TU TECHNISCHE UNIVERSITÄT WIEN

DISSERTATION

Implementation and Application of Machine Learning Techniques for the Analysis of Higgs Boson Decays to Tau Leptons with the CMS experiment

AUSGEFÜHRT ZUM ZWECKE DER ERLANGUNG DES AKADEMISCHEN GRADES
EINES DOKTORS DER TECHNISCHEN WISSENSCHAFTEN
UNTER DER LEITUNG VON

Univ.Prof. Dipl.-Phys. Dr.rer.nat. Jochen Schieck

und

Dr. Martin Flechl

als verantwortlich mitwirkender Assistent

am Atominstitut der Technischen Universität Wien (E141)
in Zusammenarbeit mit dem Institut für Hochenergiephysik (HEPHY)
der Österreichischen Akademie der Wissenschaften (ÖAW)

eingereicht an der
TECHNISCHEN UNIVERSITÄT WIEN
FAKULTÄT FÜR PHYSIK

von

Dipl.-Ing. Markus Spanring

Matrikelnummer: 00926101
Hauslehen 78, 3342 Opponitz

Wien, am 22.10.2019

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

TU Bibliothek
Your knowledge hub
WIEN

Kurzfassung

In dieser Dissertation werden zwei Messungen von Higgs-Boson-Zerfällen zu Tau-Leptonen präsentiert. Hierfür werden Daten von Proton-Proton-Kollisionen am CERN Large Hadron Collider analysiert, welche vom Compact Muon Solenoid (CMS)-Experiment in den Jahren 2016 und 2017 bei einer Schwerpunktsenergie von $13\,\mathrm{TeV}$ aufgezeichnet wurden. Die Daten entsprechen einer integrierten Luminosität von $35.9\,\mathrm{fb}^{-1}$ und $41.5\,\mathrm{fb}^{-1}$ für das Jahr 2016 bzw. 2017.

Im ersten Teil der Arbeit wird der inklusive Wirkungsquerschnitt, σ_{incl}, des Zerfalls von einem Higgs-Boson in ein Tau-Lepton-Paar gemessen. Weiters wird der Wirkungsquerschnitt abhängig von den individuellen Produktions-Mechanismen sowie in diversen kinematischen Regionen des Higgs-Bosons betrachtet. Für die Trennung von Untergrund- und Signal-Prozessen wird ein Klassifikationsalgorithmus, welcher auf einem künstlichen neuronalen Netz basiert, eingesetzt. Das Ergebnis der Messung des inklusiven Wirkungsquerschnittes lautet $\sigma_{\mathrm{incl}}\mathcal{B}(\mathrm{H} \rightarrow \tau^{-}\tau^{+}) = 2.56\,\mathrm{pb} \pm 0.48\,(\mathrm{stat})\,\mathrm{pb} \pm 0.34\,(\mathrm{syst})\,\mathrm{pb}$.

Im zweiten Teil wird die Suche nach zusätzlichen neutralen Higgs-Bosonen, welche in Tau-Leptonen zerfallen, im Rahmen des minimalen supersymmetrischen Standardmodells (MSSM) vorgestellt. Es wurde keine signifikante Abweichung von der Untergrund-Hypothese beobachtet. Die durchgeführten Messungen werden verwendet, um modellunabhängige Ausschlussgrenzen für die Produktion eines Higgs-Bosons durch Gluon-Fusion, sowie in Kombination mit Bottom-Quarks, zu bestimmen. Weiters wird das Ergebnis innerhalb von Theorien jenseits des Standardmodells interpretiert.

Abstract

Two measurements on Higgs boson decays to tau leptons are presented in this thesis. They are based on proton-proton collisions at the CERN Large Hadron Collider recorded by the Compact Muon Solenoid (CMS) experiment in 2016 and 2017, corresponding to an integrated luminosity of $35.9\,\mathrm{fb}^{-1}$ and $41.5\,\mathrm{fb}^{-1}$ at a centre-of-mass energy of $13\,\mathrm{TeV}$, respectively.

A measurement of the inclusive cross section, σ_{incl}, of a Higgs boson decay to a pair of tau leptons, $\mathrm{H} \rightarrow \tau^{-}\tau^{+}$, is presented. The results are also presented in terms of cross sections for individual Higgs boson production modes and kinematic regimes. To differentiate between signal and background processes a neural network classification algorithm is used. A value of $\sigma_{\mathrm{incl}}\mathcal{B}(\mathrm{H} \rightarrow \tau^{-}\tau^{+}) = 2.56\,\mathrm{pb} \pm 0.48\,(\mathrm{stat})\,\mathrm{pb} \pm 0.34\,(\mathrm{syst})\,\mathrm{pb}$ is obtained for the inclusive measurement.

A search for additional neutral Higgs bosons decaying to a pair of tau leptons is presented in the context of the minimal supersymmetric extension of the standard model (MSSM). No significant deviation with respect to the background-only hypothesis is observed. Model-independent upper limits are set on the product of the Higgs boson production cross section via gluon fusion or in association with bottom quarks and the branching fraction for the decay to a pair of tau leptons. Furthermore, exclusion contours in the context of selected MSSM benchmark scenarios are provided.

Acknowledgements

CONTENTS

Contents

CHAPTER 1

INTRODUCTION

The standard model (SM) of particle physics is currently the most accurate description of elementary particles and their interaction via the strong, weak and electromagnetic force. With the discovery of a scalar particle that is compatible with the SM Higgs boson at the Large Hadron Collider (LHC) in July 2012 [1–3] the last missing piece and cornerstone of the SM was found. The Higgs boson is predicted by the Brout-Englert-Higgs (BEH) mechanism [4–7], which is the mechanism for spontaneous symmetry breaking in the SM. In the SM the couplings of the Higgs boson to vector bosons enter via the BEH mechanism and the couplings to fermions are introduced as Yukawa couplings. The vector boson and fermion couplings to the Higgs boson therefore need to be characterized independently. Theories beyond the standard model (BSM), of which supersymmetry (SUSY) [8] is one of the most popular, predict deviations of the Higgs boson couplings to down type fermions, such as the tau (τ) lepton.

The Higgs boson decay to τ leptons has the second largest branching fraction of the fermionic Higgs boson decays. Both, the ATLAS [9] and CMS [10] collaboration, have reported the observation of the Higgs boson in this final state. Compared to the decay to bottom quarks, which is the dominant Higgs boson decay mode, the τ lepton decay benefits from the distinct signature in the detector, which yields better experimental accessibility. This final state is therefore an ideal candidate for precision measurements of Yukawa couplings of the SM Higgs boson and furthermore to search for signatures of physics beyond the SM. This thesis exploits data recorded by the Compact Muon Solenoid (CMS) experiment in 2016 and 2017, corresponding to an integrated luminosity of $35.9\,\mathrm{fb}^{-1}$ and $41.5\,\mathrm{fb}^{-1}$, at a centre-of-mass energy of $13\,\mathrm{TeV}$ to study the $\tau^-\tau^+$ final state.

The first analysis is a cross section measurement of $\mathrm{pp} \to \mathrm{H} \to \tau^-\tau^+$ at different levels of differentiation. This includes an inclusive measurement, a measurement split by production mode and a measurement in different kinematic regimes as well as jet multiplicities in form of simplified template cross sections [11]. Furthermore, a new method (with respect to previous publications in the $\tau^-\tau^+$ final

state) to differentiate between signal and background processes in form of a neural network (NN) classification algorithm is introduced [12].

In the second analysis, a search for additional neutral Higgs bosons decaying to a pair of τ leptons is presented in context of the minimal supersymmetric extension of the standard model (MSSM). The results are presented as model-independent upper limits on the product of the Higgs boson production cross section via gluon fusion or in association with bottom quarks and the branching fraction for the decay to a pair of τ leptons. Furthermore, the measurements are interpreted in selected MSSM benchmark scenarios in form of exclusion contours [13].

The thesis is structured as follows. In Chapter 2 a short introduction to the standard model of particle physics with focus on the Higgs boson is given. Furthermore, the Higgs sector of the minimal supersymmetric extension of the standard model is briefly discussed. An overview of the CMS detector at the LHC is given in Chapter 3. Chapter 4 outlines the Monte Carlo (MC) event generation and the physics object reconstruction. The measurement of Higgs boson production and decay to the $\tau^-\tau^+$ final state is presented in Chapter 5, followed by the search for additional neutral Higgs bosons in the $\tau^-\tau^+$ final state in Chapter 6. The thesis concludes in Chapter 7 with a summary and an outlook for future improvements of the discussed analyses.

CHAPTER 2

THEORY AND MOTIVATION

> "*The story so far: In the beginning the Universe was created. This has made a lot of people very angry and been widely regarded as a bad move.*"
>
> – THE RESTAURANT AT THE END OF THE UNIVERSE

The goal of this chapter is to briefly introduce the theory behind the standard model of particle physics, with emphasis on electroweak symmetry breaking and the Higgs boson, to motivate the analysis presented in Chapter 5. Furthermore, the minimal supersymmetric extension of the standard model is introduced which is the theory motivating the search presented in Chapter 6.

2.1 Standard Model of Particle Physics

The SM [14–18] is the result of a long period of persistent experimental work and theoretical developments in elementary particle physics, which changed the understanding of how nature interacts on the most fundamental level. It describes the strong, weak and electromagnetic interactions, which arise due the exchange of various spin-1 bosons amongst the spin-$1/2$ particles (fermions) that make up matter. Gravitation, the fourth of the currently known fundamental interactions, can be described in the framework of general relativity but is not incorporated in the SM. A central piece of the SM is the spin-0 particle, which arises due to electroweak symmetry breaking, the Higgs boson. The Higgs field is responsible for the masses of the $W^{\pm}$ and Z bosons, and of the fermions. Figure 2.1 shows a summary of all particles in the SM.

2.1.1 Particles of the Standard Model

In the following a short description of the properties of the fundamental fermions, which can be grouped into quarks and leptons, and the gauge bosons is given.
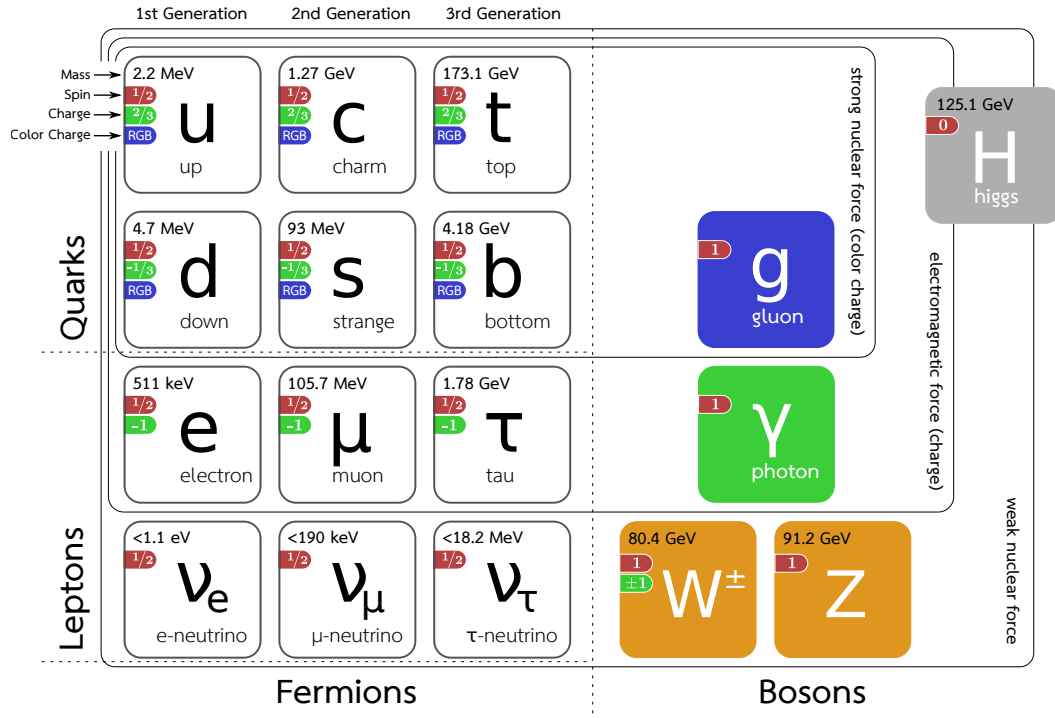
Figure 2.1 − Overview of the particles in the standard model. Numbers are taken from Refs. [19, 20]

Additionally, the tau lepton, which plays an essential role in this thesis, is discussed in more detail.

Gauge Bosons

The standard model is a gauge theory based on the

$$
\begin{array}{ccccc}
SU_c(3) & \times & SU_L(2) & \times & U_Y(1) \\
\downarrow & & \downarrow & & \downarrow \\
G_\mu^\alpha & & W_\mu^a & & B_\mu \\
\alpha = 1, \dots, 8 & & a = 1, 2, 3 & &
\end{array}
\tag{2.1}
$$

symmetry group. The eight massless spin-1 gauge bosons G_μ^α associated to the group $SU_c(3)$, where the subscript "c" denotes colour, are called *gluons*. They are the mediator particles of the strong interaction. Any particle that carries colour charge couples to gluons and is said to be "strongly interacting". The corresponding quantum field theory is called Quantum Chromodynamics (QCD), which is described in more detail in Ref. [14]. The $SU_L(2) \times U_Y(1)$ symmetry group describes the electroweak interaction, which is the unification of the weak

and electromagnetic interaction, as introduced by Glashow [16], Weinberg [17] and Salam [18]. The subscript in $SU_L(2)$ indicates that only left-handed fermions carry this quantum number. In $U_Y(1)$ the subscript indicates that the group is associated to the weak hypercharge Y, which is related to the electric charge and the third component of the weak isospin, to distinguish between the $U(1)_{\text{em}}$ of Quantum Electrodynamics (QED) [21]. The four spin-1 bosons associated to this symmetry group, W_μ^a and B_μ, are related to the physical massive mediator particles of the weak interactions, $\text{W}^\pm$ and Z, and the massless photon (A) from QED via

$$W_\mu^\pm = \frac{1}{\sqrt{2}} \left(W_\mu^1 \mp W_\mu^2 \right),$$

$$Z_\mu = W_\mu^3 \cos\theta_W - B_\mu \sin\theta_W, \qquad (2.2)$$

$$A_\mu = W_\mu^3 \sin\theta_W + B_\mu \cos\theta_W,$$

respectively. The Weinberg or weak mixing angle θ_W is related to the weak (g_1) and electromagnetic (g_2) coupling constants according to

$$\sin\theta_W = \frac{g_1}{\sqrt{g_1^2 + g_2^2}} \quad \text{and} \quad \cos\theta_W = \frac{g_2}{\sqrt{g_1^2 + g_2^2}}. \qquad (2.3)$$

Quarks

Quarks are subject to the strong, weak, and electromagnetic interaction. There are six different quark flavours, which may be subdivided into three generations as shown in Table 2.1. Quarks carry an electric charge of either $^2/_3$ or $-^1/_3$ and

Generation	Quark flavour	Mass	Charge (q_{e})	Symbol
First	Up	$2.16^{+0.49}_{-0.26}\,\text{MeV}$	$^2/_3$	u
	Down	$4.67^{+0.48}_{-0.17}\,\text{MeV}$	$-^1/_3$	d
Second	Charm	$1.27 \pm 0.02\,\text{GeV}$	$^2/_3$	c
	Strange	$93^{+11}_{-5}\,\text{MeV}$	$-^1/_3$	s
Third	Top	$4.18^{+0.03}_{-0.02}\,\text{GeV}$	$^2/_3$	t
	Bottom	$173.1 \pm 0.9\,\text{GeV}$	$-^1/_3$	b

Table 2.1 – The three quark generations and their properties. The charge is given in units of the Coulomb charge q_{e}. The masses of the u-, d- and s-quarks correspond to current quark masses at a scale $\mu \approx 2\,\text{GeV}$. The other quark masses correspond to running masses in the $\overline{\text{MS}}$ scheme [19, 22, 23].

furthermore a colour charge, which may be one out of three values. The term

colour is not to be understood in the sense of actual colours but used to describe the corresponding quantum number. Antiquarks have the corresponding negative electric charge as well as anti-colour and are indicated with a bar over the symbol, such as $\bar{t}$ for a top antiquark. Due to colour confinement, which implies that a particle with colour charge cannot be isolated, quarks can only be observed in combination with other quarks. These composite particles which do not carry any net colour charge are called *hadrons*. Hadrons are quark systems composed of three quarks, called *baryons*, and systems of a quark-antiquark pair, called *mesons*, bound by the strong interaction via the exchange of gluons. The baryon number B of a quark system is strictly conserved. It is 1 for baryons and 0 for mesons and is defined as

$$B = \frac{1}{3}(N_q - N_{\bar{q}}),\tag{2.4}$$

where N_q is the number of quarks and $N_{\bar{q}}$ is the number of antiquarks. However, the flavour of the individual quarks may change due to the weak interaction. Prominent baryon representatives are the proton (uud), which is stable within the current experimental resolution, and the neutron (udd), which decays if not in a bound state within about 15 min to a proton through the weak interaction [19]. Mesons are unstable and the lightest representatives are the *pions* or π-mesons. The π^- and π^+ have a mass of 139.57 MeV and are made up of (u$\bar{\text{d}}$) and ($\bar{\text{u}}$d) pairs respectively. The π^0 is a coherent superposition of (u$\bar{\text{u}}$) and (d$\bar{\text{d}}$) and has a mass of 134.98 MeV.

In particle detectors the signature of a quark emerges in form of a *jet*. A jet is a spray of collimated hadrons and other particles, which is produced by hadronization of quarks or gluons during particle collisions. As described above, particles with colour charge cannot be isolated due to colour confinement. This leads to the creation of other coloured particles, which move in a narrow cone around the initial particle and in the the same direction, to form hadrons. The ensemble of these particles is called jet. Out of the six quark flavours the t quark is the only flavour that decays before a jet can form.

A useful model for the interpretation of high-energy hadron collisions is the so-called *parton* model. In this model, hadrons are described as composite particles of point-like constituents, such as quarks and gluons, called *partons*. During a collision, each parton is considered to carry a certain fraction of the hadron momentum. The probability density for finding a parton with such a momentum is described by a parton density function (PDF).

Leptons

Leptons are subject to the weak interaction and may be subdivided into three generations as shown in Table 2.2. Each generation consists of a negatively charged

Generation	Name	Mass	Charge (q_e)	Symbol
First	Electron	$511\,\mathrm{keV}$	-1	e^-
	e neutrino	$< 1.1\,\mathrm{eV}$	0	ν_e
Second	Muon	$105.7\,\mathrm{MeV}$	-1	μ^-
	μ neutrino	$< 190\,\mathrm{keV}$	0	ν_μ
Third	Tau lepton	$1.78\,\mathrm{GeV}$	-1	τ^-
	τ neutrino	$< 18.2\,\mathrm{MeV}$	0	ν_τ

Table 2.2 – The three lepton generations and their properties. The charge is given in units of the Coulomb charge q_e. The neutrino masses are known to be tiny compared to other SM particles but non-zero [19, 20]

.

lepton and a corresponding neutrino. The charged leptons are, in addition to the weak interaction, also subject to the electromagnetic interaction. Each lepton has a corresponding anti-particle indicated by either a positive electric charge, e.g. e^+ for the positron, or a bar over the symbol, e.g. $\overline{\nu}_e$ for the e anti-neutrino. Throughout this thesis, ℓ refers to the light leptons, electron and muon, and τ refers to the tau lepton. Note that the indication of the charge is omitted for the sake of simplicity. Therefore, a process such as the decay of a Higgs boson to a pair of oppositely charged τ leptons, $H \rightarrow \tau^-\tau^+$, is written in the simplified form $H \rightarrow \tau\tau$.

Tau Lepton

The τ lepton is the heaviest member of the lepton family with a mass of $1.78\,\mathrm{GeV}$ and has a mean lifetime of about $2.9 \cdot 10^{-13}\,\mathrm{s}$ [19]. First evidence for its existence was found in 1974 by Martin Perl [24] and the Stanford Linear Accelerator Center-Lawrence Berkeley Laboratories (SLAC-LBL) collaboration at the SPEAR e^+e^- collider. Perl and his colleagues were searching for new charged leptons, in addition to the already discovered muon, in e^+e^- collisions, with the basic idea that they will be pair produced ($e^+e^- \rightarrow \tau^+\tau^-$) if the e^+e^- energy is larger than twice the mass of the new lepton. The signature they were looking for was of the form

$$e^+ + e^- \xrightarrow{?} e^\pm + \mu^\mp + \text{at least two undetected particles,} \tag{2.5}$$

showing apparent lepton flavour violation accompanied by missing energy and momentum. This violation may be resolved if an undiscovered lepton is pair produced in an intermediate step, subsequently decaying to one electron and one muon via the weak interaction:

$$\mathrm{e}^+ + \mathrm{e}^- \to \tau^+ + \tau^- \to \mathrm{e}^+ \nu_\mathrm{e} \overline{\nu}_\tau + \mu^- \overline{\nu}_\mu \nu_\tau. \tag{2.6}$$

In a series of experiments between 1974 and 1977 the discovery of the τ lepton was finally confirmed [25–28] and Perl was awarded the Nobel Prize in 1995.

