generators and Violation of p Invariance“. *JETP Lett.* 13 (1971). [*Pisma Zh. Eksp. Teor. Fiz.* 13.452 (1971)], pp. 323–326.

[25] Haber, H. and Kane, G. „The search for supersymmetry: Probing physics beyond the standard model“. *Phys. Rep.* 117.2 (1985), pp. 75–263.
DOI: 10.1016/0370-1573(85)90051-1.

[26] Nilles, H. „Supersymmetry, supergravity and particle physics“. *Phys. Rep.* 110.1 (1984), pp. 1–162.
DOI: 10.1016/0370-1573(84)90008-5.

[27] Djouadi, A. „The Anatomy of electro-weak symmetry breaking. II. The Higgs bosons in the minimal supersymmetric model“. *Phys. Rep.* 459 (2008), pp. 1–241.
DOI: 10.1016/j.physrep.2007.10.005.

[28] Fayet, P. „Supergauge Invariant Extension of the Higgs Mechanism and a Model for the electron and Its Neutrino“. *Nucl. Phys.* B90 (1975), pp. 104–124.
DOI: 10.1016/0550-3213(75)90636-7.

[29] Fayet, P. „Spontaneously broken supersymmetric theories of weak, electromagnetic and strong interactions“. *Phys. Lett. B* 69.4 (1977), pp. 489–494.
DOI: 10.1016/0370-2693(77)90852-8.

[30] The CMS Collaboration. „Searches for additional Higgs bosons and for vector leptoquarks in $\tau\tau$ final states in proton-proton collisions at $\sqrt{s} = 13$ TeV“ (2022). Submitted to: *J. High Energy Phys.*
DOI: 10.48550/arXiv.2208.02717.

[31] The CMS Collaboration. „Search for additional neutral MSSM Higgs bosons in the $\tau\tau$ final state in proton-proton collisions at $\sqrt{s} = 13\,\text{TeV}$“. *J. High Energy Phys.* 2018.09, 007 (2018).
DOI: 10.1007/JHEP09(2018)007.

[32] Georgi, H. „A Model of Soft CP Violation“. *Hadronic J.* 1.155 (1978).

[33] Bargmann, V. and Wigner, E. P. „Group Theoretical Discussion of Relativistic Wave Equations“. *Proc. Natl. Acad. Sci. U.S.A.* 34.5 (1948), pp. 211–223.
DOI: 10.1073/pnas.34.5.211.

[34] Pauli, W. „The Connection Between Spin and Statistics“. *Phys. Rev.* 58.8 (1940), pp. 716–722.
DOI: 10.1103/PhysRev.58.716.

[35] The Particle Data Group Collaboration. „Review of Particle Physics“. *Prog. Theor. Exp. Phys.* 2022.8, 083C01 (2022).
DOI: 10.1093/ptep/ptac097.

[36] Peskin, M. E. and Schroeder, D. V. „An introduction to quantum field theory“. 4th edition. Reading, Massachusetts: Addison-Wesley, 1997.

[37] Mandl, F. and Shaw, G. „Quantum field theory“. 2nd. Chichester, West Sussex: Wiley, 2010.

[38] Georgi, H. M. et al. „Higgs Bosons from Two-Gluon Annihilation in Proton-Proton Collisions“. *Phys. Rev. Lett.* 40.11 (1978), pp. 692–694.
DOI: 10.1103/physrevlett.40.692.

[39] LHC Higgs Cross Section Working Group. „Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector“ (2016).
DOI: 10.23731/CYRM-2017-002.

[40] Anastasiou, C. et al. „High precision determination of the gluon fusion Higgs boson cross-section at the LHC“. *J. High Energy Phys.* 2016.05 (2016).
DOI: 10.1007/jhep05(2016)058.

[41] LHC Higgs Cross Section Working Group. „Handbook of LHC Higgs Cross Sections: 1. Inclusive Observables“ (2011).
DOI: 10.5170/CERN-2011-002.

[42] The ATLAS Collaboration. „Observation of $H \to b\bar{b}$ decays and VH production with the ATLAS detector“. *Phys. Lett. B* 786 (2018), pp. 59–86.
DOI: 10.1016/j.physletb.2018.09.013.

[43] The Super-Kamiokande Collaboration. „Evidence for Oscillation of Atmospheric Neutrinos“. *Phys. Rev. Lett.* 81.8 (1998), pp. 1562–1567.
DOI: 10.1103/physrevlett.81.1562.