Due to the short lifetime of the τ lepton it cannot be detected directly and is therefore reconstructed according to the decay products in the final state. The most frequent decay modes are listed in Table 2.3. The leptonic decay to an

Decay mode	Meson resonance	$\mathcal{B}$ [%]
Leptonic decay modes		35.2
$\tau^- \to \mathrm{e}^- \overline{\nu}_\mathrm{e} \nu_\tau$		17.8
$\tau^- \to \mu^- \overline{\nu}_\mu \nu_\tau$		17.4
Hadronic decay modes		64.8
$\tau^- \to h^- \nu_\tau$		11.5
$\tau^- \to h^- \pi^0 \nu_\tau$	$\rho(770)$	25.9
$\tau^- \to h^- \pi^0 \pi^0 \nu_\tau$	$\mathrm{a}_1(1260)$	9.5
$\tau^- \to h^- h^+ h^- \nu_\tau$	$\mathrm{a}_1(1260)$	9.8
$\tau^- \to h^- h^+ h^- \pi^0 \nu_\tau$		4.8
Other		3.3

Table 2.3 – Branching fractions ($\mathcal{B}$) of different τ decay modes. The symbol h^- denotes a charged pion or charged kaon. Hadronic τ decays that proceed via an intermediate meson resonance have the corresponding resonance assigned [19].

electron as well as the decay to a muon, accompanied by the respective neutrinos, have approximately the same rate due to lepton universality. Due to the large τ mass, which exceeds the mass of the lightest hadrons, about 2/3 of all τ decays are to hadrons. The hadronic τ decays are indicated as τ_h throughout this thesis.

2.1.2 Brout-Englert-Higgs Mechanism

The $SU(2) \times U(1)$ gauge group as introduced in Section 2.1.1 describes the electroweak interactions. A major shortcoming is that the local gauge symmetry does

not allow mass terms for the gauge boson fields. However, measurements [19] found the W and Z bosons to be massive. To circumvent this shortcoming the mechanism of spontaneous symmetry breaking, achieved via the Brout-Englert-Higgs mechanism, is introduced. Symmetry is broken by a field that is symmetric under gauge transformations, but has a non-vanishing vacuum expectation value. In the following the mechanism is applied to an Abelian, $U(1)$ gauge theory, to demonstrate how the corresponding gauge boson acquires mass. A more detailed description can be found in Refs. [14, 29].

Global Symmetry Breaking and Goldstone Bosons

For a complex field Φ a possible Lagrangian density is

$$\mathcal{L} = (\partial_t \Phi^\dagger)(\partial_t \Phi) - (\nabla \Phi^\dagger \cdot \nabla \Phi + m^2 \Phi^\dagger \Phi) \quad \text{with} \quad \Phi = \frac{(\phi_1 + i\phi_2)}{\sqrt{2}}. \qquad (2.7)$$

The term $(\partial_t \Phi^\dagger)(\partial_t \Phi)$ can be regarded as the kinetic energy density and $\nabla \Phi^\dagger \cdot \nabla \Phi + m^2 \Phi^\dagger \Phi$ as the potential energy density. The derivatives vanish if Φ is constant in space and time, which leaves $m^2 \Phi^\dagger \Phi$ as the only contribution to the energy. Therefore, for positive m^2, the energy has a minimum at $\phi_1 = \phi_2 = 0$, which means that $\Phi = 0$ corresponds to the *vacuum state*. For negative m^2 the potential is unbounded from below which leads to an unstable Lagrangian density. By introducing a potential term of the form

$$V(\Phi^\dagger \Phi) = \frac{m^2}{2\phi_0^2} \left[\Phi^\dagger \Phi - \phi_0^2 \right]^2, \qquad (2.8)$$

where ϕ_0^2 is a real parameter, stability can be restored. The modified Lagrangian, given by

$$\mathcal{L} = \partial_\mu \Phi^\dagger \partial^\mu \Phi - V(\Phi^\dagger \Phi) \quad \text{with} \quad \partial_\mu = \partial_t - \nabla, \qquad (2.9)$$

has a minimum at constant Φ that is independent of space and time at $\Phi^\dagger \Phi = |\Phi|^2 = \phi_0^2$. However, this minimum is not unique since there exists an infinite number of solutions, each defined by a point on the circle $|\Phi| = \phi_0 = \sqrt{\phi_1^2 + \phi_2^2}$, indicated by the solid red contour line in Fig. 2.2. The Lagrangian density has a *global $U(1)$* symmetry:

$$\begin{aligned} \Phi \to \Phi' = e^{-i\alpha} \Phi \quad &\text{with} \quad \alpha \in \mathbb{R} \\ \mathcal{L} \to \mathcal{L}' = \mathcal{L}. \end{aligned} \qquad (2.10)$$

The transformation is a rotation in the state space (ϕ_1, ϕ_2) around a circle $|\Phi|^2 =$ constant. For a particular direction in the (ϕ_1, ϕ_2) space for which Φ is real and the
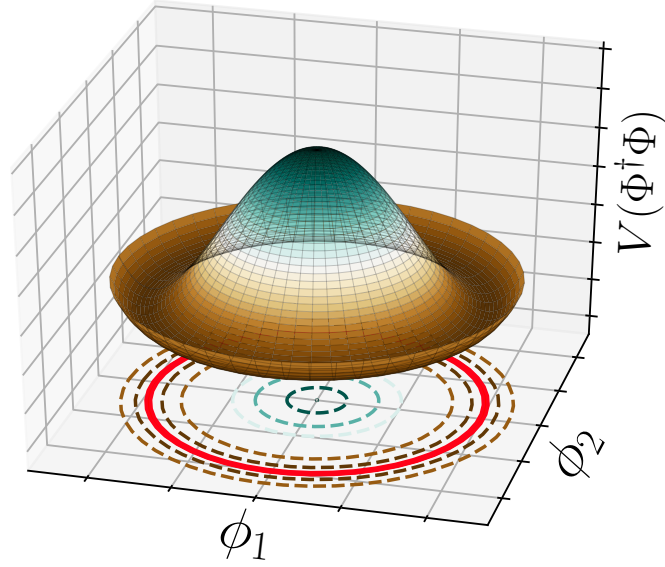
Figure 2.2 – The Higgs potential $V(\Phi^\dagger\Phi)$ with the symmetric "mexican hat" shape. There is an infinite number of vacuum states, each defined by a point on the red solid contour line.

vacuum state is chosen to be $(\phi_0, 0)$, the $U(1)$ symmetry is broken. The expansion of this ground state has the form

$$\Phi = \phi_0 + \frac{1}{\sqrt{2}}\left(\chi + i\psi\right), \tag{2.11}$$

with two scalar real fields χ and ψ, for which the Lagrangian density becomes

$$\mathcal{L} = \frac{1}{2}\partial_\mu\chi\partial^\mu\chi + \frac{1}{2}\partial_\mu\psi\partial^\mu\psi - \frac{m^2}{2\phi_0^2}\left[\sqrt{2}\phi_0\chi + \frac{\chi^2}{2} + \frac{\psi^2}{2}\right]^2. \tag{2.12}$$

The Lagrangian density can now be divided into $\mathcal{L}_{\text{free}}$, which contains all quadratic terms in $\mathcal{L}$ and represents free particle fields, and $\mathcal{L}_{\text{int}}$, which corresponds to interactions between the free particles and higher order corrections to their motion. This gives

$$\mathcal{L} = \mathcal{L}_{\text{free}} + \mathcal{L}_{\text{int}} \tag{2.13}$$

with

$$\mathcal{L}_{\text{free}} = \frac{1}{2}\left[\partial_\mu\chi\partial^\mu\chi - 2m^2\chi^2\right] + \frac{1}{2}\partial_\mu\psi\partial^\mu\psi. \tag{2.14}$$

The quadratic term $-2m^2\chi^2$ in $\mathcal{L}_{\text{free}}$ corresponds to a scalar spin-0 particle χ with mass $\sqrt{2}m$ and ψ corresponds to a massless spin-0 particle. The massless particle is called Goldstone boson [30] and arises as a result of global symmetry breaking.

Local Symmetry Breaking and the Higgs Boson

The previously described symmetry breaking can be generalised further by constructing a Lagrangian density that is invariant under a *local* $U(1)$ gauge transformation

$$\Phi(x) \to \Phi'(x) = e^{-iq\theta(x)}\Phi(x),$$
$$A_\mu(x) \to A'_\mu(x) = A_\mu(x) + \partial_\mu\theta(x),$$

(2.15)

by introducing a massless gauge field A_μ corresponding to the four-potential from QED. The resulting Lagrangian density is

$$\mathcal{L} = \left[(\partial_\mu - iqA_\mu)\Phi^\dagger\right]\left[(\partial^\mu - iqA^\mu)\Phi\right] - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} - V(\Phi^\dagger\Phi) \tag{2.16}$$

where $F_{\mu\nu} = \partial_\mu A_\nu - \partial^\mu A^\nu$ is the electromagnetic field strength tensor, q is the electromagnetic coupling constant and $V(\Phi^\dagger\Phi)$ is the potential introduced in Eq. (2.8). The Lagrangian density has a minimum when the field A_μ vanishes and Φ is constant and any gauge transformation on this field configuration is also a minimum. This yields an infinite number of vacuum states, each defined by a point on the circle $|\Phi| = \phi_0$. It is always possible to choose $\theta(x)$ such that the field $\Phi'(x) = e^{-iq\theta(x)}\Phi(x)$ is real, which breaks the symmetry. Putting

$$\Phi'(x) = \phi_0 + \frac{1}{\sqrt{2}}h(x), \tag{2.17}$$

where $h(x)$ is real, into the Lagrangian defined in Eq. (2.16) and splitting it into the free particle terms and into the interaction terms results in

$$\mathcal{L}_{\text{free}} = \frac{1}{2}\left[\partial_\mu h\partial^\mu h - 2m^2 h^2\right] - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} + q^2\phi_0^2 A_\mu A^\mu,$$
$$\mathcal{L}_{\text{int}} = q^2 A_\mu A^\mu\left(\sqrt{2}\phi_0 h + \frac{1}{2}h^2\right) - \frac{m^2 h^2}{2\phi_0^2}\left(\sqrt{2}\phi_0 h + \frac{1}{4}h^2\right).$$

(2.18)

To summarize, we constructed a Lorentz invariant Lagrangian density, which, by introducing the gauge field A_μ, is also invariant under a local $U(1)$ transformation. After local symmetry breaking, a single scalar field $h(x)$ corresponding to a spin-0 particle of mass $\sqrt{2}m$ and a vector field A_μ corresponding to a vector boson of mass $\sqrt{2}q\phi_0$ emerge. The field $h(x)$ is generally referred to as *Higgs boson* (H). This mechanism to introduce mass into a theory was first proposed by Anderson [4] and introduced by three independent groups in 1964: by Brout and Englert [5], Higgs [6], and Guralnik, Hagen and Kibble [7]. In a further generalization to the non-abelian case it is possible to construct a Lagrangian density that is invariant under local $SU(2)$ as well as a local $U(1)$ transformation, which is the basis for

the SM. By breaking the $SU(2)$ symmetry and replacing the fields W_μ^a and B_μ with the expresions from Eq. (2.2) one arrives at the Lagrangian density

$$\mathcal{L} = \mathcal{L}_1 + \mathcal{L}_2 \tag{2.19}$$

where

$$
\begin{aligned}
\mathcal{L}_1 = & \frac{1}{2} \left(\partial_\mu h \partial^\mu h - M_H^2 h^2 \right) \\
& - \frac{1}{4} Z_{\mu\nu} Z^{\mu\nu} + \frac{1}{2} M_Z^2 Z_\mu Z^\mu \\
& - \frac{1}{4} A_{\mu\nu} A^{\mu\nu} \\
& - \frac{1}{2} \left[\left(D_\mu W_\nu^+ \right)^* - \left(D_\nu W_\mu^+ \right)^* \right] \left[\left(D^\mu W^{+\nu} \right) - \left(D^\nu W^{+\mu} \right) \right] + M_W^2 W_\mu^- W^{+\mu},
\end{aligned}
\tag{2.20}
$$

with

$$D_\mu W_\nu^+ = \left(\partial_\mu + i \overbrace{g_2 \sin \theta_W}^{q} A_\mu \right) W_\nu^+. \tag{2.21}$$

The individual terms can be interpreted as a free massive neutral scalar boson field $h(x)$, a massive neutral scalar field $Z_\mu(x)$, and a pair of massive charged vector boson fields $W_\mu^+(x)$ and $W_\mu^-(x)$, interacting with the electromagnetic field $A_\mu(x)$. $\mathcal{L}_2$ describes the remaining interactions and the terms that are needed to make the theory renormalizable. We can now identify the three massive scalar fields $Z_\mu(x), W_\mu^+(x)$ and $W_\mu^-(x)$ as the Z, W^+ and W^- gauge bosons of the weak interactions, with the experimentally determined masses [19]

$$
\begin{aligned}
M_W &= 80.379 \pm 0.012 \,\text{GeV}, \\
M_Z &= 91.1876 \pm 0.0021 \,\text{GeV}.
\end{aligned}
\tag{2.22}
$$

The scalar field h corresponds to the Higgs boson. It is possible to add mass terms for fermions via Yukawa couplings between the fermions and the Higgs field. These couplings are of the form

$$\lambda_f \left(\bar{\psi}_L \Phi \psi_R + \bar{\psi}_R \Phi^\dagger \psi_L \right), \tag{2.23}$$

where ψ_L is the left-handed spin-doublet, ψ_R the right-handed singlet, and Φ corresponds to the Higgs boson field. The parameter λ_f is a coupling constant that is proportional to the mass of each fermion, $\lambda_f \propto m_f$, implicating that heavier fermions couple more strongly to the Higgs field. Note that the coupling to bosons is proportional to the squared mass of each boson $\lambda_V \propto m_V^2$.

2.1.3 Standard Model Higgs Boson Discovery and Measurements

With the advent of powerful particle colliders, starting with the Large Electron-Positron Collider (LEP) at CERN and followed by the Tevatron at Fermilab, extensive searches for the Higgs boson were conducted. The searches performed at both collider experiments did not lead to the observation of the Higgs boson, however, certain mass ranges were excluded for a Higgs boson as predicted by the SM. The combined results from LEP excluded a Higgs boson with a mass $M_H < 114.4\,\text{GeV}$ at the 95 % confidence level (CL) [31]. Searches at the Tevatron resulted in the exclusion of an additional mass range of 147–180 GeV at the 95 % CL [32]. Until 2012, results from data collected at the LHC with the ATLAS and CMS experiments [33, 34] in combination with results from the LEP and Tevatron experiments excluded a wide mass range, as shown in Fig. 2.3. However, there was evidence for an excess between 115–130 GeV.
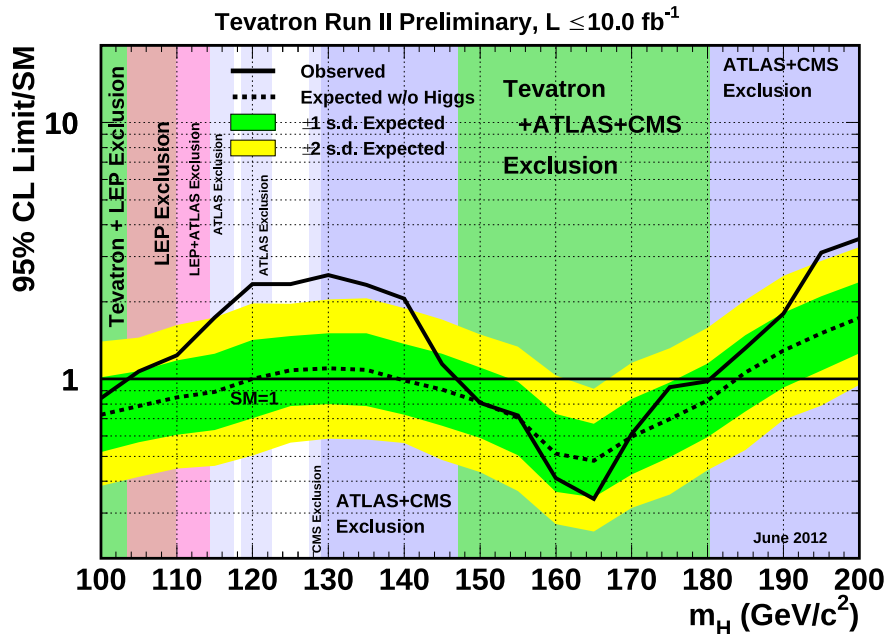
Figure 2.3 – Observed and expected 95 % CL upper cross section limits on the ratio to the SM cross section, as functions of the Higgs boson mass for the Tevatron. The mass regions that have been excluded by analyses performed at LEP, Tevatron and at the LHC are indicated by the vertical coloured bands [32].

Finally, on July 4th, 2012 the discovery of a new boson was announced by the ATLAS and CMS experiments at a mass of approximately 125 GeV [1–3]. The discovery was driven by the H $\rightarrow \gamma\gamma$ and H $\rightarrow$ ZZ $\rightarrow 4\ell$ decay channels and

was based on data collected at centre-of-mass energies of 7 TeV and 8 TeV, corresponding to $5\,\mathrm{fb}^{-1}$ and $6\,\mathrm{fb}^{-1}$, respectively. More data was collected and analysed before the LHC entered its first extensive upgrade phase. This allowed to measure several properties [35–40] of the newly found boson and increase the confidence of its compatibility with the SM Higgs boson hypothesis.

Production of the Standard Model Higgs Boson

The dominant Higgs boson production mode at the LHC is via gluon fusion (ggF). Since gluons are massless and do not couple directly to the Higgs field, the coupling proceeds via a quark loop as shown in Fig. 2.4a. This loop is dominated by the
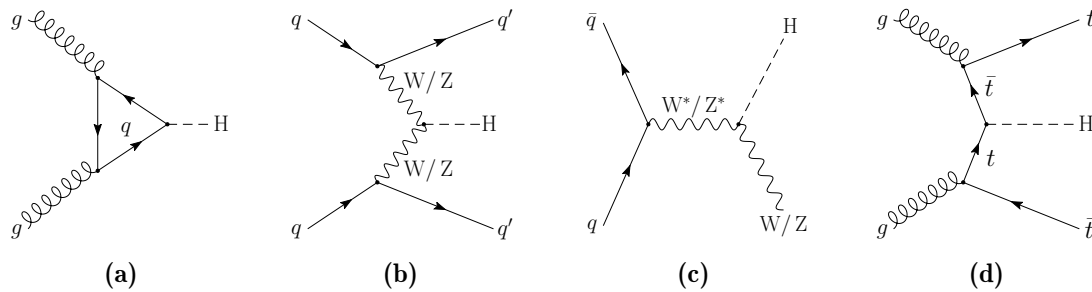
Figure 2.4 – Leading order Feynman diagrams for the dominant Higgs boson production modes at the LHC. Shown are gluon fusion (a), vector boson fusion (b), Higgs strahlung (c) and associated production with t quarks (d) .

contribution of top quarks, since they have the largest mass and therefore the largest coupling to the Higgs boson. The next important production mode with a cross-section that is almost a magnitude smaller than the ggF cross-section is vector boson fusion (VBF). In this process, which is shown in Fig. 2.4b, two quarks each radiate a W or Z boson which subsequently fuse to produce the Higgs boson. This production mode exhibits a distinct event signature due to the scattered quarks, which is rarely observed for other SM processes and may be exploited to specifically tag VBF events. Additional production modes are Higgs strahlung (VH), shown in Fig. 2.4c, where a Higgs boson is produced in association with a W or Z boson, and Higgs production in association with t quarks (ttH), which is shown in Fig. 2.4d. The corresponding production cross-section at a centre-of-mass energy of 13 TeV for each production mode is shown in Fig. 2.5.

Decay of the Standard Model Higgs Boson

The mass of the Higgs boson is a free parameter in the SM. The respective branching ratios are strongly dependent on the Higgs boson mass as shown in
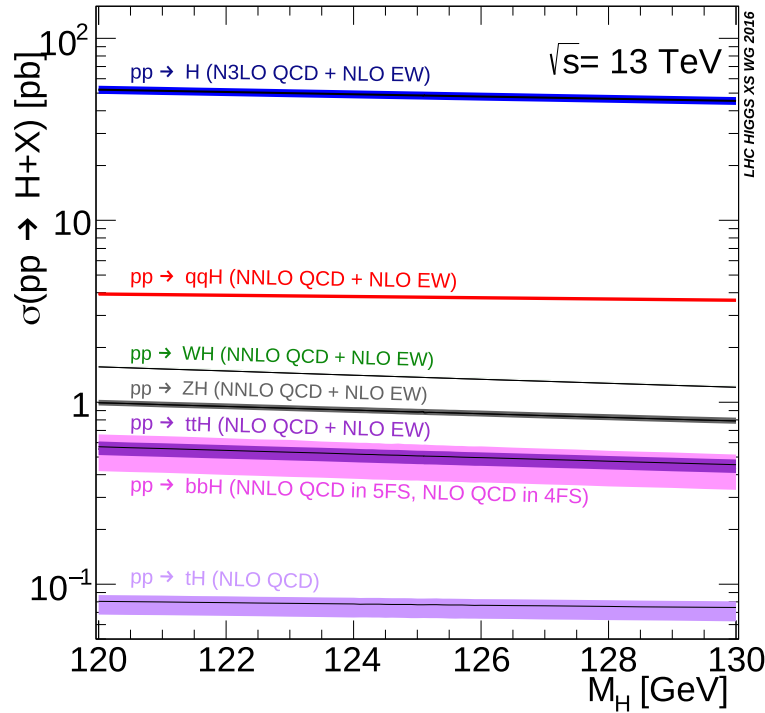
Figure 2.5 – The SM Higgs boson production cross section at a centre-of-mass energy of 13 TeV [11].

Fig. 2.6a. For the discovered Higgs boson at 125 GeV the dominant decay, with approximately 58 %, is into b quarks as shown in Fig. 2.6b. However, it is very challenging to separate the H → bb decay from the quark and gluon jets that are produced abundantly at the LHC. A similar problem arises for the decays into c quarks and gluons, which are even harder to separate from the jet background and have in addition a much lower branching fraction.

The decay of the Higgs boson into a pair of vector bosons (W or Z) is only possible if one vector boson is produced off-shell, since the Higgs boson mass is smaller than the combined mass of the decay products. Even though H → WW has the second largest branching fraction it suffers from suboptimal mass resolution due to the undetectable neutrinos from the W decay. The H → ZZ decay, however, provides excellent mass resolution and the signal arises over a mostly flat background. It is therefore a well suited decay mode for many Higgs boson property measurements and is often referred to as the "golden channel".

The decay into a pair of τ leptons is the second most common decay to fermions. The lower branching fraction compared to H → bb is compensated by the higher

experimental sensitivity of the leptonic and hadronic τ decays, which makes it the most promising candidate to measure Higgs fermion couplings.

Similar to the ggF production the Higgs boson decay to a pair of photons proceeds via an intermediate loop. This decay mode offers exceptional experimental precision due to the clean signature and the good photon energy resolution of the ATLAS and CMS detectors.

The decay modes with even smaller branching ratios, such as H $\to$ Zγ, H $\to$ $\mu\mu$, H $\to$ ee or decays involving light quark flavours, have a clean signature but require a significant amount of data in order to become sensitive.
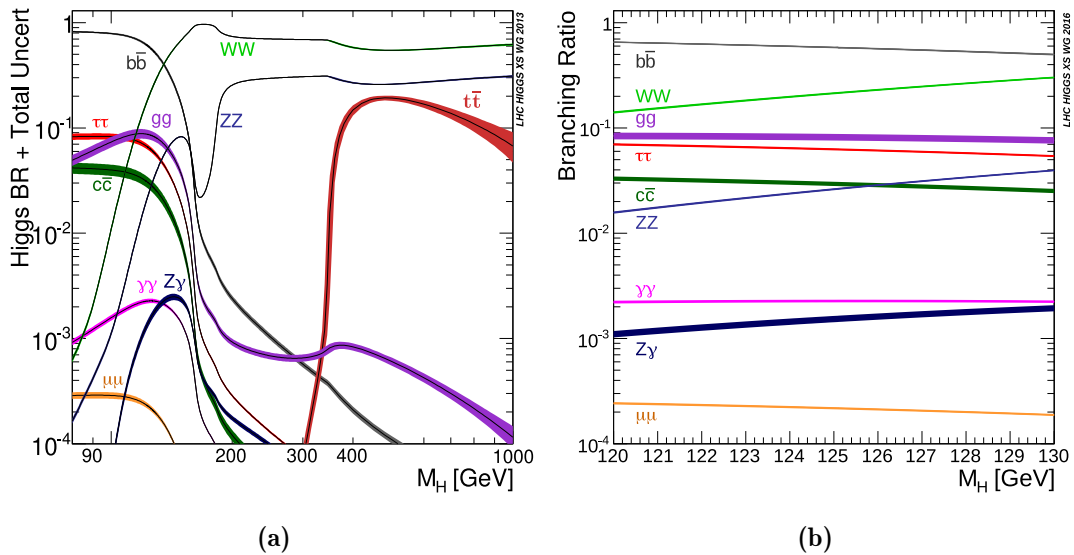
Figure 2.6 – Higgs boson branching ratios and their uncertainties for a Higgs boson mass between 80–1000 GeV (a) and around 125 GeV (b) [11].

The Kappa Framework

During LHC Run-1, which refers to the data-taking periods at a centre-of-mass energy of 7 TeV and 8 TeV from 2010-2012, the so-called *kappa framework* [41] was introduced by the ATLAS and CMS experiments. This framework was designed to explore specific aspects of the Higgs boson coupling structure that could be realistically studied with the LHC Run-1 datasets via a series of benchmark parametrizations. It should on the one hand provide a recommendation to experiments on how to perform couplings fits that are useful for theorists and on the other hand prepare the theory community for the experimental results.

In this framework all deviations from the SM are computed under the assumption that there is only one Higgs boson with a mass of 125 GeV and that the zero-width approximation is applicable for this state. With these assumptions, the signal cross section for all channels can be decomposed to

$$(\sigma \cdot \mathcal{B})(ii \to H \to ff) = \frac{\sigma_{ii} \cdot \Gamma_{ff}}{\Gamma_H}, \tag{2.24}$$

where σ_{ii} is the production cross section via the initial state ii, Γ_{ff} the partial decay width into the final state ff and Γ_H the total width of the Higgs boson.

The Higgs boson couplings are *pseudo-observables*, which means that they can not be measured directly. It is therefore necessary to extract the information on the couplings from quantities such as cross section times branching fraction via a certain *unfolding procedure*.

To take into account the best available SM predictions for Higgs boson cross sections while at the same time allowing possible deviations from the SM couplings, the predicted SM Higgs boson cross sections are dressed with scale factors κ_i. They are designed such that the corresponding cross sections σ_{ii} and decay widths Γ_{ii} of an SM particle i scale width κ_i^2 with respect to the SM prediction. To give an example, the $gg \to H \to \gamma\gamma$ cross section would be defined as

$$(\sigma \cdot \mathcal{B})(gg \to H \to \gamma\gamma) = \sigma_{SM}(gg \to H) \cdot \mathcal{B}_{SM}(H \to \gamma\gamma) \frac{\kappa_g^2 \cdot \kappa_\gamma^2}{\kappa_H^2}. \tag{2.25}$$

Several benchmark parametrizations can be defined based on this framework. The simplest way is to define one common coupling scale factor to search for deviations from the SM Higgs couplings structure, which effectively reduces to an overall signal strength measurement $\mu = \frac{\kappa^2 \cdot \kappa^2}{\kappa^2} = \kappa^2$. However, this parametrization does not take into account the different roles of the Higgs boson in the mass generation of the vector bosons and fermions. By defining two parameters, one scaling the couplings to vector bosons, $\kappa_V (= \kappa_W = \kappa_Z)$ and one scaling the couplings to fermions, $\kappa_F (= \kappa_t = \kappa_b = \kappa_\tau)$, the role of the Higgs boson in electroweak symmetry breaking and therefore the compatibility with the SM Higgs boson can be probed. Several other parametrizations as well as a more detailed description of the kappa framework can be found in Ref. [41].

Properties of the Standard Model Higgs Boson

The discovery of the SM Higgs boson by the ATLAS and CMS collaborations in 2012 was reported with significances of $5.9\,\sigma$ and $5.8\,\sigma$, respectively [1–3]. This

milestone shifted the experimental effort in the Higgs sector from a search to the measurement of the Higgs boson properties and couplings.

One important property, the mass of the Higgs boson, has been determined by a combination of results from ATLAS and CMS in the $H \to \gamma\gamma$ and $H \to ZZ$ decay channels [42], resulting in

$$M_{\mathrm{H}} = 125.09 \pm 0.21 \,(\mathrm{stat}) \pm 0.11 \,(\mathrm{syst})\,\mathrm{GeV}.$$

The ratio of the combined signal yield with respect to the SM prediction is measured to be $\mu = 1.09 \pm 0.11$, indicating compatibility with the SM prediction [40]. The total decay width of the 125 GeV Higgs boson is predicted to be $\Gamma_{\mathrm{H}} = 4\,\mathrm{MeV}$, which is out of reach for a direct measurement at the LHC due to the detector resolution. However, by combining the on- and off-shell contributions in $H \to ZZ$ decays the width is constrained to be $\Gamma_{\mathrm{H}} < 3.2^{+2.8}_{-2.2}\,\mathrm{MeV}$ [43].

The observation in the $H \to \gamma\gamma$ decay channel rules out the possibility of a Higgs boson with spin 1, since this would violate the Landau-Yang theorem [44, 45]. The data is consistent with the pure scalar hypothesis, under the assumption of spin 0, while disfavouring the pure pseudoscalar hypothesis and the spin 2 hypothesis. The measurement of the spin-parity of the Higgs boson, $J^P = 0^+$, therefore strongly favours the SM expectation [39, 46].

The Higgs boson couplings to the W and Z boson as well as the photon have been observed [35, 38, 47, 48] and also the couplings to 3rd generation fermions, b quarks, t quarks and τ leptons, have been established [9, 10, 49–52] by both collaborations. All of these measurements are consistent with the SM expectation. Direct searches for the $H \to \mu\mu$ and $H \to c\bar{c}$ decay modes have been performed [53–56], however, more data is required to establish these final states. No measurements for the couplings to the remaining 1st and 2nd generation fermions have been conducted so far.

2.2 Beyond the Standard Model

The standard model is the crown jewel of 20th-century particle physics and has so far been able to describe a huge amount of experimental data with exceptional accuracy. Nevertheless, it has several shortcomings. Some are aesthetic considerations and others are deficits to describe the experimental results. For instance, it does not incorporate gravity, provides no candidate for dark matter [57] and does

not explain neutrino oscillations [58, 59].

An aesthetic shortcoming is the so-called hierarchy problem. The physics of elementary particles and their interaction is well described by the SM at an energy scale of $\mathcal{O}(100)\,\mathrm{GeV}$. However, the SM must break down at an energy scale Λ, which is at most of the order of the Planck scale $M_\mathrm{Pl} \simeq 10^{19}\,\mathrm{GeV}$, at which quantum gravitational effects become important. For the Higgs sector of the SM the hierarchy problem refers to the large quantum corrections to the Higgs boson mass due to virtual fermion and gauge boson loops [60]:

$$M_{\mathrm{H_{SM}}}^2 = (M_\mathrm{H})_0^2 + \underbrace{\Delta M_\mathrm{H}^2}_{\mathcal{O}(\Lambda^2)}. \tag{2.26}$$

The observation of $M_{\mathrm{H_{SM}}} = 125\,\mathrm{GeV}$, under the assumption that there is no new physics up to the Planck scale, would only be compatible with $\Lambda \sim \mathcal{O}(10^{19}\,\mathrm{GeV})$ if the bare Higgs boson mass $(M_\mathrm{H})_0$ is extremely fine tuned.

The hierarchy problem is addressed by many beyond the standard model (BSM) theories and one of the most popular is supersymmetry (SUSY) [8]. This theory introduces a new symmetry, which is invariant under transformations between bosonic and fermionic states, and postulates that every SM fermion field has a bosonic superpartner and each SM boson field has a fermionic superpartner. The introduction of a symmetry between fermions and bosons leads to the cancellation of the ΔM_H^2 term, since the fermion and boson loops contribute to the correction of the Higgs boson mass with opposite sign. For an unbroken symmetry, where all SUSY particles have the same mass as their SM counterparts, the cancellation would be exact. However, since we have not observed any of the SUSY particles, which would have been inevitable for an unbroken symmetry, this implies that the symmetry must be broken and the cancellation is not complete. Apart from solving the hierarchy problem SUSY also provides a dark matter candidate if the lightest SUSY particle is stable.

2.2.1 The Higgs Sector of the MSSM

One example for a supersymmetric extension of the SM is the minimal supersymmetric extension of the standard model (MSSM) [8], which is the main motivation for the results presented in Chapter 6. For this purpose, a brief introduction to the Higgs sector of the MSSM is given, which is based on the detailed description given in Ref. [60].

The simplest extension of the Higgs sector in the MSSM requires two Higgs doublets to generate mass for both up- and down-type fermions. The two doublets can be written as

$$\Phi_u = \begin{pmatrix} \phi_u^+ \\ \phi_u^0 \end{pmatrix} \text{ with } Y = +1, \qquad \Phi_d = \begin{pmatrix} \phi_d^0 \\ \phi_d^- \end{pmatrix} \text{ with } Y = -1. \quad (2.27)$$

Similar to the electroweak symmetry breaking described in Section 2.1.2 the phases of the Higgs field can be chosen such that the vacuum expectation values are real and positive

$$\langle \Phi_u \rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v_u \end{pmatrix}, \qquad \langle \Phi_d \rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} v_d \\ 0 \end{pmatrix}. \quad (2.28)$$

Spontaneous symmetry breaking leads to five physical Higgs bosons: a charged Higgs boson pair $H^\pm$, the light and heavy neutral scalars h and H ($m_h < m_H$), and the pseudoscalar A. All parameters of the MSSM Higgs sector at tree-level can be determined by two free parameters, which are chosen to be the ratio of the vacuum expectation values from Eq. (2.28)

$$\tan \beta \equiv \frac{v_u}{v_d}, \quad (2.29)$$

and the mass of the pseudoscalar m_A. By diagonalizing the squared-mass matrix

$$\mathcal{M}_{\text{tree}}^2 = \begin{pmatrix} m_A^2 \sin^2 \beta + m_Z^2 \cos^2 \beta & -\left(m_A^2 + m_Z^2\right) \sin \beta \cos \beta \\ -\left(m_A^2 + m_Z^2\right) \sin \beta \cos \beta & m_A^2 \cos^2 \beta + m_Z^2 \sin^2 \beta \end{pmatrix}, \quad (2.30)$$

one obtains the mixing angle α. The five physical Higgs bosons arise from the mixing of the fields and are defined as:

$$H^\pm = \phi_d^\pm \sin \beta + \phi_u^\pm \cos \beta,$$

$$A = \sqrt{2} \left(\text{Im}(\phi_d^0) \sin \beta + \text{Im}(\phi_u^0) \cos \beta \right),$$

$$h = -\sqrt{2} \left[\left(\text{Re}(\phi_d^0) - \frac{v_d}{\sqrt{2}} \right) \sin \alpha - \left(\text{Re}(\phi_u^0) - \frac{v_u}{\sqrt{2}} \right) \cos \alpha \right], \quad (2.31)$$

$$H = \sqrt{2} \left[\left(\text{Re}(\phi_d^0) - \frac{v_d}{\sqrt{2}} \right) \cos \alpha + \left(\text{Re}(\phi_u^0) - \frac{v_u}{\sqrt{2}} \right) \sin \alpha \right],$$

By choosing $\tan \beta$ and m_A as the free parameters the mass of the charged Higgs bosons can conveniently be written as

$$m_{H^\pm}^2 = m_A^2 + m_W^2, \quad (2.32)$$

and the masses of the scalar Higgs bosons at tree level, which are the eigenvalues of the matrix given in Eq. (2.30) are given by

$$m_{\mathrm{H,h}}^2 = \frac{1}{2}\left(m_{\mathrm{A}}^2 + m_{\mathrm{Z}}^2 \pm \sqrt{\left(m_{\mathrm{A}}^2 + m_{\mathrm{Z}}^2\right)^2 - 4m_{\mathrm{A}}^2 m_{\mathrm{Z}}^2 \cos^2 2\beta}\right). \qquad (2.33)$$

An important consequence of Eq. (2.33) is that there is an upper bound to the mass of the light scalar h:

$$m_{\mathrm{h}} \leq m_{\mathrm{Z}} \cdot |\cos(2\beta)| \leq m_{\mathrm{Z}}, \qquad (2.34)$$

where m_{Z} is the mass of the Z boson. Furthermore, $(\beta - \alpha)$ is constrained by

$$\cos^2(\beta - \alpha) = \frac{m_{\mathrm{h}}^2 \left(m_{\mathrm{Z}}^2 - m_{\mathrm{h}}^2\right)}{m_{\mathrm{A}}^2 \left(m_{\mathrm{H}}^2 - m_{\mathrm{h}}^2\right)}. \qquad (2.35)$$

The upper bound on m_{h} seems to be in contradiction with the Higgs state found in 2012, which has a mass of 125 GeV. However, the loop corrections to the tree level masses and couplings of the Higgs bosons in the MSSM can be significant and can increase m_{h} up to a maximum of 135 GeV [60]. The dominant one-loop contribution to the light-scalar mass arises from loops of t quarks and their scalar superpartners. Such higher order corrections depend on additional parameters. For this purpose, searches targeting the Higgs sector use a set of predefined values, referred to as *benchmark scenario*, for these parameters and keep $\tan \beta$ and m_{A} as free parameters. A selection of different benchmark scenarios can be found in Refs. [61–63]. The model-dependent interpretation of the results presented in Chapter 6 is carried out in the $m_{\mathrm{h}}^{\mathrm{mod}+}$ and hMSSM benchmark scenarios. In both scenarios the observed Higgs boson at 125 GeV is interpreted as the light scalar h. In the $m_{\mathrm{h}}^{\mathrm{mod}+}$ scenario a large fraction of the phase space is compatible with the observed Higgs boson at 125 GeV within a variation of ± 3 GeV. This variation corresponds to the theoretical uncertainty on the MSSM prediction of the h boson mass. In the hMSSM scenario the Higgs boson found at 125 GeV is interpreted as the light scalar h and the mass measurement is used to estimate the main corrections and therefore the masses of the remaining MSSM Higgs bosons.

The couplings of the neutral scalars h and H are modified with respect to the SM couplings by the factors given in Table 2.4. In the *decoupling limit*, $m_{\mathrm{A}} \gg m_{\mathrm{Z}}$, $\cos(\beta - \alpha)$ tends to zero, which simplifies the mixing angle in the scalar sector to $\alpha \approx \beta - \pi/2$. In this limit, which is approached at values $m_{\mathrm{A}} \gtrsim 300$ GeV, the couplings of the light scalar h to gauge bosons and fermions become SM-like. Furthermore, the couplings of the heavy scalar H to down-type fermions are enhanced

	h	H	A
Gauge bosons	$\sin(\beta - \alpha)$	$\cos(\beta - \alpha)$	0
up-type fermion	$\frac{\cos\alpha}{\sin\beta}$	$\frac{\sin\alpha}{\sin\beta}$	$\frac{1}{\tan\beta}$
down-type fermion	$-\frac{\sin\alpha}{\cos\beta}$	$\frac{\cos\alpha}{\cos\beta}$	$\tan\beta$

Table 2.4 − Coupling modifiers of the three neutral MSSM Higgs bosons to vector bosons, up-type and down-type fermions at tree level. The factors are given with respect to the SM couplings [11].

by a factor $\tan\beta$, the couplings to up-type fermions are suppressed by a factor $\frac{1}{\tan\beta}$ and the couplings to gauge bosons vanish. The enhanced coupling to down-type fermions motivates the choice of the Higgs boson decay to a pair of τ leptons in the search presented in Chapter 6. This enhancement also has consequences for the Higgs boson production. At large values of $\tan\beta$ the kinematic properties of the Higgs boson produced via ggF are modified due to the enhanced contribution of b quarks in the fermion loop. Furthermore, the Higgs production in association with b quarks gains importance and dominates over the production via ggF. The tree-level Feynman diagrams for both processes are shown in Fig. 2.7.
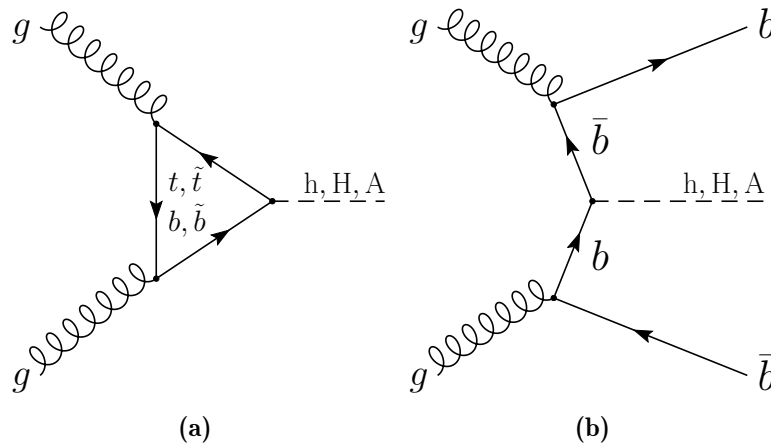
Figure 2.7 − Tree-level Feynman diagrams for the MSSM Higgs production via ggF (a) and in association with b quarks (b).

2.3 The Precision Quest on Higgs Couplings

One important question regarding the measurement of Higgs boson couplings is how well they need to be measured [64]. Assuming that the 125 GeV Higgs boson is a "pure" SM Higgs boson a rigorous answer is that the measurements should be at

least at the level of the theory uncertainties. However, since the SM is known to be incomplete as discussed in Section 2.1.1 the 125 GeV Higgs boson also needs to be embedded in a more comprehensive theory, potentially with deviating predictions for the Higgs sector. A more alluring motivation to measure the Higgs boson couplings is therefore to find such deviations. Consequently, the Higgs boson couplings need to be measured with high enough precision to be sensitive to these deviations. Table 2.5 shows the largest deviation from the SM expectation of three BSM physics scenarios: mixed-in hidden sector Higgs bosons [65, 66], composite Higgs bosons [67–69] and MSSM Higgs bosons [8]. The presented numbers are based on the assumption that the 125 GeV Higgs boson is the "SM-like" state and that there are no additional Higgs states observable at the LHC. The current

	$\Delta\kappa_{\mathrm{V}}$	$\Delta\kappa_{\tau}$
Mixed-in Singlet	6 %	6 %
Composite Higgs	8 %	tens of %
Minimal Supersymmetry	< 1 %	5 %–20 %
Current precision (2019)	≈ 10 %	≈ 15 %

Table 2.5 – Summary of the possible maximum deviations of the couplings to vector bosons and τ leptons for three different BSM physics scenarios. The last row represents the current precision achieved by experiments at the LHC. The values given for MSSM refer to the preferred $\tan\beta$ region but deviations can be of $\mathcal{O}(100\,\%)$ for κ_{τ}. Adapted from Refs. [64, 70]

precision on the Higgs bosons couplings to τ leptons and vector bosons is in the same ballpark as most of the expected deviations from the SM couplings. However, to be sensitive to the deviations in the percent level the precision on the Higgs couplings needs to be improved further.

CHAPTER 3

LHC AND THE CMS EXPERIMENT

"Everything not saved will be lost."

— NINTENDO "QUIT SCREEN" MESSAGE

In this chapter, an overview of the experimental setup utilized in this thesis will be given. Section 3.1 outlines the accelerator complex operated by the European Organisation for Nuclear Research (CERN) and the largest particle collider ever built, the LHC. In Section 3.2, the CMS experiment is introduced by outlining the major detector parts and their functionality.