[44] Djouadi, A. „The Anatomy of electro-weak symmetry breaking. I: The Higgs boson in the standard model“. *Phys. Rep.* 457 (2008), pp. 1–216.
DOI: 10.1016/j.physrep.2007.10.004.

[45] Bertone, G. et al. „Particle dark matter: evidence, candidates and constraints“. *Phys. Rep.* 405.5-6 (2005), pp. 279–390.
DOI: 10.1016/j.physrep.2004.08.031.

[46] Haber, H. E. and Stål, O. „New LHC benchmarks for the $\mathcal{CP}$-conserving two-Higgs-doublet model“. *Eur. Phys. J. C* 75.10, 491 (2015).
DOI: 10.1140/epjc/s10052-015-3697-x.

[47] Glashow, S. L. and Weinberg, S. „Natural conservation laws for neutral currents“. *Phys. Rev. D* 15.7 (1977), pp. 1958–1965.
DOI: 10.1103/physrevd.15.1958.

[48] Eriksson, D. et al. „2HDMC – two-Higgs-doublet model calculator“. *Comput. Phys. Commun.* 181.1 (2010), pp. 189–205.
DOI: 10.1016/j.cpc.2009.09.011.

[49] Harlander, R. V. et al. „SusHi: A program for the calculation of Higgs production in gluon fusion and bottom-quark annihilation in the Standard Model and the MSSM“. *Comput. Phys. Commun.* 184.6 (2013), pp. 1605–1617.
DOI: 10.1016/j.cpc.2013.02.006.

[50] Harlander, R. V. et al. „SusHi Bento: Beyond NNLO and the heavy- top limit“. *Comput. Phys. Commun.* 212 (2017), pp. 239–257.
DOI: 10.1016/j.cpc.2016.10.015.

[51] Harlander, R. V. and Kant, P. „Higgs production and decay: analytic results at next-to-leading order QCD“. *J. High Energy Phys.* 2005.12, 015 (2005).
DOI: 10.1088/1126-6708/2005/12/015.

[52] Harlander, R. V. and Kilgore, W. B. „Next-to-Next-to-Leading Order Higgs Production at Hadron Colliders“. *Phys. Rev. Lett.* 88.20, 201801 (2002).
DOI: 10.1103/physrevlett.88.201801.

[53] Aglietti, U. et al. „Two-loop light fermion contribution to Higgs production and decays“. *Phys. Lett. B* 595.1-4 (2004), pp. 432–441.
DOI: 10.1016/j.physletb.2004.06.063.

[54] Bonciani, R. et al. „On the generalized harmonic polylogarithms of one complex variable“. *Comput. Phys. Commun.* 182.6 (2011), pp. 1253–1264.
DOI: 10.1016/j.cpc.2011.02.011.

[55] Anastasiou, C. et al. „Higgs boson gluon-fusion production beyond threshold in N3LO QCD“. *J. High Energy Phys.* 2015.03 (2015).
DOI: 10.1007/jhep03(2015)091.

[56] Anastasiou, C. et al. „Soft expansion of double-real-virtual corrections to Higgs production at N3LO“. *J. High Energy Phys.* 2015.8 (2015).
DOI: 10.1007/jhep08(2015)051.

[57] Harlander, R. V. and Kilgore, W. B. „Higgs boson production in bottom quark fusion at next-to-next-to-leading order“. *Phys. Rev. D* 68.1, 013001 (2003).
DOI: 10.1103/PhysRevD.68.013001.

[58] Chetyrkin, K. et al. „RunDec: a Mathematica package for running and decoupling of the strong coupling and quark masses“. *Comput. Phys. Commun.* 133.1 (2000), pp. 43–65.
DOI: 10.1016/S0010-4655(00)00155-7.

[59] Butterworth, J. et al. „PDF4LHC recommendations for LHC Run II“. *J. Phys. G* 43.2, 023001 (2016).
DOI: 10.1088/0954-3899/43/2/023001.

[60] The CMS Collaboration. „Summary results of high mass BSM Higgs searches using CMS Run-I data“. Tech. rep. Geneva: CERN, 2016.
CDS: 2142432.

[61] The CMS Collaboration. „Search for high mass resonances decaying into W^+W^- in the dileptonic final state with 138 fb^{-1} of proton-proton collisions at $\sqrt{s} = 13$ TeV“. Tech. rep. Geneva: CERN, 2022.
CDS: 2803723.

[62] Gunion, J. F. et al. „The Higgs Hunter's Guide“. Frontiers in physics, 80. Boulder, Colorado: Westview Pr., 2003.