3.1 The Large Hadron Collider

The LHC [71] is a superconducting hadron accelerator and collider installed in a 26.7 km tunnel between 45 m and 170 m below the surface close to Geneva (Switzer-
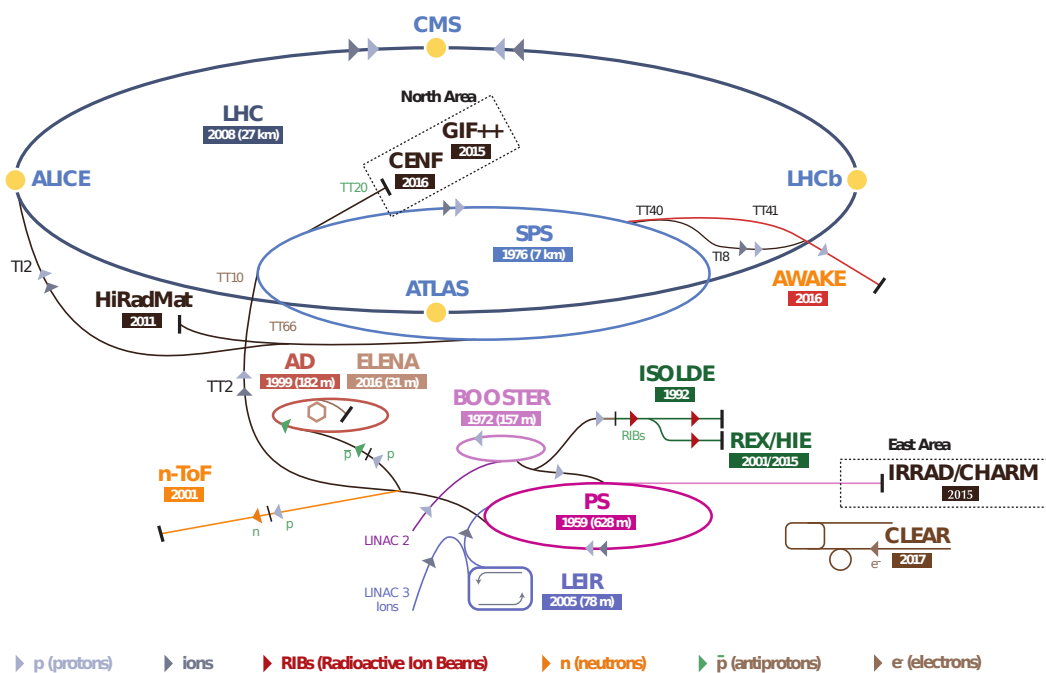
Figure 3.1 − Overview of the CERN accelerator complex [72].

land). Being a particle-particle collider, two rings with counter-rotating beams are needed, unlike particle-antiparticle beams that can share the same phase-space in a single ring. Due to the limited size of the tunnel that was initially designed for LEP the LHC adopted the twin-bore magnet design [73], shown in Fig. 3.2, instead of two separate rings. The LHC was designed to collide protons at a centre-of-mass
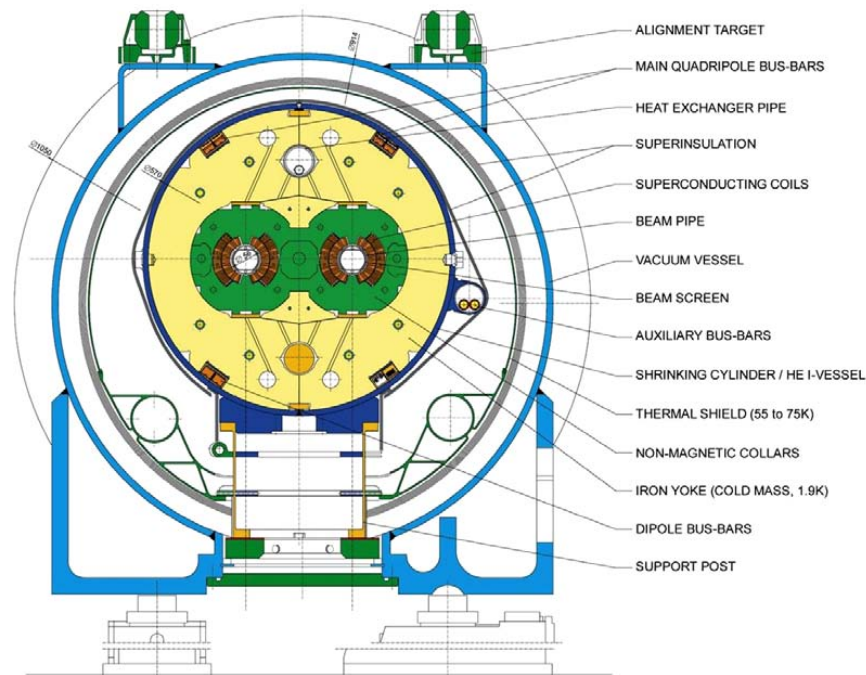
Figure 3.2 – Cross-section of a cryodipole [71].

energy of up to 14 TeV and is also used for proton-lead and lead-lead collisions. The following sections only describe proton-proton collisions as they correspond to the data used in the physics results in Chapters 5 and 6.

The LHC is the last element of a chain of particle accelerators shown in Fig. 3.1. Protons are produced using a Duoplasmatron, by ionizing hydrogen gas, and accelerated to about 50 MeV in the linear accelerator LINAC 2. In the next step, they are injected in the Proton Synchrotron Booster (PSB) before entering the Proton Synchrotron (PS) where they reach an energy of 25 GeV. The Super Proton Synchrotron (SPS) increases the proton energy to 450 GeV before injecting them in the LHC where they reach their final energy. The four initial acceleration processes take about 20 min and it takes another 20 min in the LHC to reach the final energy. In total, 2808 bunches circulate in the LHC when it is completely filled. These bunches cross at four different interaction points, shown in Fig. 3.3, where the four main experiments are located:
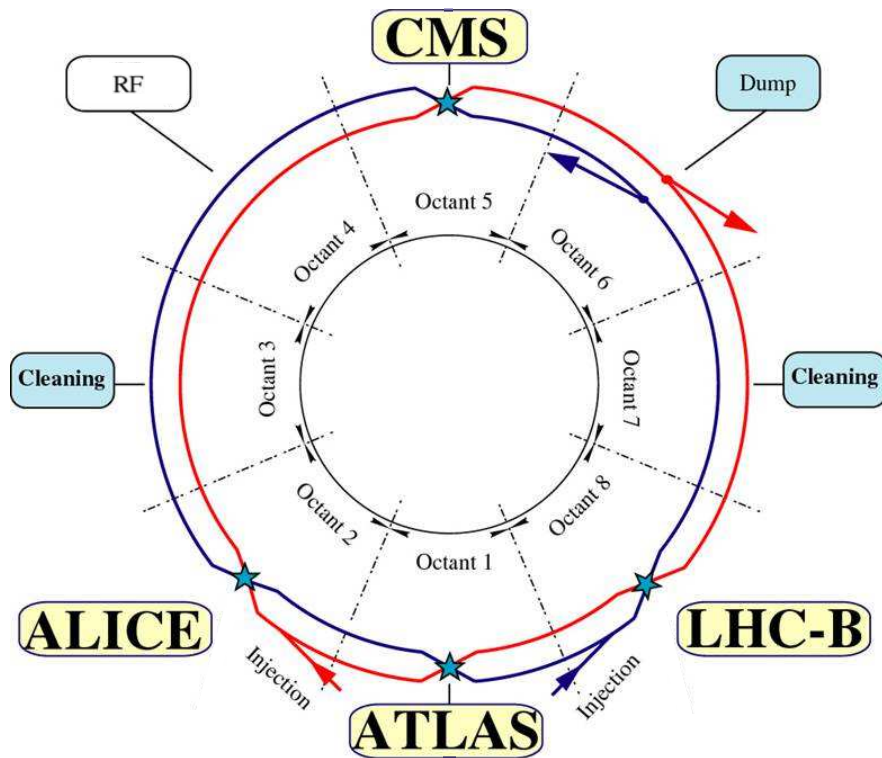
Figure 3.3 – Schematic layout of the LHC. Insertions at Octant 3 and 7 each contain two collimation systems. The insertion at Octant 4 contains one independent radio frequency (RF) cavity for each LHC beam. The insertion at Octant 6 houses the beam abort systems for Beam 1 and Beam 2. Adapted from Ref. [71].

ALICE – A Large Ion Collider Experiment [74].
Designed to address the physics of strongly interacting matter and the quark-gluon plasma at extreme values of energy density and temperature in nucleus-nucleus collisions.

ATLAS – A Toroidal LHC Apparatus [75].
A multi-purpose detector covering a wide physics program.

CMS – Compact Muon Solenoid [76].
A multi-purpose detector like ATLAS but with a different design philosophy outlined in more detail in Section 3.2.

LHC-B – Large Hadron Collider beauty [77].
Performs precision measurements of charge-parity (CP) violation and rare B hadron decays.

3.2 The Compact Muon Solenoid Experiment

The CMS detector is a multi-purpose apparatus located close to the French village of Cessy at Point 5 along the LHC. It is designed to elucidate the nature of electroweak symmetry breaking, probe physics at the TeV scale and search for evidence of physics beyond the SM. The central feature of CMS is a superconducting solenoid, surrounding the tracker and both the Electromagnetic Calorimeter (ECAL) and the Hadron Calorimeter (HCAL). The solenoid, which produces a magnetic field of $3.8\,\mathrm{T}$, is encapsulated by the iron return yoke interspersed with muon chambers. Figure 3.4 shows an exploded view of the key elements of the CMS detector. In the following, a brief overview of the different detector systems is given. A detailed description can be found in Ref. [76]. The requirements for CMS to meet the goals of the physics program can be summarized as follows:
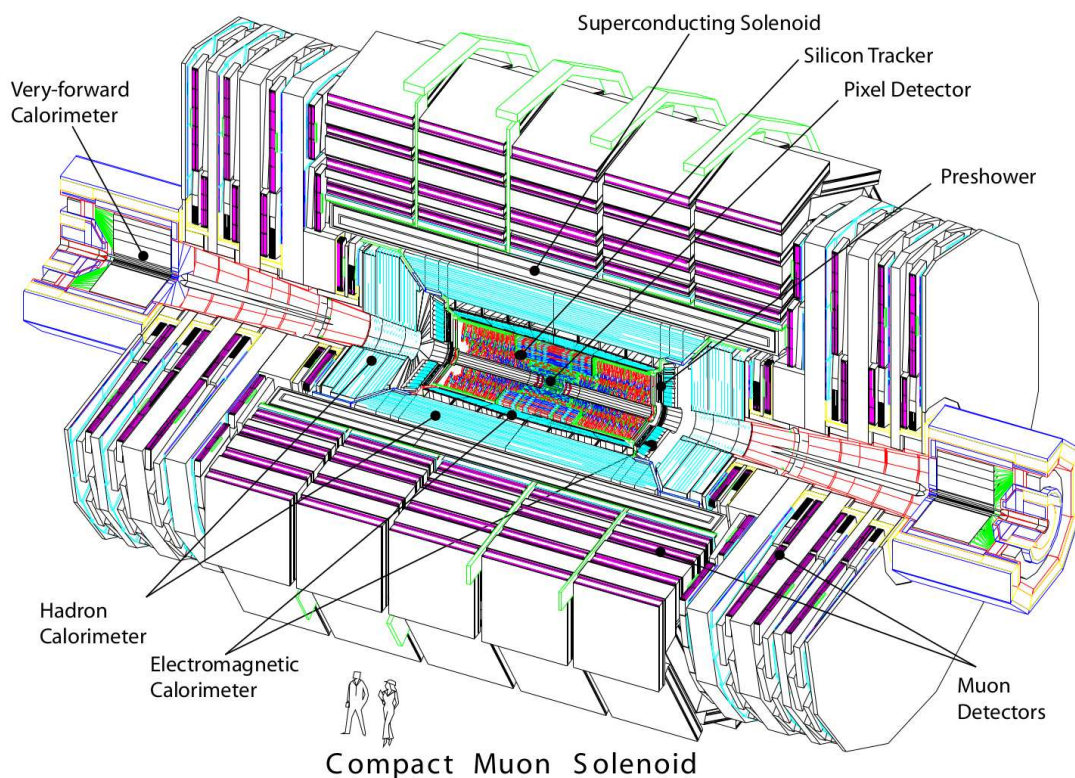
Figure 3.4 – Exploded view of the CMS detector showing the typical onion-like structure of the particle detectors. Located closest to the interaction point is the tracking system, which is surrounded by the electromagnetic and hadron calorimeters. All three systems are installed inside the superconducting solenoid. The muon detectors are the final layer of CMS [76].

- Over a wide range of momenta and angles a good muon identification and momentum resolution is required. The dimuon mass resolution should be about 1 % at 100 GeV, and the charge of muons needs to be determined unambiguously for momenta $p < 1$ TeV.

- The inner tracker is required to have good charged particle momentum resolution and reconstruction efficiency. For efficient triggering and offline tagging of τ leptons and b-jets, pixel detectors close to the interaction point are needed.

- The ECAL needs good electromagnetic energy resolution as well as good diphoton and dielectron mass resolution of about 1 % at 100 GeV. Furthermore, a wide geometric coverage, π^0 rejection, and efficient photon and lepton isolation is required.

- The HCAL is required to have large hermetic geometric coverage with fine lateral segmentation to ensure good missing transverse energy and dijet-mass resolution.

The origin of the coordinate system adopted by CMS is centered at the nominal collision point with the y-axis pointing vertically upward, and the x-axis radially inward. The z-axis points in the direction of the beam running counterclockwise. The polar angle θ is measured from the z-axis, while the azimuthal angle ϕ is measured starting from the x-axis in the x-y plane. Commonly the pseudorapidity defined as

$$\eta = -\ln\left(\tan\frac{\theta}{2}\right) \tag{3.1}$$

is used instead of the polar angle since for $E \gg m$ the difference of the pseudorapidities of two particles is invariant under a Lorentz boost in the z-direction.

The distance ΔR between two objects in the $\eta \times \phi$ plane is defined as

$$\Delta R = \sqrt{\Delta\eta + \Delta\phi}. \tag{3.2}$$

3.2.1 Inner Tracking System

The tracker [76] is the subdetector closest to the interaction point with the purpose of measuring the trajectories of charged particles as well as secondary vertices of long-lived particle decays. At design luminosity, $\mathcal{O}(1000)$ charged particles travel through the tracker every 25 ns and it is therefore required to have high granularity and fast response time. Furthermore, the components need to withstand the high
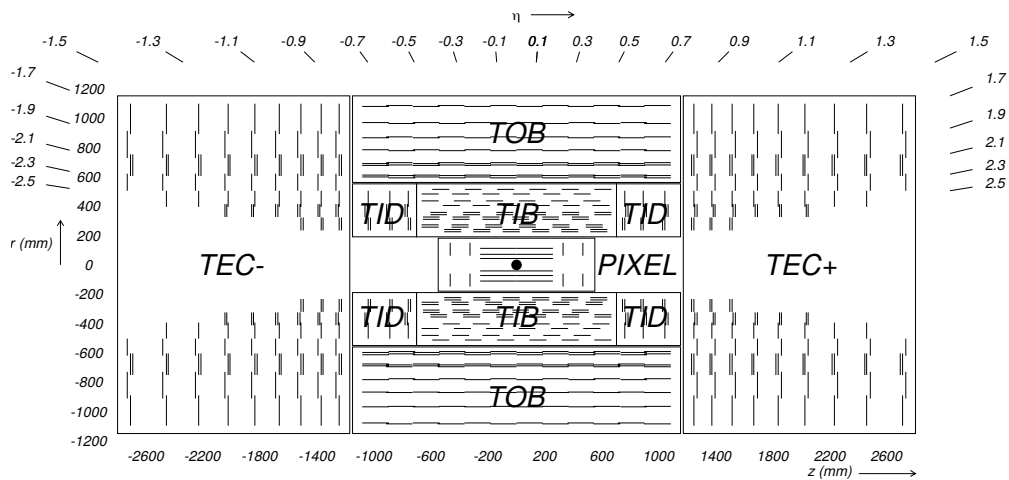
Figure 3.5 – Cross section of the CMS tracker showing the position of the pixel detector, the barrel detection layers (TIB and TOB), and the endcap layers (TID and TEC). Each line represents a detector module. Double lines indicate back-to-back modules [76].

radition due to the intense particle flux. The tracker is therefore solely based on silicon detector technology to satisfy the requirements on granularity, speed and radiation hardness.

The tracker is composed of a pixel and a strip tracker with cylindrical barrel layers, completed by endcaps which extend the acceptance up to a pseudorapidity of $|\eta| < 2.5$. Figure 3.5 shows a cross section of the CMS tracking systems. The innermost part is a pixel detector with three barrel layers at radii between $4.4\,\mathrm{cm}$ and $10.2\,\mathrm{cm}$ and two disks in the endcap. It was upgraded during the 2016 - 2017 extended year-end technical stop with a fourth barrel layer and a third disk in the endcap. The pixel detector is surrounded by the silicon strip tracker with 10 barrel detection layers (TIB and TOB) extending outwards to a radius of $110\,\mathrm{cm}$ and 3 plus 9 disks (TID and TEC) in the endcap.

3.2.2 Electromagnetic Calorimeter

The Electromagnetic Calorimeter (ECAL) [76] is a homogeneous calorimeter made of $75\,848$ scintillating lead tungstate ($PbWO_4$) crystals of which $61\,200$ are mounted in the central barrel part and the remaining crystals in the two endcaps. The radiation hardness and short radiation length of $0.89\,\mathrm{cm}$ of lead tungstate makes it an ideal choice to built a compact and fine granular calorimeter. The crystals emit about 80% of the light in $25\,\mathrm{ns}$ which is of the same order of magnitude as the LHC bunch crossing time.
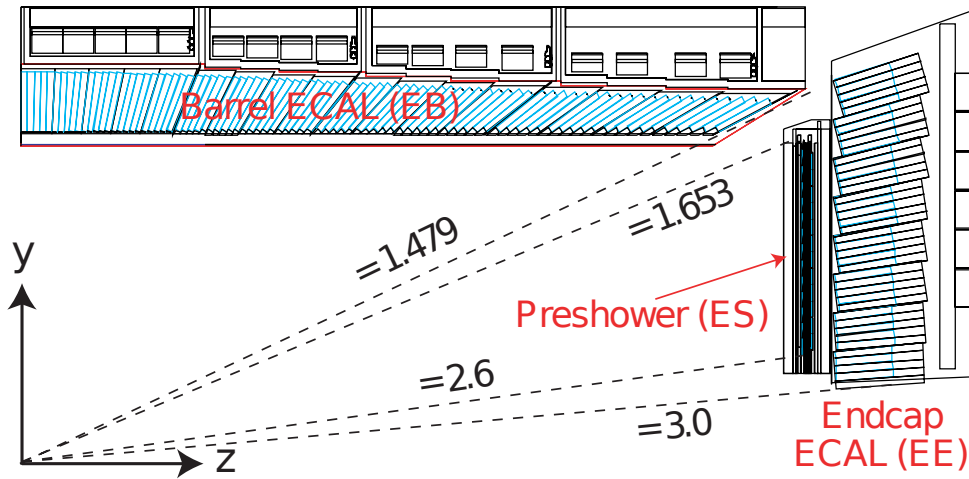
Figure 3.6 – Transverse section through the ECAL showing the location of the ECAL barrel (EB), the Preshower detector and the ECAL endcap (EE) [78].

The ECAL Barrel (EB) covers the range $|\eta| < 1.479$ and the ECAL Endcap (EE) extends the range to $1.479 < |\eta| < 3.0$, as can be seen in Fig. 3.6. The cross-section of a barrel crystal is $22 \times 22\,\mathrm{mm}^2$ at the front face, with $22\,\mathrm{mm}$ also being the Molière radius of $PbWO_4$, and a length of $230\,\mathrm{mm}$ corresponding to 25.8 radiation lengths X_0. Similar lead tungstate crystals, grouped in mechanical units of 5×5 crystals, are used in the EE with a cross section of $28.62 \times 28.62\,\mathrm{mm}^2$ at the front face and a length of $220\,\mathrm{mm}$ (24.7 X_0). This results in a compact calorimeter with fine granularity allowing for precise energy measurement of electrons, photons and neutral pions. A preshower detector is mounted in front of the endcaps in the region $1.653 < |\eta| < 2.6$ to measure the position of electromagnetic showers precisely, which allows an improved differentiation between electrons and photons from neutral pion decays.

3.2.3 Hadron Calorimeter

The Hadron Calorimeter (HCAL) [76] is a sampling calorimeter, providing energy measurements for the reconstruction of neutral hadrons and for apparent missing transverse energy due to neutrinos or possible exotic particles. It consists of brass absorber layers, chosen due to its short interaction length and since it is non-magnetic, interspersed with plastic scintillator tiles. Figure 3.7 shows the four different parts of the HCAL which are explained in the following.
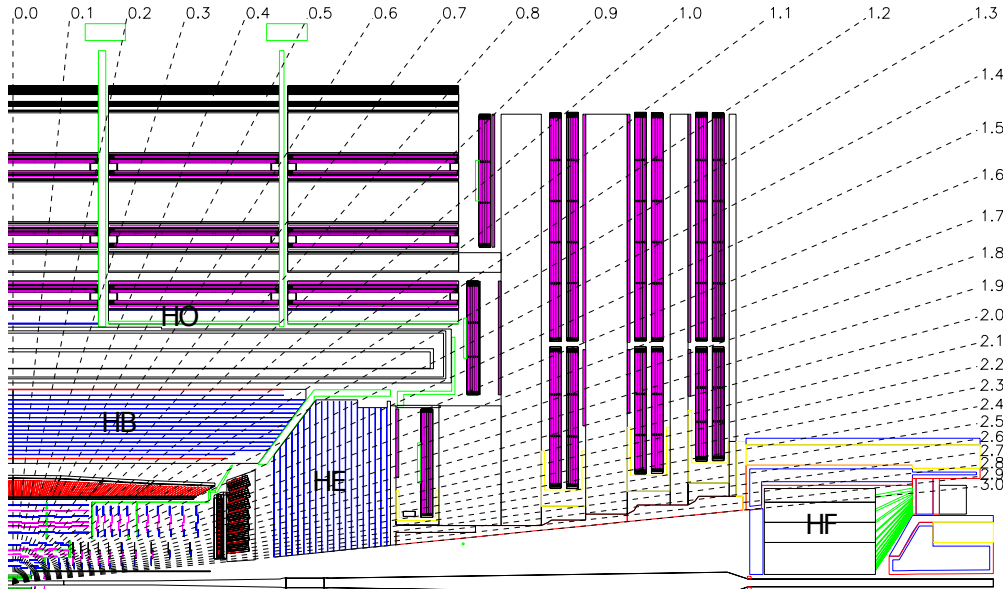
Figure 3.7 – Transverse section through the CMS detector showing the locations of the hadron barrel (HB), endcap (HE), outer (HO) and forward (HF) calorimeters [76].

The Hadron Barrel (HB) is radially restricted by the outer extent of the ECAL and the inner extent of the magnetic coil. It covers a pseudorapidity range of $|\eta| < 1.3$ and has a spatial resolution of $(\Delta\eta, \Delta\phi) = (0.087, 0.087)$. The total absorber thickness at $\eta = 0$ is 5.82 interaction lengths (λ_l) and increases up to 10.6 λ_l at $|\eta| < 1.3$, with the ECAL adding about 1.1 λ_l of material in front of the HB. Since not all hadrons can be stopped in the limited range of the HB, a small part, called Hadron Outer (HO) calorimeter, is located outside the solenoid. It covers a range of $|\eta| < 1.26$ and serves as tail catcher of showers penetrating the magnetic coil. The HO uses the solenoid as absorber material and increases the total calorimeter depth to a minimum of 11.8 λ_l. The Hadron Endcap (HE) covers a substantial portion of the rapidity range, $1.3 < |\eta| < 3$, a region containing about 34 % of the particles produced, and has a spatial resolution of $(\Delta\eta, \Delta\phi) = (0.17, 0.17)$. At $|\eta| \simeq 3$ the HE is required to have high radiation tolerance due to the high luminosity and counting rates of the LHC. The Hadron Forward (HF) calorimeter, at $2.8 < |\eta| < 5.2$, covers even higher rapidities, and therefore particle fluxes, and is located at 11.2 m from the interaction point.

3.2.4 Muon System

The detection of muons is of central importance to the CMS experiment as implied by its middle name. They are a powerful tool for recognizing signatures of interesting SM and BSM processes. The muon system [76] uses three types of
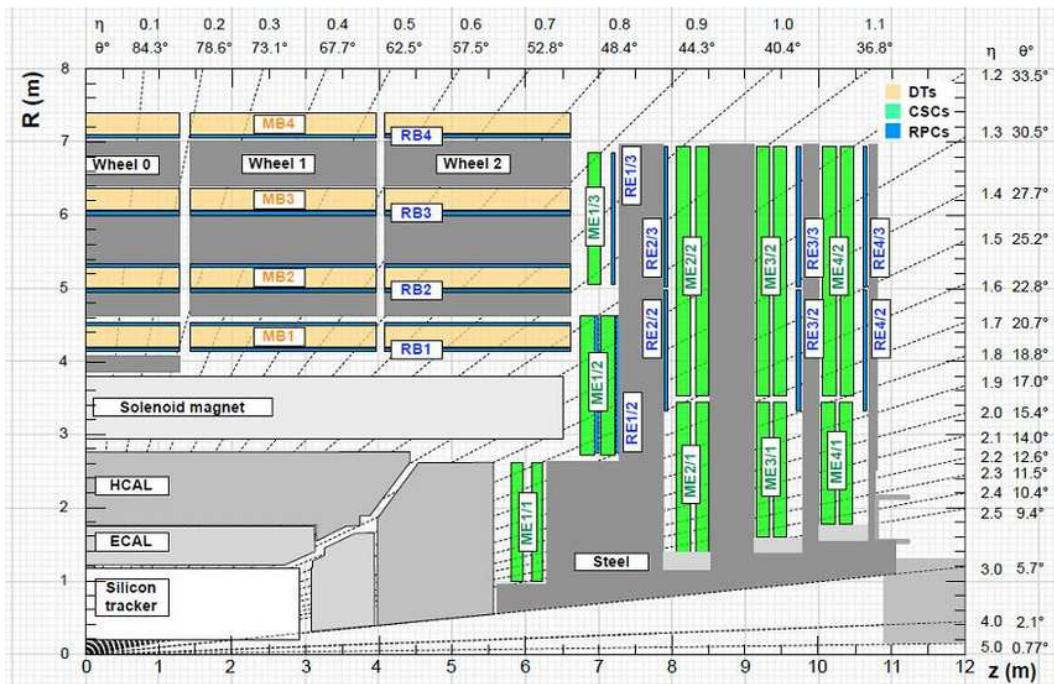
Figure 3.8 – Longitudinal layout of one quadrant of the CMS detector showing the position of the DT, CSC and RPC layers in the muon chambers [79].

gaseous particle detectors chosen due to the large surface to be covered and the different radiation regions. It is located outside of the solenoid interspersed in the iron return yoke. Drift Tubes (DT) are installed in the barrel region covering a rapidity range of $|\eta| < 1.2$. In the endcap region in the η range $0.9 < |\eta| < 2.4$, where muon rates and background levels are high, Cathode Strip Chambers (CSC) are used. These have a fine segmentation, a fast response and are radiation hard. In both the barrel and the endcap region Resistive Plate Chambers (RPC) are installed for $|\eta| < 1.6$ to complement the excellent spacial resolution of the DTs and CSCs with a high precision timing measurement.

Figure 3.8 illustrates the cross-section of one quadrant of the muon system, showing the three muon detection systems. The fourth muon endcap station on the top providing additional layers of CSC and RPC was added during long shutdown 1 (LS1).

3.2.5 Trigger

With a bunch spacing of 25 ns during proton-proton collisions at the LHC the crossing frequency is 40 MHz. Due to the high crossing rate, and the fact that typically tens of collisions occur at each bunch crossing simultaneously, it is impos-

sible to transfer, store and process every event. Therefore, a drastic rate reduction has to be achieved which is accomplished with the trigger system. The trigger system is used to selectively enhance events of physical interest, which reduces the rate to $\mathcal{O}(1\,\text{kHz})$. The rate reduction is performed in two stages: the Level-1 (L1) Trigger and the High-Level Trigger (HLT).

The L1 Trigger uses coarse information from the calorimeters and the muon system, and has a design output rate of 100 kHz. It was replaced by an upgraded system during long shutdown 1 and the year-end technical stop in 2015 [80], providing amongst other things improved jet finding, object isolation, muon transverse momentum resolution and hadronic τ lepton identification.

The L1 Trigger is split into two paths, the calorimeter trigger and the muon trigger. The calorimeter trigger is based on a time-multiplexed design feeding energy deposits from the ECAL and HCAL to the first layer of the trigger. This layer maps data received from many bunch crossings onto slices of the detector. The data is re-mapped from the first to the second trigger layer so that the second layer receives data from the entire calorimeter for a single bunch crossing. In this step a sorted list of the best candidates is passed to the global trigger, which is based on a basic object identification from the energy deposits.

Hits in the muon chambers from all three muon detection systems, DT, CSC, and RPC, are passed to the track-finder layer and processed regionally in sectors of $\eta-\phi$. The information from different regions is combined in further layers of the muon trigger which is further passed to the global muon trigger. The global muon trigger returns a sorted list of the best muon candidates to the global trigger. The global trigger combines in the last step the information from the calorimeter and muon triggers and decides if the event is kept for further processing in the HLT.

The HLT is able to perform more complex calculations on events passing the L1 Trigger since it has access to the complete detector read-out data. The events passing the HLT are kept for storage, which ultimately reduces the event rate down to the target rate of $\mathcal{O}(1\,\text{kHz})$.

CHAPTER 4

EVENT SIMULATION AND RECONSTRUCTION

> "*When you have eliminated all which is impossible, then whatever remains, however improbable, must be the truth.*"
>
> – SHERLOCK HOLMES

The analyses in the following chapters rely on predictions based on Monte Carlo (MC) event generation which will be described in Section 4.1. This is followed by an overview of the algorithms used to reconstruct collision events with the CMS detector. First, the reconstruction of charged tracks is described in Section 4.2.1. Sections 4.2.2 and 4.2.3 focus on the reconstruction of electrons and muons from charged tracks and deposits in the ECAL and muon chambers. Finally, the reconstruction of jets, missing transverse momentum and τ leptons, based on Particle-Flow (PF) candidates, is described in Sections 4.2.4–4.2.6.

4.1 Monte Carlo Simulation

The main process of interest in proton-proton collisions at the LHC is a hard-scattering of partons from two colliding protons. Factorization [81] allows the event simulation to be separated into several steps. The following is based on Ref. [82].

Hard-scattering process The hard scattering of the incoming partons leads to a relatively small number of primary outgoing particles. Incoming parton momenta are sampled from parton density functions, which give the probability for a given parton to carry a certain fraction of the proton momentum. The hard scattering is computed with the matrix element formalism within the framework of perturbative QCD usually performed at leading order (LO) or next-to-LO (NLO) depending on the used event generator.

Underlying event   Interactions between partons of the colliding protons that are not directly associated with the hard-scattering are called underlying event. These typically give rise to soft QCD interactions of which the most common is the interaction of two partons resulting in a single outgoing parton.

Parton showering   The radiation of soft gluons from colour charged particles in the initial and final states is described by parton shower models. Each of the gluons can also radiate further partons. This evolution of the parton shower continues until the partons reach a minimum energy threshold.

Hadronization   If the energy of partons is below about $1\,\mathrm{GeV}$ perturbative QCD is not applicable and they cannot be treated as free anymore. At this stage partons undergo hadronization where colour-neutral hadrons are formed out of colour-charged quarks and gluons.

Pile-up   There are several proton-proton collisions per bunch crossing of which one is the hard process of interest. Most of these collisions are however soft-scattering events, called pile-up (PU). For this purpose, soft proton-proton collisions need to be added to the simulation of the main hard-scattering process. The simulated events are adjusted after simulation to match the PU distribution in data.

Detector Simulation and Digitization   The final step is to simulate how the generated particles interact with the CMS detector. For the detector simulation a description of the detector is needed, which is done with the GEANT4 toolkit [83]. The detector definition requires the representation of its elements and their spatial position as well as their material and electronic properties. The detector response, which is the physical signal in the respective detector parts, is transformed via electronics into a digital signal and the simulation of this step is called digitization.

4.2 Object Reconstruction

Figure 4.1 shows the distribution of all reconstructed particles in a simulated $t\bar{t}$ event. This makes clear that impressive use and understanding of the detector is needed in order to precisely measure such complex events. For this purpose, CMS makes use of the Particle-Flow (PF) algorithm [85] to reconstruct and identify all particles visible to the detector by combining the information from the different subdetectors. The interaction of the particles with each subdetector is
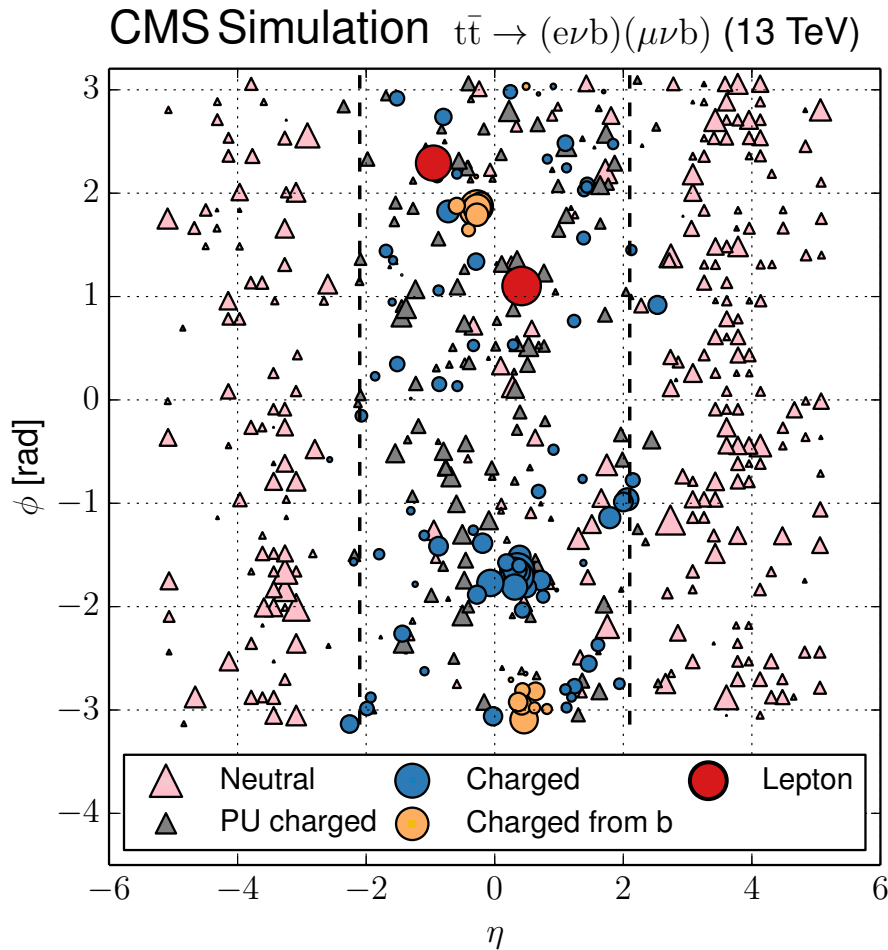
Figure 4.1 − Distribution of all PF candidates reconstructed in a simulated t$\bar{\text{t}}$ event [84].

schematically shown in Fig. 4.2. Each particle is classified as either photon, electron, muon, neutral hadron or charged hadron, which are further combined to higher-level physics object such as jets, hadronic τ leptons and missing energy. The highly-segmented tracker, fine-grained ECAL, hermetic HCAL, strong magnetic field and excellent muon spectrometer of the CMS detector are well suited for PF reconstruction. The algorithm starts with charged particle tracks, tracks in the muon chamber and calorimeter energy clusters. At the core of the PF reconstruction is the *link algorithm* which is used to connect the signals in the individual subdetectors in blocks. Tracks are associated to calorimeter clusters if they intersect one of the calorimetric cells. Similarly are clusters in the ECAL and HCAL linked if the measurement of the subdetector with finer granularity lies within the envelope of the coarser measurement.
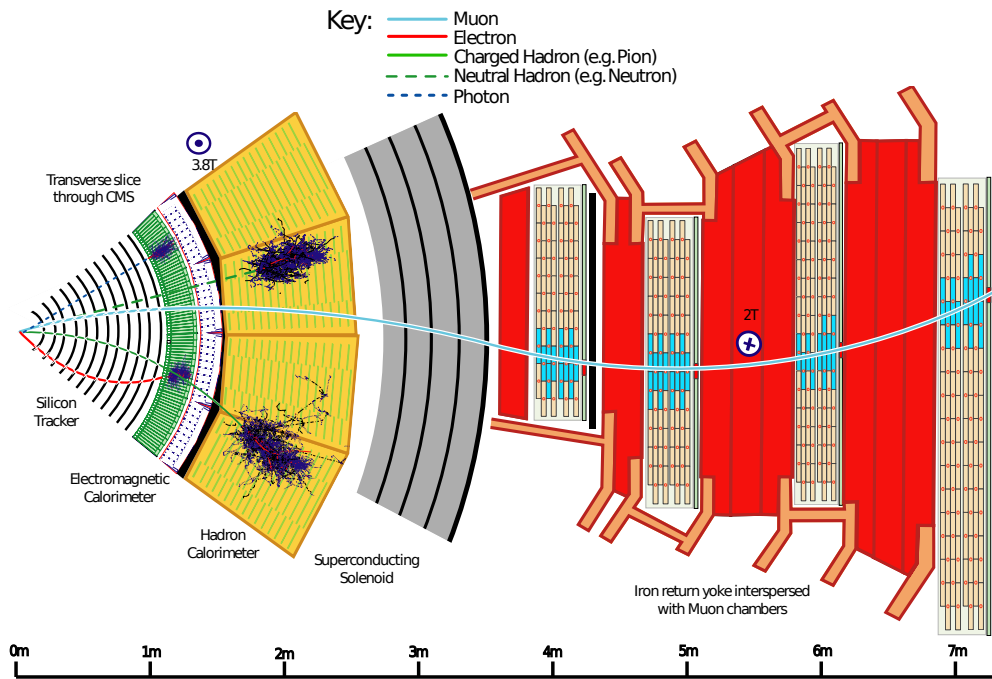
Figure 4.2 – Sketch of the specific particle interactions in a transverse slice of the CMS detector. The muon and the charged pion are positively charged, and the electron is negatively charged [85].

Blocks of linked objects are then interpreted as a particle of particular type. At first electrons and muons are selected and reconstructed. Their expected energy deposits in the calorimeters are subtracted from the relevant clusters. Each remaining track is interpreted as charged hadron that gets assigned the pion mass and as momentum a weighted average between the track and cluster measurements. In the case that an excess of calorimeter energy is found with respect to the associated track momentum, the residual energy is identified as neutral hadron or photon. This is also done with deposits in the calorimeters not associated to a track.

4.2.1 Tracks and Vertices

The *tracks*, and therefore momentum and position parameters, of charged particles are reconstructed from hits they leave in the inner tracker using the Combinatorial Track Finder (CTF) algorithm [86]. The collection of reconstructed tracks is produced in a process called *iterative tracking* in which the CTF algorithm is executed multiple times, each iteration passing the following four steps:

1. An initial estimate of the trajectories is given by track seeds, originating from hits found in two or three layers of the tracker.

2. The track finding, which is based on a Kalman Filter (KF) [87], extrapolates the seed trajectories along the expected flight path of a charged particle. Additional hits found along this trajectory are added to the track.

3. The KF is run again with the full set of hits that are found after the extrapolation has reached the final layer of the tracker. This provides the best set of parameters of each trajectory.

4. Tracks not satisfying certain quality criteria are rejected.

To reconstruct as many tracks as possible these steps are repeated six times. In each iteration the hits associated to already reconstructed tracks are removed and the settings of the algorithm are updated. Each iteration is configured to reconstruct tracks with a specific signature. At first prompt tracks, which are tracks originating near the interaction point, are reconstructed. The next iteration is used to recover prompt tracks with only two pixel hits. The third iteration is configured to find prompt tracks with low transverse momentum. The remaining iterations are used to reconstruct tracks outside the beam spot or to recover tracks not found in the previous steps.

After the track reconstruction the location of all proton-proton interaction vertices in each event are measured (primary-vertex (PV) reconstruction). For this purpose, tracks that are compatible with originating from the same vertex are clustered and fitted for the position of the vertex. Track clustering is performed using the Deterministic Annealing (DA) algorithm [88] which identifies and assigns tracks to the most probable vertex candidate. A candidate with at least two associated tracks is fitted with an *adaptive vertex fitter* [89] to compute the best estimate of the vertex parameters. The primary vertex with the largest scalar sum of transverse momenta of the associated tracks is chosen as the primary vertex of the hard scatter in the analyses shown in Chapters 5 and 6.

4.2.2 Electrons

Electrons are reconstructed from a cluster of energy in the ECAL that is associated to a track [90]. To maximize performance a stand-alone approach [91] complemented by the PF approach is used.

The electron energy is usually spread over several crystals and depending on the thickness of the traversed material up to $86\,\%$ is lost due to bremsstrahlung before reaching the ECAL. To recover and fully reconstruct the electron energy it is

necessary to capture the radiated photon energy which is mainly spread in the ϕ direction due to the electron trajectory in the magnetic field. For this purpose *supercluster* algorithms are used. The clustering is performed by defining the crystal containing the largest energy deposit above a certain threshold as the seed crystal. Slightly different algorithms are used in the ECAL barrel and the endcap due to different subdetector geometries. In the barrel arrays of 5×1 crystals in $\eta \times \phi$ are added in steps around the seed crystal, in a range of $\approx 0.3\,\mathrm{rad}$ in both ϕ directions. A supercluster is then formed by adding all arrays in which a minimum energy requirement is satisfied. This is done similarly in the endcaps by building superclusters from arrays of 5×5 crystals within a range of $\pm 0.7\,\mathrm{rad}$ in η and $\pm 0.3\,\mathrm{rad}$ in ϕ.

A dedicated tracking algorithm is used to take into account the change in curvature of the electron track due to the radiation of bremsstrahlung photons. Electron track reconstruction is done in two steps in CMS: seeding and tracking. The results of two complementary algorithms are used and combined in the seeding step. The tracker-based seeding uses tracks reconstructed with the general algorithm for charged particles extrapolated and matched to a supercluster in the ECAL. The ECAL-based seeding starts from a supercluster and extrapolates its energy and position to the collision vertex. In the tracking step the tracker seeds found with the algorithms above are extrapolated and smoothed using the Gaussian Sum Filter (GSF) [92]. This gives a better approximation for the non-Gaussian energy loss in each layer compared to the KF.

To separate genuine prompt electrons from photon conversion electrons, misidentified jets or electrons from b and c quark decays, a Multivariate Analysis (MVA) approach based on a Boosted Decision Tree (BDT) is used. It combines information on shower shape, track quality and kinematic quantities. To further reduce electrons coming from heavy flavour decays and misidentified jets, the electrons are required to be isolated from hadronic activity. For this purpose a *relative isolation* is introduced

$$I_{\mathrm{rel}}^{\mathrm{e}} = \frac{1}{p_{\mathrm{T}}^{\mathrm{e}}} \left[\sum p_{\mathrm{T}}^{\mathrm{charged}} + \max\left(0, \sum p_{\mathrm{T}}^{\mathrm{neutral}} + \sum p_{\mathrm{T}}^{\gamma} - p_{\mathrm{T}}^{\mathrm{PU}}\right) \right], \qquad (4.1)$$

with $p_{\mathrm{T}}^{\mathrm{e}}$ being the transverse momentum of the electron. The sums run over the transverse momenta of charged hadrons $p_{\mathrm{T}}^{\mathrm{charged}}$, neutral hadrons $p_{\mathrm{T}}^{\mathrm{neutral}}$ and photons p_{T}^{γ}. The estimate for the contribution from PU is given by $p_{\mathrm{T}}^{\mathrm{PU}}$. The isolation combines the transverse momentum sum of charged hadrons with the PU corrected sums of neutral hadrons and photons inside a cone $\Delta R = \sqrt{(\Delta\eta^2) + (\Delta\phi)^2}$ around the direction of the electron. The cone size that is used in the analyses

presented in Chapters 5 and 6 is $\Delta R = 0.3$.