[63] Bagnaschi, E. A. et al. „Benchmark Scenarios for MSSM Higgs Boson Searches at the LHC“. Tech. rep. Geneva: CERN, 2021.
CDS: 2791954.

[64] Bagnaschi, E. et al. „MSSM Higgs boson searches at the LHC: benchmark scenarios for Run 2 and beyond“. *Eur. Phys. J. C* 79.7, 617 (2019).
DOI: 10.1140/epjc/s10052-019-7114-8.

[65] Bahl, H. et al. „MSSM Higgs benchmark scenarios for Run 2 and beyond: the low $\tan\beta$ region“. *Eur. Phys. J. C* 79.3, 279 (2019).
DOI: 10.1140/epjc/s10052-019-6770-z.

[66] Bahl, H. et al. „HL-LHC and ILC sensitivities in the hunt for heavy Higgs bosons“. *Eur. Phys. J. C* 80.10, 916 (2020).
DOI: 10.1140/epjc/s10052-020-08472-z.

[67] LHC Higgs Working Group – MSSM Subgroup. *LHCHWG MSSM ROOT files*. 2022.
DOI: 10.5281/ZENODO.6334713.

[68] Bagnaschi, E. et al. „Benchmark scenarios for low $\tan\beta$ in the MSSM“. Tech. rep. Geneva: CERN, 2015.
CDS: 2039911.

[69] Heinemeyer, S. et al. „FeynHiggs: A Program for the calculation of the masses of the neutral CP even Higgs bosons in the MSSM“. *Comput. Phys. Commun.* 124 (2000), pp. 76–89.
DOI: 10.1016/S0010-4655(99)00364-1.

[70] Slavich, P. et al. „Higgs-mass predictions in the MSSM and beyond“. *Eur. Phys. J. C* 81.5, 450 (2021).
DOI: 10.1140/epjc/s10052-021-09198-2.

[71] Djouadi, A. et al. „HDECAY: A Program for Higgs boson decays in the standard model and its supersymmetric extension“. *Comput. Phys. Commun.* 108 (1998), pp. 56–74.
DOI: 10.1016/S0010-4655(97)00123-9.

[72] The CMS Collaboration. „The CMS experiment at the CERN LHC. The Compact Muon Solenoid experiment“. *J. Instrum.* 3.08, S08004 (2008).
DOI: 10.1088/1748-0221/3/08/s08004.

[73] Evans, L. and Bryant, P. „LHC Machine“. *J. Instrum.* 3, S08001 (2008).
DOI: 10.1088/1748-0221/3/08/S08001.

[74] The ATLAS Collaboration. „The ATLAS Experiment at the CERN Large Hadron Collider“. *J. Instrum.* 3, S08003 (2008).
DOI: 10.1088/1748-0221/3/08/S08003.

[75] The LHCb Collaboration. „The LHCb Detector at the LHC“. *J. Instrum.* 3, S08005 (2008).
DOI: 10.1088/1748-0221/3/08/S08005.

[76] The ALICE Collaboration. „The ALICE Experiment at the CERN LHC“. *J. Instrum.* 3, S08002 (2008).
DOI: 10.1088/1748-0221/3/08/S08002.

[77] The CMS Collaboration. „Interactive Slice of the CMS detector“ (2016).
CDS: 2205172.

[78] Dominguez, A. et al. „CMS Technical Design Report for the Pixel Detector Upgrade“. Tech. rep. CERN-LHCC-2012-016. CMS-TDR-11. Geneva: CERN, 2012.
CDS: 1481838.

[79] Adzic, P. et al. „Energy resolution of the barrel of the CMS Electromagnetic Calorimeter“. *J. Instrum.* 2.04, P04004 (2007).
DOI: 10.1088/1748-0221/2/04/p04004.

[80] Abdullin, S. et al. „The CMS barrel calorimeter response to particle beams from 2 to 350 GeV/c“. *Eur. Phys. J. C* 60.3 (2009), pp. 359–373.
DOI: 10.1140/epjc/s10052-009-0959-5.

[81] „The CMS trigger system“. *J. Instrum.* 12.01, P01020 (2017).
DOI: 10.1088/1748-0221/12/01/p01020.

[82] The CMS Collaboration. „Performance of the CMS Level-1 trigger in proton-proton collisions at $\sqrt{s} = 13$ TeV“. *J. Instrum.* 15.10, P10017 (2020).
DOI: 10.1088/1748-0221/15/10/p10017.

[83] The CMS Collaboration. „Particle-flow reconstruction and global event description with the CMS detector“. *J. Instrum.* 12.10, P10003 (2017).
DOI: 10.1088/1748-0221/12/10/p10003.