4.2.3 Muons

Two approaches are used to reconstruct muons. They are based on the information of the tracker and the muon system and are called *global* and *tracker* reconstruction [93]. The global muon reconstruction is an *outside-in* approach which starts from tracks in the muon system and searches for matching tracks in the inner tracker. If a matching track is found the KF is reapplied to the combination of hits in the inner tracker and the muon system. It is especially efficient for muons that penetrate more than one muon station but has the downside of degraded resolution if the muon has lost energy while traversing the full detector. The tracker reconstruction on the other hand is an *inside-out* approach which starts from tracks in the inner tracker. These tracks are extrapolated to the muon system looking for a matching track stub in the DTs or CSCs. The extrapolation takes into account the magnetic field, average expected energy loss and scattering in the detector. It is more efficient for muons with $p_\mathrm{T} < 5\,\mathrm{GeV}$ since it only requires a single muon segment in the muon system. The efficiency to reconstruct a muon produced in proton-proton collisions within the geometrical acceptance of the muon system as either a global or tracker muon is approximately 99 %. Muons are called *global muons* if they are reconstructed by both the inside-out as well as the outside-in approach.

Similar as described in Section 4.2.2 for electrons an isolation requirement is placed on muons. The isolation variable I_rel^μ is similar to Eq. (4.1) but with the cone ΔR around the muon chosen to be 0.4 for the analyses presented in Chapters 5 and 6.

4.2.4 Jets

Quarks and gluons are produced abundantly in the QCD-dominated collisions at the LHC. However, due to colour confinement they cannot be observed directly and form jets as described in Section 2.1.1. To reconstruct the properties of the initial parton the constituents of these jets are combined and measured using a clustering algorithm [94]. This requires a definition of how the four-momentum of the jet should be assigned and how the single objects should be grouped involving some measure of distance. In CMS a class of algorithms known as sequential recombination is used, which are not affected by parton splitting or emission of soft collinear gluons (collinear- and infrared-safe). The algorithm defines the distances d_{ij} between two objects (particles or proto-jets) and d_{iB} between an

object and the beam as:

$$
d_{ij} = \min\left(k_{\mathrm{T}i}^{2p}, k_{\mathrm{T}j}^{2p}\right) \frac{\left(\eta_i - \eta_j\right)^2 + \left(\phi_i - \phi_j\right)^2}{R^2},
$$
$$
d_{iB} = k_{\mathrm{T}i}^{2p},
$$

(4.2)

with k_{T} being the the transverse momentum of a given object. At CMS dominantly the anti-k_{T} algorithm [95] is deployed in which the parameter p in Eq. (4.2) is set to -1. During LHC Run 2 the radius parameter R is chosen to be 0.4 for most analysis [96]. The clustering proceeds by defining the smallest of the d_{ij} and d_{iB} parameters. If the smallest parameter is d_{ij} the objects i and j are combined. If it is d_{iB}, object i is defined as jet and removed from the list.

Jet Energy Calibration

The jet energy calibration [97] is used to relate the energy measured for a detector jet to the energy of the corresponding true particle jet at hadron level and is applied to both simulation and data. A chain of factorized corrections is applied to the raw jet four-momentum p_μ^{raw} according to

$$
p_\mu^{\mathrm{corr}} = \left[C_{\mathrm{offset}}(p_{\mathrm{T}}^{\mathrm{raw}}) \cdot C_{\mathrm{MC}}(p_{\mathrm{T}}', \eta) \cdot C_{\mathrm{rel}}(\eta) \cdot C_{\mathrm{abs}}(p_{\mathrm{T}}'')\right] \cdot p_\mu^{\mathrm{raw}}
$$

(4.3)

where p_{T}' is the transverse momentum after applying the offset correction and p_{T}'' is the transverse momentum after the offset correction, the Monte Carlo calibration and the relative jet energy scale is applied. The definition of each individual correction is:

Offset Correction — C_{offset}
The purpose of the offset correction is to estimate and subtract the energy not associated with the hard scatter which includes contributions from pile-up and electronics noise.

Monte Carlo Calibration — C_{MC}
This calibration is based on simulation and corrects for the difference between the momentum of the reconstructed jet and the generated jet. For this purpose the response variable $\mathcal{R} = \frac{p_{\mathrm{T}}^{\mathrm{reco}}}{p_{\mathrm{T}}^{\mathrm{gen}}}$ is introduced. The inverse of the average response $C_{\mathrm{MC}} = \frac{1}{\langle \mathcal{R} \rangle}$ is the correction factor to be applied.

Relative Jet Energy Scale — C_{rel}
For the measurement of the response of a jet at any η relative to the jet energy response the dijet p_{T}-balancing method, as described in Ref. [97], is used. This is done by requiring one jet in the central region ($|\eta| < 1.3$) and

one jet at arbitrary η. The relative residual correction makes the response flat as a function of $|\eta|$ and furthermore ensures that jets are described similarly in data and simulation.

Absolute Jet Energy Scale — C_{abs}

The absolute jet energy scale is measured in the reference region $|\eta| < 1.3$ using γ/Z+jets events and is designed to make the response uniform in p_{T}. Events from γ and Z are used because their energy can be accurately measured in the tracker and muon systems (Z $\rightarrow \mu\mu$) or in the ECAL (γ, Z $\rightarrow$ ee). For the measurement the missing transverse momentum projection fraction (MPF) [97] method is used, with the result being verified with the p_{T}-balancing method.

b-Jet Identification

The identification of jets originating from b-quark hadronization is an important element of many SM measurements and searches for BSM physics. Several properties of b hadrons, such as their lifetime and high mass, can be exploited to distinguish b-jets from the dominant light quark and gluon jet production. The lifetime of b quarks is about $c\tau \approx 450\,\mu\mathrm{m}$. This leads to measurable displaced secondary vertices in the tracker. Due to their high mass the decay products tend to have larger transverse momentum relative to the jet axis compared to decay products from light partons. The b-tagging algorithms that have been developed in CMS rely on the information from secondary vertices and the corresponding decay products. For the analysis of data collected during 2016 the CSVv2 tagging algorithm [98] is used and for 2017 the DeepCSV algorithm [99].

4.2.5 Missing Transverse Momentum

Since the initial transverse momentum from proton-proton collisions at the LHC is negligible, the missing transverse momentum $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$ parametrizes the momentum imbalance in the transverse plane. A non-negligible $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$ may have several reasons including neutrinos since they are invisible for the detector. Other reasons are detector resolution, detector misalignment or defects but also hypothetical new particles that go undetected.

In general, the missing transverse momentum $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$ is defined as the negative vector sum of the transverse momenta from all reconstructed PF candidates.

$$\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}} = - \sum_{\mathrm{PF\ cand.}} \vec{p}_{\mathrm{T}} \tag{4.4}$$

Recoil Corrections

To mitigate the mismodeling of $\vec{p}_\mathrm{T}^{\,\mathrm{miss}}$ due to hadronic recoil in the simulation of Drell-Yan, W+jets, and Higgs boson production corrections are applied. For these *recoil corrections* the parallel ($U_\parallel$) and perpendicular ($U_\perp$) components of the hadronic recoil U are used. The difference of $U_\parallel$ and $U_\perp$ between simulation and data is used to calibrate the simulation. The corrections are derived in Z $\rightarrow$ $\mu\mu$ events where no genuine $\vec{p}_\mathrm{T}^{\,\mathrm{miss}}$ is expected. The parallel and perpendicular components are assumed to be Gaussian distributed. For both simulation and observed data the mean of the parallel component $\langle U_\parallel \rangle$ and the resolutions $\sigma\left(U_\parallel\right)$ and $\sigma\left(U_\perp\right)$ are determined. The correction factors that are propagated to $\vec{p}_\mathrm{T}^{\,\mathrm{miss}}$ are therefore

$$
\begin{aligned}
U_\parallel' &= \langle U_\parallel \rangle_\mathrm{data} + \left(U_\parallel - \langle U_\parallel \rangle_\mathrm{sim}\right) \frac{\sigma\left(U_\parallel\right)_\mathrm{data}}{\sigma\left(U_\parallel\right)_\mathrm{sim}}, \\
U_\perp' &= U_\perp \frac{\sigma\left(U_\perp\right)_\mathrm{data}}{\sigma\left(U_\perp\right)_\mathrm{sim}}.
\end{aligned}
\tag{4.5}
$$

4.2.6 Tau Leptons

Leptonic τ decays to electrons and muons are reconstructed and identified through the techniques discussed in Sections 4.2.2 and 4.2.3, respectively. The hadronic decays are reconstructed from PF particles using the Hadron Plus Strips (HPS) algorithm [100, 101]. This algorithm combines information from charged hadrons and from π^0 candidates obtained by clustering electron and photon candidates in so called *strips*. The HPS algorithm is seeded by jets clustered from PF candidates using the anti-k_T algorithm as described in Section 4.2.4.

More than $60\,\%$ of the hadronic τ decay final states contain at least one π^0, which promptly decay to pairs of photons. These photons convert with a large probability to $\mathrm{e^+e^-}$ pairs when traversing the tracker material, which then get spatially separated due to the magnetic field. These electrons and photons, in the following referred to as e/γ, originating from a π^0 decay are clustered in strips by merging all e/γ candidates with $p_\mathrm{T} > 0.5\,\mathrm{GeV}$ in a certain $\Delta\eta \times \Delta\phi$ window around the highest-p_T candidate. In Run 1 a *fixed strip* size of 0.05×0.2 in $\eta \times \phi$ was used which was replaced in Run 2 by the improved *dynamic strip reconstruction* explained in more detail below. Charged particles are required to originate from the primary vertex and have $p_\mathrm{T} > 0.5\,\mathrm{GeV}$ which ensures that the tracks have sufficient quality.

Dynamic Strip Reconstruction

A fixed window size is not always able to capture all electron and photon candidates from the hadronic τ decays due to the following reasons.

- Due to multiple scattering and bremsstrahlung of the e^+e^- pairs in the tracker material some of the particles may end up outside the fixed window.

- A charged pion undergoing nuclear interaction in the tracker material can produce low-p_T secondary particles. This may result in cascades of electrons and photons that appear outside the fixed window despite originating from τ decay remnants.

A countermeasure is to suitably increase the window size such that all these decay products can be integrated into the strip. However, for hadronic τ leptons with large transverse momenta the decay products are more collimated. A smaller strip size is therefore sufficient to take full account of the decay products while also reducing background contributions e.g. from PU to that strip. Figure 4.3 shows a simplified sketch of the disadvantage of a fixed strip window compared to a dynamic one. Based on these considerations the strip clustering has been changed from a fixed to a dynamic window size in the following way:
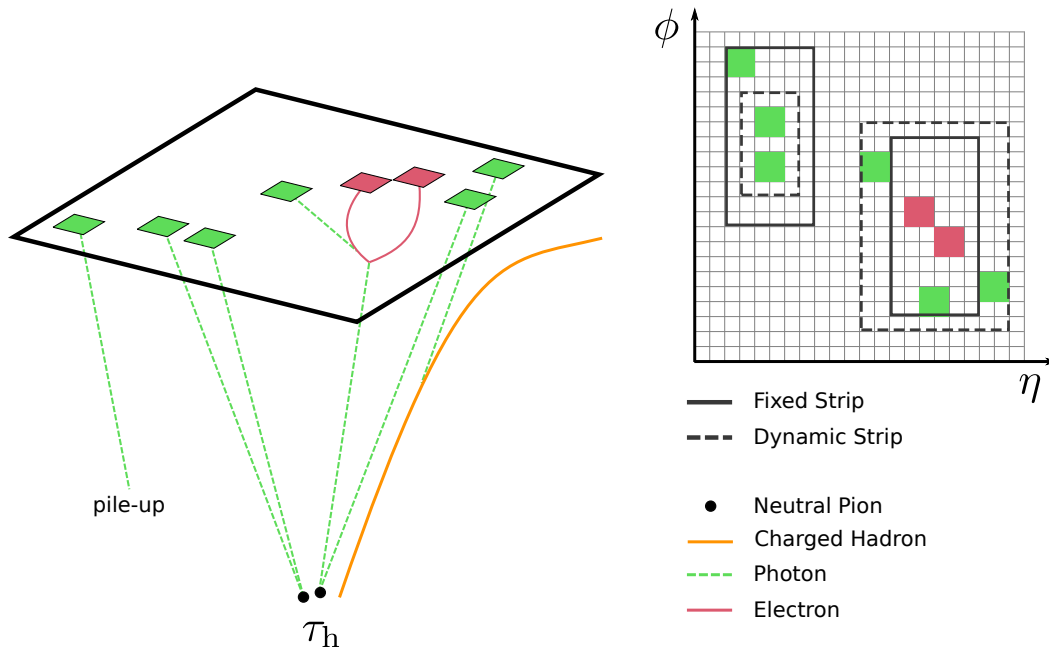
Figure 4.3 – Simplified schematic of the benefit of using the dynamic over the fixed strip reconstruction. In one case the fixed strip window size is too large and therefore includes photons not originating from the τ_h decay. In the second case not all τ decay remnants are collected by the fixed strip window since it is too small.

1. The e/γ candidate with the highest p_T that is not yet clustered in any strip is used as seed for a new strip.

2. A new e/γ candidate is added to the strip if it falls within:

$$\Delta\eta = f\left(p_T^{e/\gamma}\right) + f\left(p_T^{strip}\right),$$

$$\Delta\phi = g\left(p_T^{e/\gamma}\right) + g\left(p_T^{strip}\right). \tag{4.6}$$

The functions f and g are determined from simulated single τ lepton events such that 95 % of all electrons and photons are contained in the strip. Figure 4.4 shows the 95 % envelope of points in each bin which are fitted using $a \cdot p_T^{-b}$, resulting in:

$$f(p_T) = 0.20\,p_T^{-0.66}$$

$$g(p_T) = 0.35\,p_T^{-0.71}, \tag{4.7}$$

with upper limits on the strip size set to 0.15 (0.3) in $\Delta\eta$ ($\Delta\phi$) and the lower limits set to 0.05 in both cases. If the candidate is added the position and momentum of the strip is recomputed.

3. The strip construction ends when all e/γ candidates within $\Delta\eta \times \Delta\phi$ are clustered. If the p_T-sum of all clustered particles is above 2.5 GeV it is considered as a π^0 candidate.
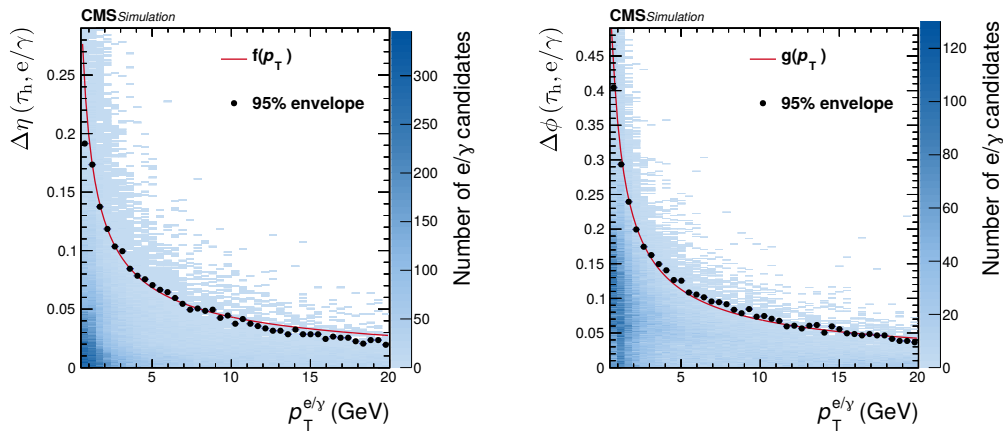
Figure 4.4 – Distance in η (left) and in ϕ (right) between the τ_h and the e/γ candidates associated to the hadronic τ decay. The τ leptons are simulated with uniform p_T in the range 20–400 GeV and $|\eta| < 2.3$. The points show the 95 % envelope for a given bin, and the solid red lines represent the fitted functions f and g given in Eq. (4.7) [101].

The strips are combined with charged particles within a signal cone with $R_{\text{sig}} < \frac{3.0}{p_{\text{T}}[\text{GeV}]}$ into different signatures compatible with a decay mode of the τ_{h} candidate. The cone size is limited to the range $0.05 < R_{\text{sig}} < 0.1$. If the τ_{h} candidate is compatible with more than one of the decay modes below the one that yields the largest p_{T} is chosen.

1-prong $+\, 0\, \pi^0$: One charged particle, no strips.

1-prong $+\, 1\, \pi^0$: One charged particle and one strip. The mass of the τ_{h} candidate is required to be within $0.3\,\text{GeV} < m_\tau < \sqrt{p_{\text{T}}^\tau} \cdot 0.13\,\text{GeV}$ with the upper mass limit constrained to lie between 1.3 and 4.2 GeV.

1-prong $+\, 2\, \pi^0$: One charged particle and two strips. The mass of the τ_{h} candidate is required to be within $0.4\,\text{GeV} < m_\tau < \sqrt{p_{\text{T}}^\tau} \cdot 0.12\,\text{GeV}$ with the upper mass limit constrained to lie between 1.2 and 4.0 GeV.

3-prong $+\, 0\, \pi^0$: Three charged particles with mass $0.8\,\text{GeV} < m_\tau < 1.5\,\text{GeV}$.

To reduce the misidentification of electrons ($\text{e} \mapsto \tau_{\text{h}}$), muons ($\mu \mapsto \tau_{\text{h}}$) and jets ($\text{jet} \mapsto \tau_{\text{h}}$) as τ_{h} candidates several discriminants have been developed.

The misidentification rate of electrons, which have a high probability to be reconstructed as 1-prong or 1-prong$+\pi^0$ decays, is reduced with the *anti-electron discriminator* [101]. This discriminant is based on a BDT and utilizes observables that are sensitive to the amount of bremsstrahlung along the leading track, observables that quantify distribution and energy deposition in the ECAL, observables to distinguish electromagnetic from hadronic showers, and observables quantifying energy, spatial distribution and multiplicities of photons in the strips. The BDT is trained on simulated events that contain genuine τ leptons or electrons. Working points ranging from `VeryLoose` to `Tight` are provided with the misidentification rates and efficiencies shown in Figs. 4.5a and 4.5b.

The *anti-muon discriminator* [101] reduces the misidentification rate of muons, which have a high probability to be reconstructed as 1-prong decays. The discriminator is cut-based with a $\mu \mapsto \tau_{\text{h}}$ misidentification probability of $\mathcal{O}(10^{-3})$ in $\text{Z} \to \mu\mu$ events. The τ_{h} identification efficiency in $\text{Z} \to \tau\tau$ events exceeds 99 %. Hadronic τ candidates are rejected when track segments in the muon detector are found near the τ_{h} direction or when the comparison between ECAL and HCAL energy deposits with the leading τ_{h} track momentum is below a certain threshold.

A strong handle to reduce the jet $\mapsto \tau_{\mathrm{h}}$ misidentification rate is to require the τ_{h} candidate to be isolated. For this purpose the isolation sum $I_{\tau_{\mathrm{h}}}$ and the p_{T} sum of the e/γ candidates outside the signal cone ($p_{\mathrm{T}}^{\mathrm{strip, outer}}$) can be used, which is explained in detail in Ref. [101]. Furthermore, a MVA-based τ_{h} identification discriminant [101] is available, which is used in the analyses presented in Chapters 5 and 6. This discriminant is based on a BDT which combines several isolation variables, differential variables sensitive to the τ lifetime and variables related to the τ_{h} decay products. The complete list of input variables to the BDT is given in Refs. [100, 101]. The training is performed using simulated τ_{h} candidates with $p_{\mathrm{T}} > 20\,\mathrm{GeV}$ and $|\eta| < 2.3$ in Z $\to \tau\tau$, H $\to \tau\tau$, Z$' \to \tau\tau$ and W$' \to \tau\nu$ events. To model quark and gluon jets simulated QCD multijet, W+jets and $t\bar{t}$ events are used. Several working points are provided for the MVA-based τ_{h} identification discriminant, corresponding to different τ_{h} identification efficiencies. The working points have efficiencies between 40–90 %, in steps of 10 %, and are called `VeryLoose`, `Loose`, `Medium`, `Tight`, `VeryTight` and `VeryVeryTight`. Figures 4.6a and 4.6b show the misidentification probability and efficiency of the provided working points.

(a)   (b)
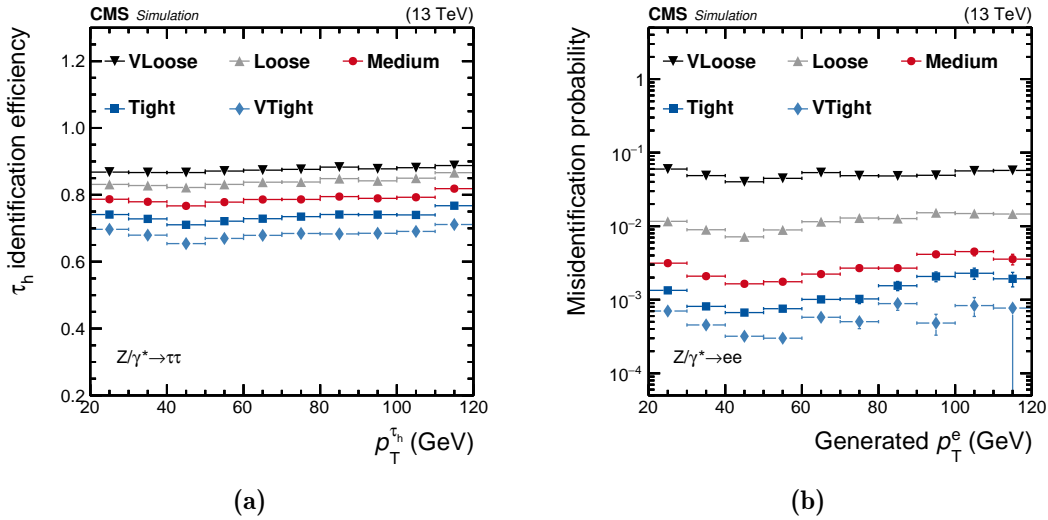
Figure 4.5 – τ_{h} identification efficiency (a) and misidentification probability of e $\mapsto \tau_{\mathrm{h}}$ events (b) given for the predefined working points of the MVA-based electron discriminator. The efficiency is calculated for generated τ_{h} candidates with $p_{\mathrm{T}} > 20\,\mathrm{GeV}$, $|\eta| < 2.3$, and matched to a reconstructed τ_{h} candidate with $p_{\mathrm{T}} > 18\,\mathrm{GeV}$ in simulated Z $\to \tau\tau$ events. The misidentification probability is calculated for generated electrons with $p_{\mathrm{T}} > 20\,\mathrm{GeV}$, $|\eta| < 2.3$, and matched to a reconstructed τ_{h} candidate with $p_{\mathrm{T}} > 18\,\mathrm{GeV}$ in simulated Z $\to$ ee events.