[84] CMS Collaboration. „Description and performance of track and primary-vertex reconstruction with the CMS tracker“. *J. Instrum.* 9.10, P10009 (2014).
DOI: 10.1088/1748-0221/9/10/p10009.

[85] Adam, W. et al. „Track Reconstruction in the CMS tracker“. Tech. rep. CMS-NOTE-2006-041. Geneva: CERN, 2006.
CDS: 934067.

[86] Frühwirth, R. et al. „Adaptive Vertex Fitting“. Tech. rep. CMS-NOTE-2007-008. Geneva: CERN, 2007.
CDS: 1027031.

[87] Cacciari, M. et al. „The anti-kt jet clustering algorithm“. *J. High Energy Phys.* 2008.04, 063 (2008).
DOI: 10.1088/1126-6708/2008/04/063.

[88] Cacciari, M. et al. „FastJet user manual“. *Eur. Phys. J. C* 72.3, 1896 (2012). DOI: 10.1140/epjc/s10052-012-1896-2.

[89] Contardo, D. et al. „Technical Proposal for the Phase-II Upgrade of the CMS Detector“. Tech. rep. Geneva: CERN, 2015. DOI: 10.17181/CERN.VU8I.D59J.

[90] The CMS Collaboration. „Performance of the CMS muon detector and muon reconstruction with proton-proton collisions at $\sqrt{s} = 13$ TeV“. *J. Instrum.* 13.06, P06015 (2018). DOI: 10.1088/1748-0221/13/06/p06015.

[91] The CMS Collaboration. „Electron and photon reconstruction and identification with the CMS experiment at the CERN LHC“. *J. Instrum.* 16.05, P05014 (2021). DOI: 10.1088/1748-0221/16/05/p05014.

[92] Adam, W. et al. „Reconstruction of electrons with the Gaussian-sum filter in the CMS tracker at the LHC“. *J. Phys. G* 31.9 (2005). DOI: 10.1088/0954-3899/31/9/n01.

[93] Khachatryan, V. et al. „Jet energy scale and resolution in the CMS experiment in pp collisions at 8 TeV“. *J. Instrum.* 12.02, P02014 (2017). DOI: 10.1088/1748-0221/12/02/p02014.

[94] The CMS Collaboration. „Performance of tau-lepton reconstruction and identification in CMS“. *J. Instrum.* 7.01, P01001 (2012). DOI: 10.1088/1748-0221/7/01/P01001.

[95] The CMS Collaboration. „Reconstruction and identification of τ lepton decays to hadrons and ν_τ at CMS“. *J. Instrum.* 11.01, P01019 (2016). DOI: 10.1088/1748-0221/11/01/p01019.

[96] The CMS Collaboration. „Performance of reconstruction and identification of τ leptons decaying to hadrons and ν_τ in pp collisions at $\sqrt{s} = 13$ TeV“. *J. Instrum.* 13.10, P10005 (2018). DOI: 10.1088/1748-0221/13/10/p10005.

[97] The CMS Collaboration. „Identification of hadronic tau lepton decays using a deep neural network“. *J. Instrum.* 17.07, P07023 (2022). DOI: 10.1088/1748-0221/17/07/p07023.

[98] The CMS Collaboration. „Performance of missing transverse momentum reconstruction in proton-proton collisions at $\sqrt{s} = 13$ TeV using the CMS detector“. *J. Instrum.* 14.07, P07004 (2019). DOI: 10.1088/1748-0221/14/07/p07004.

[99] Bertolini, D. et al. „Pileup per particle identification“. *J. High Energy Phys.* 2014.10 (2014). DOI: 10.1007/jhep10(2014)059.

[100] The CMS Collaboration. „Pileup mitigation at CMS in 13 TeV data“. *J. Instrum.* 15.09, P09018 (2020).
DOI: 10.1088/1748-0221/15/09/p09018.

[101] The CMS Collaboration. „Identification of heavy-flavour jets with the CMS detector in pp collisions at 13 TeV“. *J. Instrum.* 13.05, P05011 (2018).
DOI: 10.1088/1748-0221/13/05/p05011.

[102] Bols, E. et al. „Jet flavour classification using DeepJet“. *J. Instrum.* 15.12, P12012 (2020).
DOI: 10.1088/1748-0221/15/12/p12012.

[103] The CMS Collaboration. „Precision luminosity measurement in proton–proton collisions at $\sqrt{s} = 13\,\text{TeV}$ in 2015 and 2016 at CMS“. *Eur. Phys. J. C* 81.9, 450 (2021).
DOI: 10.1140/epjc/s10052-021-09538-2.

[104] The CMS Collaboration. „CMS luminosity measurement for the 2017 data-taking period at $\sqrt{s} = 13$ TeV“. Tech. rep. Geneva: CERN, 2018.
CDS: 2621960.

[105] The CMS Collaboration. „CMS luminosity measurement for the 2018 data-taking period at $\sqrt{s} = 13$ TeV“. Tech. rep. Geneva: CERN, 2019.
CDS: 2676164.

[106] Gottmann, A. „Global Interpretation of $\tau\tau$ Events in the Context of the Standard Model and Beyond“. PhD thesis. Karlsruhe Institute of Technology (KIT), 2020.
DOI: 10.5445/IR/1000124886.

[107] Drell, S. D. and Yan, T.-M. „Massive Lepton-Pair Production in Hadron-Hadron Collisions at High Energies“. *Phys. Rev. Lett.* 25.5 (1970), pp. 316–320.
DOI: 10.1103/physrevlett.25.316.

[108] Drell, S. D. and Yan, T.-M. „Massive Lepton-Pair Production in Hadron-Hadron Collisions at High Energies“. *Phys. Rev. Lett.* 25.13, 902 (1970).
DOI: 10.1103/physrevlett.25.902.2.

[109] The CMS Collaboration. „An embedding technique to determine $\tau\tau$ backgrounds in proton-proton collision data“. 14.06, P06032 (2019).
DOI: 10.1088/1748-0221/14/06/p06032.

[110] The CMS Collaboration. „Measurement of the $\text{Z}/\gamma^* \to \tau\tau$ cross section in pp collisions at $\sqrt{s} = 13\,\text{TeV}$ and validation of τ lepton analysis techniques“. *Eur. Phys. J. C* 78.9, 708 (2018).
DOI: 10.1140/epjc/s10052-018-6146-9.

[111] Abulencia, A. et al. „Search for Neutral Higgs Bosons of the Minimal Supersymmetric Standard Model Decaying to τ Pairs in $p\bar{p}$ Collisions at $\sqrt{s} = 1.96\,\text{TeV}$“. *Phys. Rev. Lett.* 96.1, 011802 (2006).
DOI: 10.1103/physrevlett.96.011802.

[112] The ATLAS Collaboration. „Search for neutral Higgs bosons of the minimal supersymmetric standard model in pp collisions at $\sqrt{s} = 8\,\text{TeV}$ with the ATLAS detector“. *J. High Energy Phys.* 2014.11, 056 (2014).
DOI: 10.1007/JHEP11(2014)056.

[113] Bianchini, L. et al. „Reconstruction of the Higgs mass in $H \to \tau\tau$ Events by Dynamical Likelihood techniques“. *J. Phys. Conf. Ser.* 513.2, 022035 (2014).
DOI: 10.1088/1742-6596/513/2/022035.

[114] Alwall, J. et al. „The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations“. *J. High Energy Phys.* 2014.07, 079 (2014).
DOI: 10.1007/JHEP07(2014)079.

[115] Nason, P. „A New Method for Combining NLO QCD with Shower Monte Carlo Algorithms“. *J. High Energy Phys.* 2004.11, 040 (2004).
DOI: 10.1088/1126-6708/2004/11/040.

[116] Frixione, S. et al. „Matching NLO QCD computations with Parton Shower simulations: the POWHEG method“. *J. High Energy Phys.* 11.11, 070 (2007).
DOI: 10.1088/1126-6708/2007/11/070.

[117] Alioli, S. et al. „A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX“. *J. High Energy Phys.* 2010.06, 043 (2010).
DOI: 10.1007/JHEP06(2010)043.

[118] Sjöstrand, T. et al. „An introduction to PYTHIA 8.2“. *Comput. Phys. Commun.* 191 (2015), pp. 159–177.
DOI: 10.1016/j.cpc.2015.01.024.

[119] Alwall, J. et al. „Comparative study of various algorithms for the merging of parton showers and matrix elements in hadronic collisions“. *Eur. Phys. J. C* 53.3 (2008), pp. 473–500.
DOI: 10.1140/epjc/s10052-007-0490-5.

[120] Frederix, R. and Frixione, S. „Merging meets matching in MC@NLO“. *J. High Energy Phys.* 2012.12, 061 (2012).
DOI: 10.1007/JHEP12(2012)061.