(a)

(b)
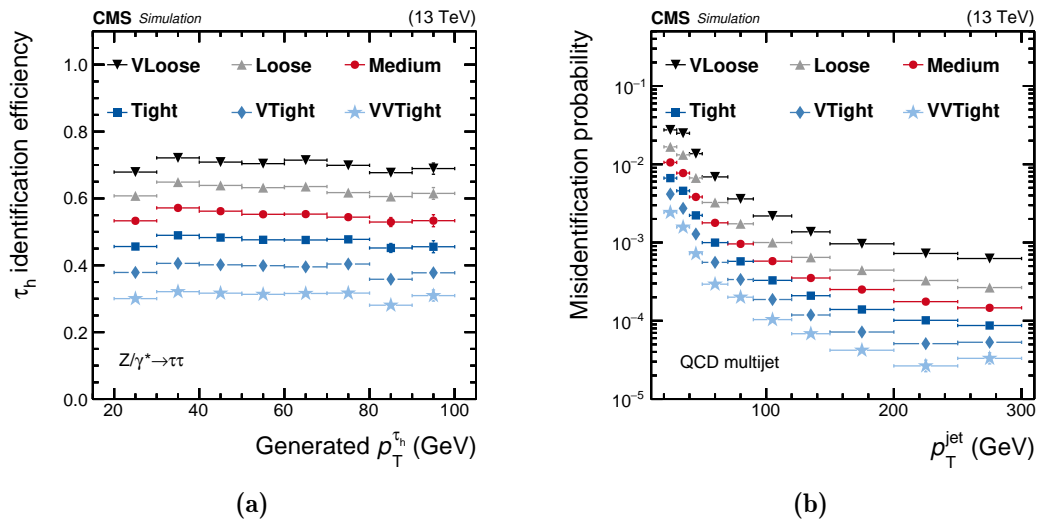
Figure 4.6 − τ_h identification efficiency (a) and misidentification probability of jet $\mapsto \tau_h$ events (b) given for the predefined working points of the MVA-based τ_h identification discriminator. The efficiency is calculated for generated τ_h candidates with $p_T >$ 20 GeV, $|\eta| < 2.3$, and matched to a reconstructed τ_h candidate with $p_T > 18$ GeV in simulated $Z \to \tau\tau$ events. The misidentification probability is calculated for jets with 20 GeV $< p_T < 300$ GeV, $|\eta| < 2.3$, and matched to a reconstructed τ_h candidate with $p_T > 18$ GeV in simulated QCD multijet events.

CHAPTER 5

MEASUREMENT OF HIGGS BOSON PRODUCTION AND DECAY TO THE $\tau\tau$ FINAL STATE

"*This summer I have discovered something totally useless.*"

– PETER HIGGS

In this chapter, a measurement of the inclusive pp $\to$ H $\to \tau\tau$ cross section is presented. Furthermore, cross sections split by production modes and different kinematic regimes are measured and presented as simplified template cross sections, as defined in Ref. [11] by the LHC Higgs Cross Section Working Group. The measurement is based on proton-proton collisions recorded in 2016 and 2017 by CMS at a centre-of-mass energy of 13 TeV, corresponding to an integrated luminosity of 77.4 fb^{-1}. In this analysis four different $\tau\tau$ final states, eμ, eτ_{h}, $\mu\tau_{\mathrm{h}}$ and $\tau_{\mathrm{h}}\tau_{\mathrm{h}}$, are studied. To improve the discrimination between the individual signal and background sources, with respect to previous CMS publications, a NN classification algorithm is introduced. Furthermore, the τ-embedding technique [102] and the Fake Factor (F_{F}) method [103] are used, allowing to estimate about 90 % of the background processes from data.

The author performed the analysis in the eτ_{h}, $\mu\tau_{\mathrm{h}}$ and $\tau_{\mathrm{h}}\tau_{\mathrm{h}}$ final states and estimated the backgrounds arising from quark or gluon jets that are misinterpreted as hadronic τ lepton decays with the fake factor method. The presented work has been published in Ref. [12].

5.1 Introduction

In 2012 the ATLAS and CMS collaborations at the LHC announced the discovery of a particle which was compatible with the long sought Higgs boson. After the

discovery both experiments determined properties of the new particle such as spin, CP, and the coupling strengths to other SM particles. In the SM the Higgs boson couples to gauge bosons via the Brout-Englert-Higgs mechanism and to fermions via Yukawa couplings. The Higgs boson was discovered in the H $\rightarrow$ ZZ, H $\rightarrow \gamma\gamma$ and H $\rightarrow$ WW decay channels, which showed that it couples to gauge bosons. An important step is to now establish the mass generation mechanism for fermions by probeing their coupling to the Higgs boson.
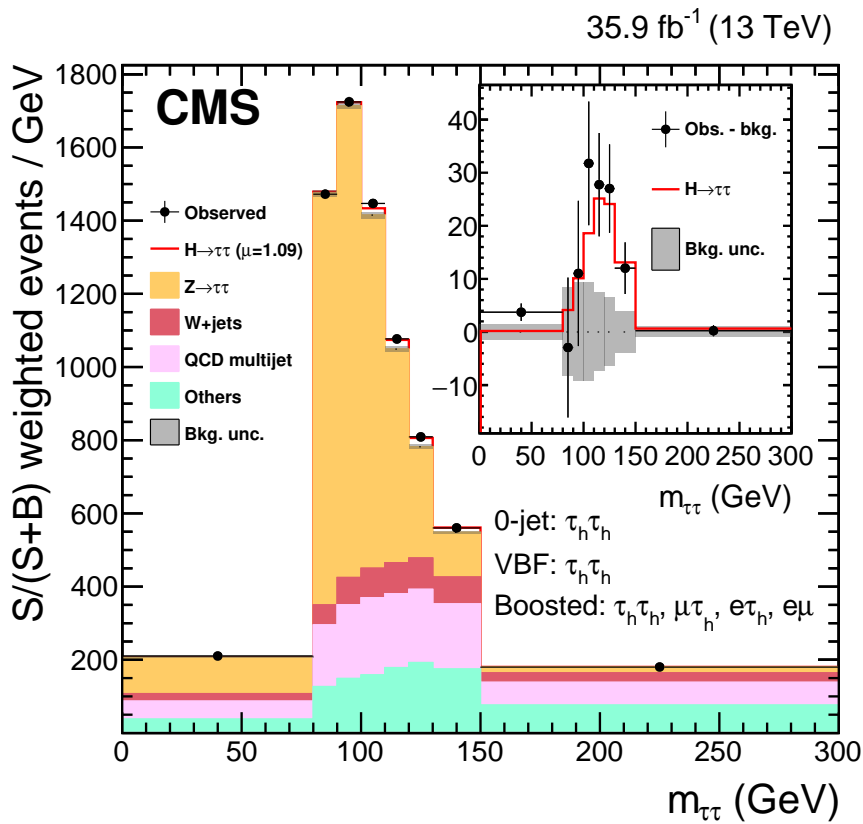
Figure 5.1 – Observed and predicted distribution of the invariant mass of the τ pair $m_{\tau\tau}$. The events are weighted according to $\frac{S}{(S+B)}$, where S and B are the expected signal or background contribution in a given bin. The inset shows the corresponding difference between the observed data and the expected background distributions [10].

The most promising fermionic decay channel is H $\rightarrow \tau\tau$, which combines a large event rate with better experimental sensitivity compared to H $\rightarrow$ bb. At a centre-of-mass energy of 7 and 8 TeV evidence was found for the decay to a τ pair by both ATLAS [104] and CMS [105]. The first observation by a single experiment was announced by the CMS collaboration after analyzing data corresponding to an integrated luminosity of 35.9 fb^{-1} collected at a centre-of-mass energy of 13 TeV

using a cut-based approach [10]. Figure 5.1 shows the observed and predicted mass distribution with a visible excess of events at the Higgs boson mass of 125 GeV.

After the observation of the Higgs boson in the $\tau\tau$ final state the next step is to provide cross section measurements at different level of differentiation, which is the target of the presented analysis. This includes an inclusive measurement of the product of the cross section for the production of the Higgs boson and the branching fraction for its subsequent decay to τ leptons, $\sigma_{\mathrm{incl}}(\mathrm{pp} \to \mathrm{h})\mathcal{B}(\mathrm{H} \to \tau\tau)$, cross section measurements split by production modes and in different kinematic regimes. The latter measurements are presented as simplified template cross sections.

In Section 5.3 the event selection in the respective $\mathrm{H} \to \tau\tau$ final states is introduced. Section 5.4 gives a short introduction to artificial neural networks, which are utilized in the event categorization in Section 5.5. In Section 5.6 the data-driven background estimation methods, τ-embedding and the F_{F} method, are discussed. The data-to-simulation corrections, uncertainty model and procedure to extract the signal are introduced in Sections 5.7–5.9. Finally, in Section 5.10, the results of the measurements are presented.

5.2 Dataset and Simulated Samples

Several MC event generators are used to generate signal and background events. Signal samples for the production of an SM Higgs boson via gluon fusion (ggH), vector boson fusion (VBF), or in association with a vector boson (VH, with V either Z or W) or a top quark pair (ttH) are simulated using POWHEG [106–111]. Single t and $\mathrm{t}\bar{\mathrm{t}}$ background samples are generated at NLO precision using POWHEG 2.0. The diboson production processes are simulated using MADGRAPH5_aMC@NLO [112, 113] or PYTHIA [114]. W+jets events and Drell-Yan events in the ee, $\mu\mu$ and $\tau\tau$ final state are generated at LO precision with MADGRAPH5_aMC@NLO. To increase the number of simulated events in regions of high signal purity additional samples with up to four outgoing jets are generated for the W+jets and Drell-Yan processes.

The NNPDF 3.0 [115] parton density functions with the underlying event tune parametrized according to the CUETP8M1 [116] tune are used for the generation of all signal and background processes. Hadronic showering and hadronization, τ lepton decays as well as additional interactions according to the expected pile-up

distribution are modelled using PYTHIA. All generated events are passed through a simulation of the CMS detector based on GEANT4 [83]. Drell-Yan, W+jets, $t\bar{t}$ and single t events in the tW-channel are normalized to their respective cross section calculated at next-to-next-to-LO (NNLO) precision [117–119]. Diboson events and single t events in the t-channel are normalized to their cross section at NLO precision or higher [119–121].

5.3 Event Selection

In this section an overview of the event selection from data recorded during the 2016 and 2017 data-taking periods by CMS and utilized in this analysis is given. The selection is aimed at identifying events in the leptonic (eμ), semi-leptonic (eτ_{h}, $\mu\tau_{\mathrm{h}}$) and fully-hadronic ($\tau_{\mathrm{h}}\tau_{\mathrm{h}}$) decay channels by using reconstructed objects as described in Section 4.2. Events are selected online during proton-proton collisions by using a trigger, followed by an offline selection of electron, muon and τ_{h} candidates. In a final step the two most isolated leptons are selected as the candidate pair originating from a Higgs boson decay. Events with extra electrons or muons, with a $p_{\mathrm{T}} > 10\,\mathrm{GeV}$ and passing loose isolation and ID requirements, are rejected to prevent overlap with other channels. In the eτ_{h} ($\mu\tau_{\mathrm{h}}$) channel events are rejected if an opposite-sign electron (muon) pair with $p_{\mathrm{T}} > 15\,\mathrm{GeV}$ and very loose isolation and ID requirements can be formed. This veto is designed to reduce the contribution from $Z/\gamma^* \to$ ee and $Z/\gamma^* \to \mu\mu$ events.

In the leptonic and semi-leptonic channels the transverse mass is used to largely suppress W+jets events. It is defined as:

$$m_{\mathrm{T}}^i = \sqrt{2\,p_{\mathrm{T}}^i\,p_{\mathrm{T}}^{\mathrm{miss}}\left(1 - \cos\Delta\phi^i\right)},$$
$$i = \mathrm{e}, \mu, \tau_{\mathrm{h}}, (\mathrm{e}+\mu),$$
$$(5.1)$$

with p_{T}^i being either the transverse momentum of the visible τ decay products or the sum of $\vec{p}_{\mathrm{T}}^{\,\mathrm{e}}$ and $\vec{p}_{\mathrm{T}}^{\,\mu}$ in the eμ final state. The difference in ϕ between p_{T}^i and the missing transverse momentum $p_{\mathrm{T}}^{\mathrm{miss}}$ is denoted as $\Delta\phi^i$.

5.3.1 Event Selection in the eμ Final State

In the eμ final state events are selected using two trigger items that require the presence of an electron and a muon. The electron is required to have $p_{\mathrm{T}} > 13\,\mathrm{GeV}$ and $|\eta| < 2.5$ and the muon $p_{\mathrm{T}} > 10\,\mathrm{GeV}$ and $|\eta| < 2.4$. To avoid events in the

lower part of the trigger turn-on curve being selected when the event only passes one trigger the higher-p_{T} trigger object is required to have $p_{\mathrm{T}} > 24\,\mathrm{GeV}$. Both leptons are required to be associated to the PV within $d_{xy} < 0.045\,\mathrm{cm}$ in the transverse plane and $d_z < 0.2\,\mathrm{cm}$ along the beam axis. Furthermore, the electron and muon are required to be separated by $\Delta R > 0.3$, have opposite charge and need to be isolated according to $I_{\mathrm{rel}}^{\mathrm{e}(\mu)} < 0.15\,(0.2)$. The working points of the electron MVA identification discriminator used in this analysis have an identification efficiency between $80\,\%$ and $90\,\%$. For muons the *medium* identification criterion [122] is used.

To supress background events from $\mathrm{t\bar{t}}$ and W+jets a requirement on the D_ζ variable [123] is imposed which is defined as:

$$D_\zeta = P_\zeta^{\mathrm{miss}} - 1.85\,P_\zeta^{\mathrm{vis}}; \qquad P_\zeta^{\mathrm{miss}} = \vec{p}_{\mathrm{T}}^{\,\mathrm{miss}} \cdot \hat{\zeta}; \qquad P_\zeta^{\mathrm{vis}} = \left(\vec{p}_{\mathrm{T}}^{\,\mathrm{e}} + \vec{p}_{\mathrm{T}}^{\,\mu}\right) \cdot \hat{\zeta}. \quad (5.2)$$

The D_ζ variable is a linear combination of projections onto the $\hat{\zeta}$ axis as shown
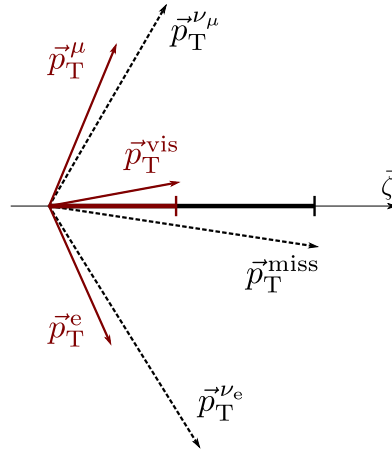
Figure 5.2 − Definition of the parameters used in the calculation of the D_ζ variable.

in Fig. 5.2. One component is the projection of $\vec{p}_{\mathrm{T}}^{\,\mathrm{e}(\mu)}$, the transverse momentum vectors of the electron and muon, and the other component is the projection of the transverse missing momentum $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$. The $\hat{\zeta}$ axis points in the direction of the bisectional direction of the electron and the muon transverse plane. All events are required to pass the requirement $D_\zeta > 35\,\mathrm{GeV}$. To ensure orthogonality between the $\tau\tau$ decay of the H boson from its decay to W bosons in the $\mathrm{e}\mu$ channel events have to pass the requirement on the transverse mass $m_{\mathrm{T}}^{\mathrm{e}^+\mu} < 60\,\mathrm{GeV}$.

5.3.2 Event Selection in the $e\tau_h$ and $\mu\tau_h$ Final State

In the semi-leptonic channels either a single-lepton or a lepton + tau trigger item is required where in the latter case at least one $e\,(\mu)$ and one τ_h candidate is required in the $e\tau_h\,(\mu\tau_h)$ channel. To suppress events from the $t\bar{t}$ and W+jets backgrounds the events have to pass a requirement of $m_T^{e(\mu)} < 50\,\text{GeV}$ in both channels.

The τ_h candidate needs to have $p_T > 30\,\text{GeV}$ and $|\eta| < 2.3$. It is required to be reconstructed by the HPS algorithm and pass the `Tight` working point of the τ_h MVA identification discriminant. Furthermore, it is required to pass the `Tight` (`VeryLoose`) working point of the anti-electron discriminator and the `Loose` (`Tight`) working point of the anti-muon discriminator in the $e\tau_h\,(\mu\tau_h)$ final state. Finally, the τ_h candidate needs to be associated to the PV along the beam axis within $d_z < 0.2\,\text{cm}$.

In the $\mu\tau_h$ channel the muon and the τ_h need to be separated by $\Delta R > 0.5$ and have opposite charge. The muon is required to have $|\eta| < 2.1$ and, depending on the data-taking period and trigger that selected the event, the p_T has to be above a threshold given in Table 5.1. Furthermore, it should pass the medium identification requirements and needs to be associated to the PV within $d_{xy} < 0.045\,\text{cm}$ in the transverse plane and $d_z < 0.2\,\text{cm}$ along the beam axis. It is also required to be isolated according to $I_{rel}^{\mu} < 0.15$ as defined in Section 4.2.3 using a cone size of $\Delta R = 0.4$.

Similar to the $\mu\tau_h$ channel the electron in the $e\tau_h$ channel is required to be separated from the τ_h by $\Delta R > 0.5$ and needs to have opposite charge. The p_T thresholds depending on the trigger are given in Table 5.1 and the electron is required to have $|\eta| < 2.1$. For data recorded in 2016 the electron needs to pass the $80\,\%$ working point of the electron MVA identification with an isolation of $I_{rel}^{e} < 0.1$ within a cone size of $\Delta R = 0.3$. In 2017 the isolation is required to be $I_{rel}^{e} < 0.15$ and the electron needs to pass the $90\,\%$ working point of the electron MVA identification. Finally, it needs to be associated to the PV within $d_{xy} < 0.045\,\text{cm}$ and $d_z < 0.2\,\text{cm}$.

Muon	$\mu\tau_{\mathrm{h}}$ trigger	Single-μ trigger
2016	$20\,\mathrm{GeV} < p_{\mathrm{T}}^{\mu} \leq 23\,\mathrm{GeV}$	$p_{\mathrm{T}}^{\mu} > 23\,\mathrm{GeV}$
2017	$21\,\mathrm{GeV} < p_{\mathrm{T}}^{\mu} \leq 25\,\mathrm{GeV}$	$p_{\mathrm{T}}^{\mu} > 25\,\mathrm{GeV}$

Electron	$\mathrm{e}\tau_{\mathrm{h}}$ trigger	Single-e trigger
2016	Not used	$p_{\mathrm{T}}^{\mathrm{e}} > 26\,\mathrm{GeV}$
2017	$25\,\mathrm{GeV} < p_{\mathrm{T}}^{\mathrm{e}} \leq 28\,\mathrm{GeV}$	$p_{\mathrm{T}}^{\mathrm{e}} > 28\,\mathrm{GeV}$

Table 5.1 – Transverse momentum requirement of the electron and muon candidates in the $\mathrm{e}\tau_{\mathrm{h}}$ and $\mu\tau_{\mathrm{h}}$ channel for each data-taking period. The requirement depends on the trigger that selected the respective event.

5.3.3 Event Selection in the $\tau_{\mathrm{h}}\tau_{\mathrm{h}}$ Final State

The triggers used to select the events in the full-hadronic final state require the presence of two τ_{h} candidates. Both candidates are required to have $p_{\mathrm{T}} > 40\,\mathrm{GeV}$, $|\eta| < 2.1$ and should pass the `Tight` working point of the τ_{h} MVA identification discriminant. The τ_{h} candidates need to be separated by $\Delta R > 0.5$, be of opposite charge and need to be associated to the PV within $d_z < 0.2\,\mathrm{cm}$. To suppress the misidentification of electrons and muons as hadronic τ decays the anti-electron and anti-muon discriminants are applied with the `VeryLoose` and `Loose` working points, respectively. The most dominant background in the $\tau_{\mathrm{h}}\tau_{\mathrm{h}}$ channel is due to QCD multijet production. To suppress this background a requirement on the vector sum of the reconstructed τ momenta and missing transverse momentum, $p_{\mathrm{T}}^{\tau\tau+\mathrm{miss}} > 50\,\mathrm{GeV}$, is required.