[121] The NNPDF Collaboration. „Parton distributions for the LHC run II“. *J. High Energy Phys.* 2015.04 (2015).
DOI: 10.1007/jhep04(2015)040.

[122] The NNPDF Collaboration. „Parton distributions from high-precision collider data“. *Eur. Phys. J. C* 77.10, 663 (2017).
DOI: 10.1140/epjc/s10052-017-5199-5.

[123] The CMS Collaboration. „Event generator tunes obtained from underlying event and multiparton scattering measurements“. *Eur. Phys. J. C* 76, 155 (2016).
DOI: 10.1140/epjc/s10052-016-3988-x.

[124] The CMS Collaboration. „Extraction and validation of a new set of CMS PYTHIA8 tunes from underlying-event measurements“. *Eur. Phys. J. C* 80.1, 4 (2020).
DOI: 10.1140/epjc/s10052-019-7499-4.

[125] Agostinelli, S. et al. „Geant4 - a simulation toolkit“. *Nucl. Instrum. Methods Phys. Res. A* 506.3 (2003), pp. 250–303.
DOI: 10.1016/S0168-9002(03)01368-8.

[126] Frixione, S. et al. „A positive-weight next-to-leading-order Monte Carlo for heavy flavour hadroproduction“. *J. High Energy Phys.* 2007.09, 126 (2007).
DOI: 10.1088/1126-6708/2007/09/126.

[127] Alioli, S. et al. „NLO single-top production matched with shower in POWHEG:s-andt-channel contributions“. *J. High Energy Phys.* 2009.09, 111 (2009).
DOI: 10.1088/1126-6708/2009/09/111.

[128] Alioli, S. et al. „Erratum: NLO single-top production matched with shower in POWHEG: s- and t-channel contributions“. *J. High Energy Phys.* 2010.02, 011 (2010).
DOI: 10.1007/JHEP02(2010)011.

[129] Bagnaschi, E. et al. „Higgs production via gluon fusion in the POWHEG approach in the SM and in the MSSM“. *J. High Energy Phys.* 2012.02, 088 (2012).
DOI: 10.1007/JHEP02(2012)088.

[130] Nason, P. and Oleari, C. „NLO Higgs boson production via vector-boson fusion matched with shower in POWHEG“. *J. High Energy Phys.* 2010.02, 037 (2010).
DOI: 10.1007/JHEP02(2010)037.

[131] Luisoni, G. et al. „$HW^{\pm}/HZ + 0$ and 1 jet at NLO with the POWHEG BOX interfaced to GoSam and their merging within MiNLO“. *J. High Energy Phys.* 2013.10, 083 (2013).
DOI: 10.1007/JHEP10(2013)083.

[132] Jäger, B. et al. „Higgs boson production in association with b jets in the POWHEG BOX“. *Phys. Rev. D* 93.1, 014030 (2016).
DOI: 10.1103/physrevd.93.014030.

[133] Zimmer, C. „Monte Carlo simulation in the context of the 2HDM $H \to \tau\tau$ analysis at CMS“. MA thesis. Karlsruhe Institute of Technology (KIT), 2020.
ETP-KA: 22020.

[134] Bechtel, J. „A novel search for di-Higgs events in the $\tau^-\tau^+ + b\bar{b}$ final state in pp collisions at 13 TeV at the LHC“. PhD thesis. Karlsruhe Institute of Technology (KIT), 2021.
DOI: 10.5445/IR/1000130103.

[135] Andrejkovic, J. et al. „BSM $H \to \tau\tau$ analysis on full Run 2 CMS data at $\sqrt{s} = 13\,\text{TeV}$“. 2020.
AN: CMSAN-2020/218.

[136] The CMS Collaboration. „Measurements of inclusive W and Z cross sections in pp collisions at $\sqrt{s} = 7\,\text{TeV}$“. *J. High Energy Phys.* 2011.01, 80 (2011).
DOI: 10.1007/JHEP01(2011)080.

[137] The CMS Collaboration. „Search for nonresonant Higgs boson pair production in final state with two bottom quarks and two tau leptons in proton-proton collisions at $\sqrt{s} = 13\,\text{TeV}$“ (2022). Submitted to: *Phys. Lett. B.* arXiv: 2206.09401 [hep-ex].

[138] Cardini, A. et al. „Study of the misidentification of jets, electrons and muons as hadronically decaying tau leptons with the DeepTau ID for the full Run II data“. 2020.
AN: CMSAN-2020/115.

[139] The CMS Collaboration. „Measurement of the differential cross section for top quark pair production in pp collisions at $\sqrt{s} = 8\,\text{TeV}$“. *Eur. Phys. J. C* 75.11, 542 (2015).
DOI: 10.1140/epjc/s10052-015-3709-x.

[140] The CMS Collaboration. „Measurement of the Higgs boson production rate in association with top quarks in final states with electrons, muons, and hadronically decaying tau leptons at $\sqrt{s} = 13\,\text{TeV}$“. *Eur. Phys. J. C* 81.378 (2021).
DOI: 10.1140/epjc/s10052-021-09014-x.

[141] Cowan, G. et al. „Asymptotic formulae for likelihood-based tests of new physics“. *Eur. Phys. J. C* 71, 1554 (2011).
DOI: 10.1140/epjc/s10052-011-1554-0.

[142] The ATLAS Collaboration. „Procedure for the LHC Higgs boson search combination in Summer 2011“. Tech. rep. Geneva: CERN, 2011.
CDS: 1379837.

[143] Conway, J. S. „Incorporating Nuisance Parameters in Likelihoods for Multisource Spectra“. *Proceedings of the PHYSTAT 2011 Workshop on Statistical Issues Related to Discovery Claims in Search Experiments and Unfolding.* 2011, pp. 115–120.
DOI: 10.5170/CERN-2011-006.115.

[144] Barlow, R. and Beeston, C. „Fitting using finite Monte Carlo samples“. *Comput. Phys. Commun.* 77.2 (1993), pp. 219–228.
DOI: 10.1016/0010-4655(93)90005-w.

[145] LHC Higgs Cross Section Working Group. „Handbook of LHC Higgs Cross Sections: 2. Differential Distributions“ (2012).
DOI: 10.5170/CERN-2012-002.

[146] Bagnaschi, E. et al. „Resummation ambiguities in the Higgs transverse-momentum spectrum in the Standard Model and beyond“. *J. High Energy Phys.* 2016.01 (2016).
DOI: 10.1007/jhep01(2016)090.

[147] Harlander, R. V. et al. „Transverse momentum resummation for Higgs production via gluon fusion in the MSSM“. *J. High Energy Phys.* 2014.11 (2014). DOI: 10.1007/jhep11(2014)116.

[148] Bagnaschi, E. and Vicini, A. „The Higgs transverse momentum distribution in gluon fusion as a multiscale problem“. *J. High Energy Phys.* 2016.01 (2016). DOI: 10.1007/jhep01(2016)056.

[149] Read, A. L. „Linear interpolation of histograms“. *Nucl. Instrum. Methods Phys. Res. A* 425.1 (1999), pp. 357–360. DOI: 10.1016/S0168-9002(98)01347-3.

[150] Cowan, G. et al. „Erratum to: Asymptotic formulae for likelihood-based tests of new physics“. *Eur. Phys. J. C* 73.7, 2501 (2013). DOI: 10.1140/epjc/s10052-013-2501-z.

[151] Read, A. L. „Presentation of Search Results: the CLs Technique“. *J. Phys. G: Nucl. Part. Phys.* 28.10, 2693 (2002). DOI: 10.1088/0954-3899/28/10/313.

[152] Junk, T. „Confidence level computation for combining searches with small statistics“. *Nucl Instrum Meth A* 434.2-3 (1999), pp. 435–443. DOI: 10.1016/S0168-9002(99)00498-2.

[153] Wilks, S. S. „The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses“. *Ann. Math. Stat.* 9.1 (1938), pp. 60–62. DOI: 10.1214/aoms/1177732360.

[154] Cousins, R. D. „Generalization of Chisquare Goodness-of-Fit Test for Binned Data Using Saturated Models, with Application to Histograms“. 2013. URL: https://www.physics.ucla.edu/~cousins/stats/cousins_saturated.pdf (visited on 04/12/2022).

[155] The ATLAS Collaboration. „Search for Heavy Higgs Bosons Decaying into Two Tau Leptons with the ATLAS Detector Using pp Collisions at $\sqrt{s} = 13\,\text{TeV}$“. *Phys. Rev. Lett.* 125.5, 051801 (2020). DOI: 10.1103/PhysRevLett.125.051801.

[156] The CMS Collaboration. „Precise determination of the mass of the Higgs boson and tests of compatibility of its couplings with the standard model predictions using proton collisions at 7 and 8 TeV“. *Eur. Phys. J. C* 75.5, 212 (2015). DOI: 10.1140/epjc/s10052-015-3351-7.

[157] Carena, M. et al. „Effective Lagrangian for the interaction in the MSSM and charged Higgs phenomenology“. *Nucl. Phys. B* 577.1-2 (2000), pp. 88–120. DOI: 10.1016/s0550-3213(00)00146-2.

[158] Guasch, J. et al. „MSSM Higgs boson decays to bottom quark pairs reexamined“. *Phys. Rev. D* 68.11, 115001 (2003). DOI: 10.1103/physrevd.68.115001.

Danksagung

Ich möchte Priv.-Doz. Dr. Roger Wolf für die Möglichkeit zur Anfertigung dieser Dissertation und die Übernahme der Rolle als Doktorvater, sowie für viele wertvolle Diskussionen und die hervorragende Betreuung danken. Auch Prof. Dr. Günter Quast möchte ich für die Möglichkeit zum Verfassen dieser Dissertation am ETP und die Übernahme des Korreferats danken. Beiden möchte ich für die Begleitung meiner akademischen Laufbahn von der Bachelorarbeit im Fall von Prof. Quast über die Masterarbeit bis hin zur Dissertation von Herzen danken.

Weiterhin möchte ich der CMS Kollaboration für den reibungslosen Betrieb des CMS Experiments während der Datennahme in den Jahren von 2016 bis 2018 und die Bereitstellung der Daten in ausgezeichneter Qualität danken. Im Speziellen zu nennen sind hier auch Dr. Artur Gottmann, Sebastian Brommer, Dr. Daniel Winterbottom, George Uttley und Dr. Aliaksei Raspiareza, mit denen ich an der Publikation der Analyse arbeiten durfte.

Ein Dankeschön gilt auch dem ETP als Ganzem und den verantwortlichen Personen für die zuverlässige Bereitstellung der IT-Infrastruktur, ohne die meine Dissertation nicht in diesem Rahmen hätte angefertigt werden können. Ein weiterer Dank gebührt den Sekretärinnen des ETP, Frau Bräunling und Frau Chen, die bei organisatorischen Problemen stets hilfsbereit waren und dafür gesorgt haben, dass alle Erneuerungen meines Arbeitsvertrages während der dreieinhalb Jahre reibungslos funktioniert haben. Ein besonderer Dank gilt Dr. Ralf Florian von Cube, Ralf Schmieder, Robin Hofsaess und Svenja Moench für das Korrekturlesen der Dissertation und ihre hilfreichen Anmerkungen. Desweiteren möchte ich allen, den bisher namentlich erwähnten, sowie unerwähnt gebliebenen, Mitgliedern der Arbeitsgruppe am ETP für das sehr gute Arbeitsklima während der vergangenen Jahre und die schöne Zeit danken. Besonders zu nennen sind hier die Kollegen aus 8/21, mit denen ich mir zu den Präsenzzeiten während der Pandemie das Büro teilen durfte.

Meinen Eltern würde ich gerne dafür danken, dass sie mir den Weg von der Schule und dem Studium bis hin zur Dissertation ermöglicht und mich immer unterstützt haben. Zu guter Letzt möchte ich noch meiner Freundin Svenja dafür danken, dass sie in den vergangen Jahren immer für mich da war und mich großartig unterstützt hat.

Eidesstattliche Versicherung gemäß der Promotionsordnung des Karlsruher Instituts für Technologie (KIT) für die KIT-Fakultät für Physik

1. Bei der eingereichten Dissertation zu dem Thema

»*A Search for Additional Neutral Higgs Bosons in $\tau\tau$ Final States in pp Collisions at $\sqrt{s} = 13$ TeV*«

handelt es sich um meine eigenständig erbrachte Leistung.

2. Ich habe nur die angegebenen Quellen und Hilfsmittel benutzt und mich keiner unzulässigen Hilfe Dritter bedient. Insbesondere habe ich wörtlich oder sinngemäß aus anderen Werken übernommene Inhalte als solche kenntlich gemacht.

3. Die Arbeit oder Teile davon habe ich bislang nicht an einer Hochschule des In- oder Auslands als Bestandteil einer Prüfungs- oder Qualifikationsleistung vorgelegt.

4. Die Richtigkeit der vorstehenden Erklärungen bestätige ich.

5. Die Bedeutung der eidesstattlichen Versicherung und die strafrechtlichen Folgen einer unrichtigen oder unvollständigen eidesstattlichen Versicherung sind mir bekannt.

Ich versichere an Eides statt, dass ich nach bestem Wissen die reine Wahrheit erklärt und nichts verschwiegen habe.

Maximilian Burkart
Karlsruhe, den 12. Oktober 2022