5.4 Artificial Neural Networks

An artificial neural network (NN) [124–126] is a machine learning (ML) model that is, to some extent, inspired by the human brain. It is able to perform a classification task by adapting connections between several artificial neurons, in order to learn some underlying model from data. The *artificial neuron* is an information-processing unit and the basic building block of a neural network. Figure 5.3 shows a mathematical model of a single neuron devised by McCulloch and Pitts [127]. Roughly speaking, it is a vector-to-scalar function that uses the input parameters x_i to calculate its output or *activation* $y(x)$. To be more specific,
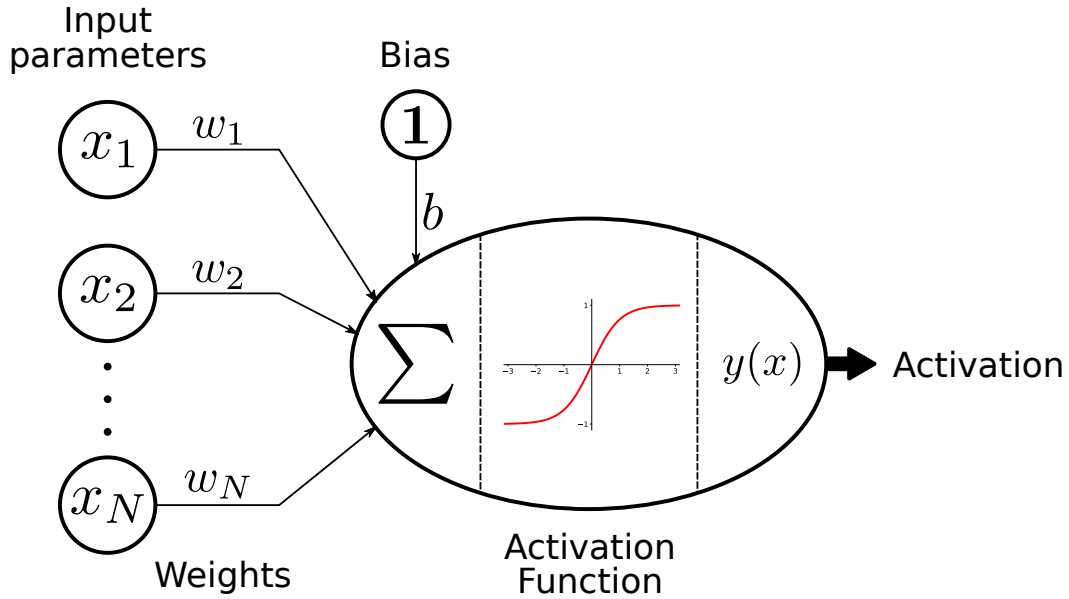
Figure 5.3 − Sketch of an artificial neuron. The input parameters x_i are connected via the weights w_i and the bias b to the neuron. The sum of these inputs is transformed using an activation function. The result of the transformation is the resulting activation of the artificial neuron.

the x_i are connected to the neuron via directed links, with each link having a numeric weight w_i. The value used to calculate the activation is the sum of all input parameters multiplied by their respective weight and the bias b. This can be expressed mathematically in the following way:

$$p(\vec{x}) = b + \sum_{i=1}^{N} w_i x_i. \tag{5.3}$$

The weighted sum $p(\vec{x})$ is the input to an activation function $a(\vec{x})$, which finally defines the activation $y(\vec{x})$ of a neuron in terms of the input parameters x_i:

$$y(\vec{x}) = a(p(\vec{x})) = a\left(b + \sum_{i=1}^{N} w_i x_i\right). \tag{5.4}$$

There different types of activation functions with the most common being discussed in Ref. [128]. Relevant for this analysis are the hyperbolic tangent and softmax functions which will be introduced in the following.

Hyperbolic tangent

The hyperbolic tangent is a smooth, zero-centred, strictly monotonic function whose range lies between -1 and 1. The output of this function is given by

$$y(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}. \tag{5.5}$$

It is chosen since it has a well defined second-order derivative, which is exploited to analyse the sensitivity of the neural network to the input parameters [129].

Softmax function

The softmax function is used to compute probability distributions from a vector of real numbers and is used in multi-class models. In such models, the softmax function returns the probability of x_i to belong to a specific class. The range of the output lies between 0 and 1 with the sum of the probabilities being equal to 1. The function is defined as

$$y(x_i) = \frac{e^{x_i}}{\sum_j e^{x_j}}. \tag{5.6}$$

The *network architecture*, which describes the interconnection between neurons, utilised in Section 5.5 is a *feed-forward neural network*. In this architecture the output of the neurons of one layer are connected with the input of the neurons in the next layer. Furthermore, there are no cyclic connections or connections within the same layer of a feed-forward network. A *layer* is a group of neurons, or *nodes*, as shown in the example in Fig. 5.4. Multilayer networks have one or
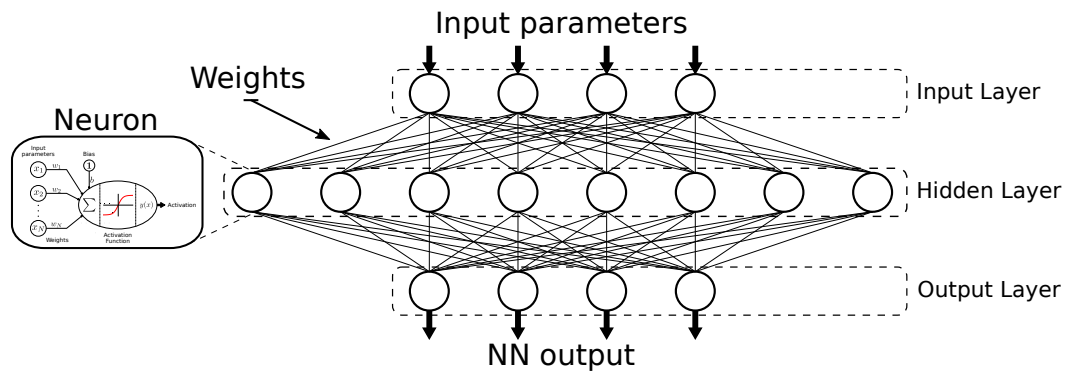
Figure 5.4 – Sketch of a neural network. The basic building blocks are artificial neurons which are grouped in layers. The input layer, hidden layer and output layer are connected via weights.

more layers of *hidden nodes*, called *hidden layers*, that are not directly connected to the input and output. These hidden layers enable the NN to act as a universal approximator according to the eponymous theorem [130]. If all nodes in each layer are connected to all nodes in the next layer the network is called *fully connected*.

5.4.1 Supervised Learning

After defining the network architecture a neural network is *trained* on labelled data $(\vec{x}^{(j)}, y^{(j)})$, called *training sample*. This type of training is called *supervised learning*. In this process, the weights and biases of the neural network are iteratively adjusted such that the NN output, for a given input $\vec{x}^{(j)}$, approaches the target value $y^{(j)}$. The goal of the training is that the neural network learns the underlying model of the labelled data and how to distinguish between entries with different labels. For this purpose, a *cost function C* is introduced that measures the accuracy of the NN output with respect to the target value via a *loss function L*. The cost function maps the parameter vector $\theta = (w_1, b_1, \ldots w_N, b_N)$, which represents the weights and biases of the neural network, to a scalar $C(\theta)$ and is typically defined as the average over the loss function values of all $(\vec{x}^{(j)}, y^{(j)})$ pairs:

$$C(\theta) = \frac{1}{n} \sum_{j=1}^{n} L\left(\hat{y}_\theta^{(j)}, y^{(j)}\right). \tag{5.7}$$

Here, $\hat{y}_\theta^{(j)}$ is the NN output computed from the input vector $\vec{x}^{(j)}$ using the configuration θ, and $y^{(j)}$ the target value. The optimal configuration of θ is obtained by minimizing the cost function with respect to the parameter θ. The minimization of $C(\theta)$ is usually performed using a gradient descent algorithm which requires the cost function to be differentiable.

A widely used loss function when using the softmax function in the output layer, and utilized for the classification task described in Section 5.5, is the cross entropy loss. It is defined as

$$L(\hat{y}_{\theta l}, y_m) = -\sum_{k=1}^{N} c_k \, y_m^k \log\left(\hat{y}_{\theta l}^k\right),$$

$$y_m^k = \begin{cases} 1 & \text{if event } k \text{ belongs to event class } m \\ 0 & \text{else} \end{cases} \tag{5.8}$$

where the index k refers to the events in a subset of $(\vec{x}^{(j)}, y^{(j)})$ of size N called batch, y_m^k is the target value for event k from event class m and $\hat{y}_{\theta l}^k$ is the NN output computed with the configuration θ, corresponding to the probability for event k to be from category l. Note that both $\hat{y}_{\theta l}^k$ and y_m^k are functions of $\vec{x}^{(j)}$. The parameter c_k is used as weight to alter the importance of a given event k during the calculation of L. This can be utilized in the case of imbalanced training samples, where the number of events in the different event classes strongly differ from each other. In this case, the parameter c_k can be used to give all event

classes the same weight during training.

After the best possible set of weights and biases is found by the training the question is if the network is able to perform well on previously unseen data. This ability of the NN is called generalization [124] and is a central problem in machine learning. For this purpose, a so-called *test sample*, which is independent from the training sample but based on the same underlying model, is used to test the NN on how well it can generalize. The ideal outcome is that the training error, which is the value of $C(\theta)$ after the training, is similar to the generalization error, which is the value of $C(\theta)$ obtained with the test sample. The performance of the NN is therefore based on its ability to minimize the training error as well as the difference between training and test error. Two phenomena, underfitting and overfitting, are the main reasons for suboptimal performance of a neural network, which may be controlled by altering the *capacity* of the NN. A generic definition of the capacity is
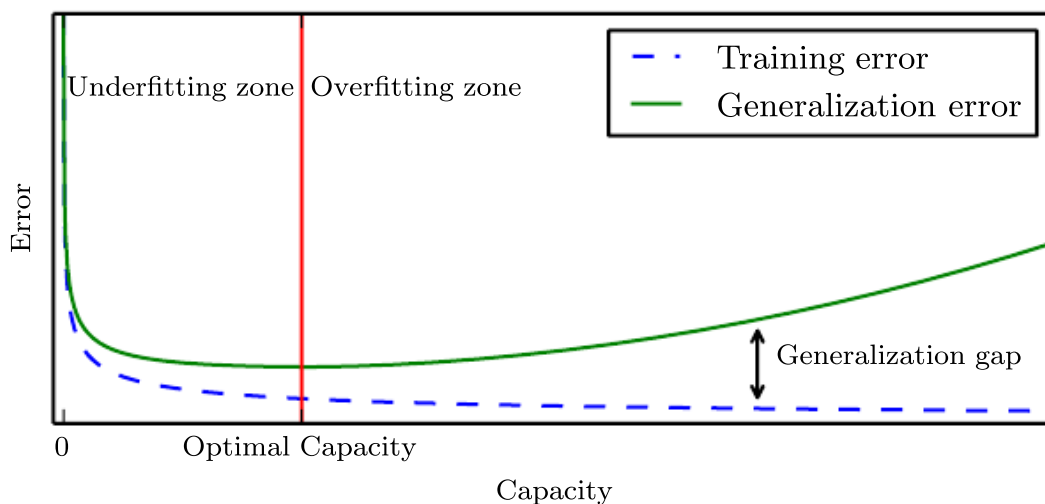
Figure 5.5 – Typical relationship between capacity and error. Training and generalization error behave differently. In the underfitting regime both training and generalization error are high. In the overfitting regime the generalization error increases, eventually outweighing the decrease in the training error [125].

the number of functions a model can fit. Underfitting occurs when the complexity of the network is not able to fit the underlying model in the training sample, leading to a high training error. The reason may be that the network architecture is chosen too simple or that the training was stopped too early (low capacity). Overfitting leads to a large gap between training and test error and occurs if the network memorizes features not related to the underlying model, e.g. statistical fluctuations (high capacity). The optimal capacity is between the underfitting and overfitting regime as shown in Fig. 5.5. Many strategies, collectively known

as regularization [124], are specifically designed to reduce the generalization error at the possible cost of increased training error.

5.4.2 Input Parameter Selection

The parameters $\vec{x}$, that are used as input to the NN classification task described in Section 5.5 are required to satisfy two quality criteria. One is the level of accuracy with which the input parameter is described by the background model and the second one characterizes its importance for the NN output y_i.

Validation

The first criterion is to validate that the distributions of the input parameters are described well by the statistical model used for the signal extraction. The motivation is that a well-described input parameter space is needed for a similar agreement for the NN output. For this purpose the background estimation model as described in Section 5.6 and the full uncertainty model as described in Section 5.8 are used. A saturated goodness-of-fit test [131], which is a generalization of a χ^2 test including all systematics and their correlation, is performed on the marginal distribution of each parameter. The single parameter is considered sufficiently well described if this test results in a p-value larger than $5\,\%$. The test considers each bin of the input distribution as an independent measurement and indicates if the model is statistically compatible with the observation. In addition to the marginal distribution of each parameter also the correlations among all input parameters are examined by applying the goodness-of-fit test on the two-dimensional distributions of pairs of variables. Parameters are removed if a systematic mismodeling in several correlations is found. Figure 5.6 shows an example for two marginal distributions on the 2016 dataset in the $\tau_\mathrm{h}\tau_\mathrm{h}$ final state and the two-dimensional distribution unrolled along a single axis to test their correlation.

Sensitivity to the Input Parameter Space

To characterize the influence of each parameter x_l in the input parameter space on the NN output the procedure described in Ref. [129] is used. This is done by performing a Taylor expansion of the NN output in the x_l up to the second order. In this approach, the Taylor coefficients are related to either single input parameters (first order) or to self- and cross-correlations of pairs of input parameters (second order). The influence of each single input parameter as well as the self- and cross-correlations with other input parameters on the NN output at a
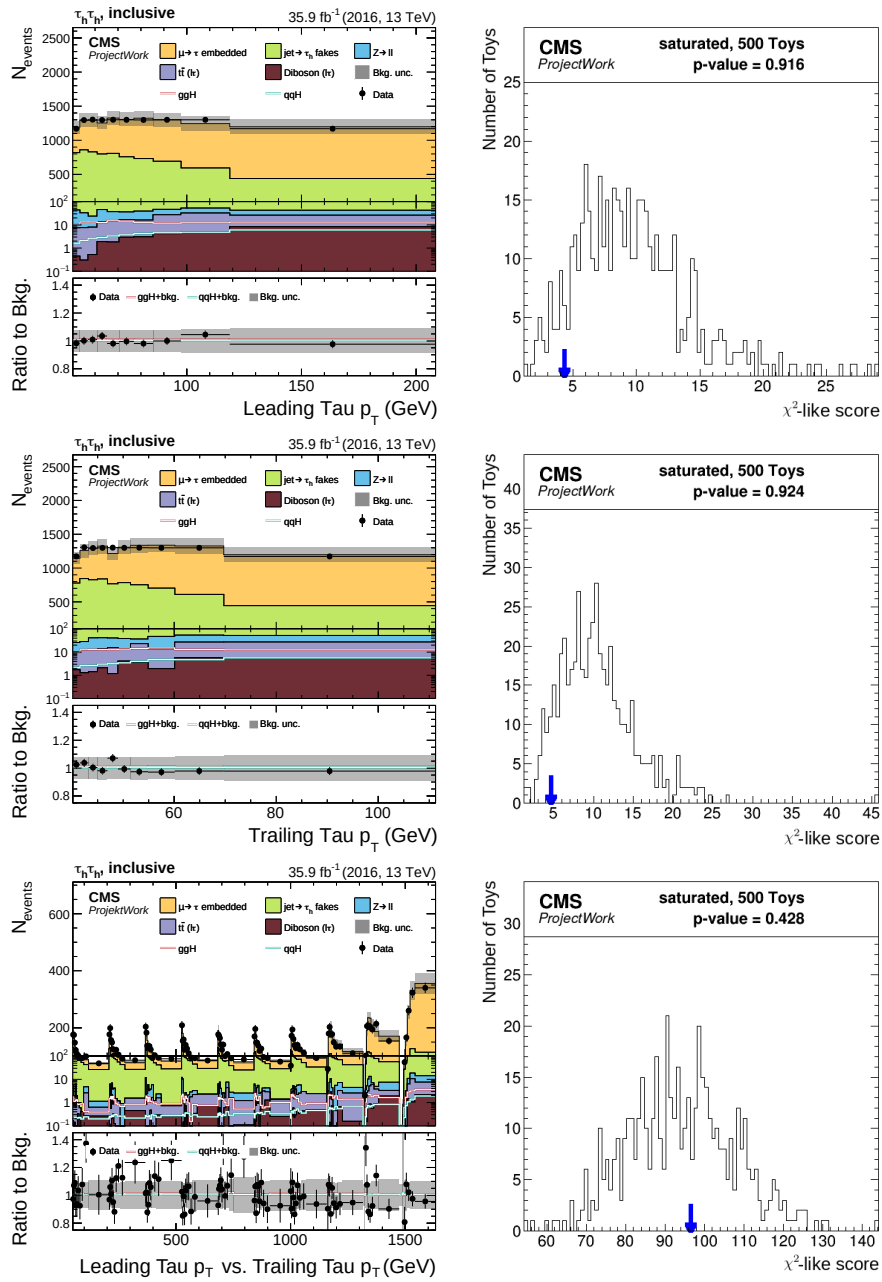
Figure 5.6 – The grey bands in the plots on the left visualize the prefit uncertainties on the simulation implemented in the uncertainty model. Each distribution of a single variable is evaluated in ten bins with bin borders at the 1%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, and 99% percentiles of the data distribution. For testing the two-dimensional distributions of variable pairs, the binning of the marginal distributions of both variables is taken and unrolled along a single axis such as shown in the bottom plot on the left. The plots on the right show the respective p-values extracted from the distribution of the test statistic determined with 500 toy experiments.

given point in the input parameter space is indicated by the magnitude of the coefficients. Based on the mean magnitude of each Taylor coefficient a ranking of the $\vec{x}$ in the input parameter space can be determined. Parameters with negligible influence on the NN output may be omitted, which therefore allows to effectively reduce the dimensionality of the input parameter space without reducing the performance. Furthermore, this procedure allows to quantify the impact of a possible mismodeling indicated by the saturated goodness-of-fit test on the NN output.

5.4.3 Training of the Neural Network

For the categorization described in Section 5.5, fully connected feed-forward neural networks with two hidden layers and 200 nodes each are used. The hyperbolic tangent is used as activation function since it has a well-defined second derivative, which can be exploited for variable selection as described in Section 5.4.2. The input parameters are preprocessed via a linear transformation by subtracting, for each input parameter, the mean and scaling it to unit variance. Each output node corresponds to an output function y_i and is activated via a softmax function. This way, the output functions y_i can be interpreted as conditional probability for an event to be associated to event class i. The set of variables used as input parameter to the NN for each final state and data-taking period is listed in Table 5.2. The definition of the individual input parameters is given in Appendix A.1. Depending on the final state the neural network has five or eight output nodes, each corresponding to one of the event classes shown in Table 5.3. Each event class is targeted at one specific process with the exception of the miscellaneous (`misc`) event classes. They comprise all processes that are either to small to be treated as single process for training or that are difficult to separate from each other.

The NN is trained on the simulated processes described in Section 5.2 except for QCD multijet production, which is taken from a control region in data where the individual candidates of the selected τ pair have the same electric charge. During training the parameter vector $\theta = (w_1, b_1, \ldots w_N, b_N)$, representing the weights and biases of the neural network, is optimized as discussed in Section 5.4. At the beginning of the training the weights are initialized with the Glorot initialization scheme [132] and the biases are initialized to zero. The training is performed on randomly sampled batches, each having 100 events, from the training sample. The Adam algorithm [133], with a constant multiplicative learning rate of 10^{-4}, is used to minimize the cross entropy loss function L given in Eq. (5.8). During

Variable	$e\mu$	$e\tau_h$	$\mu\tau_h$	$\tau_h\tau_h$	Variable	$e\mu$	$e\tau_h$	$\mu\tau_h$	$\tau_h\tau_h$
$m_{\tau\tau}^{\mathrm{SV}}$	✓✓	✓✓	✓✓	✓✓	$p_{\mathrm{T}}^{\mathrm{jj}}$	✓✓	✓✓	✓✓	✓✓
$p_{\mathrm{T}}(\mathrm{jet}_2)$	✓✓	✓✓	✓✓	✓✓	$p_{\mathrm{T}}^{\tau\tau+\mathrm{miss}}$	✓✓	✓✓	✓✓	✓✓
$p_{\mathrm{T}}(\mathrm{jet}_1)$	✓✓	✓✓	✓✓	✓−	N_{jet}	✓✓	✓✓	✓✓	−✓
m_{jj}	✓✓	✓✓	✓✓	✓✓	$N_{\mathrm{b\,jet}}$	−−	✓✓	✓✓	✓✓
$\Delta\eta_{\mathrm{jj}}$	✓✓	✓✓	✓✓	✓✓	$p_{\mathrm{T}}(\mathrm{b\,jet}_1)$	−−	✓✓	✓✓	✓✓
$p_{\mathrm{T}}^{\mathrm{vis}}$	✓✓	✓✓	✓−	✓−	$p_{\mathrm{T}}(\mathrm{b\,jet}_2)$	−−	✓✓	✓✓	✓✓
m_{vis}	✓−	✓−	✓−	✓✓	$p_{\mathrm{T}}^{\mathrm{miss}}$	−✓	✓✓	✓−	✓−
$p_{\mathrm{T}}^{\tau_1}$	−−	−−	✓−	✓✓	$m_{\mathrm{T}}^{\tau_h}$	−−	✓✓	✓−	✓✓
$p_{\mathrm{T}}^{\tau_2}$	✓−	✓✓	✓✓	✓−	$m_{\mathrm{T}}^{\mathrm{e}}$	−−	✓✓	−−	−−
$\Delta R^{\mathrm{e}\mu}$	✓✓	−−	−−	−−	m_{T}^{μ}	✓✓	−−	✓✓	−−
$p_{\mathrm{T}\tau\tau}^{\mathrm{SV}}$	✓✓	−−	−−	−−	$m_{\mathrm{T}}^{\mathrm{e}+\mu}$	✓−	−−	−−	−−
$m_{\mathrm{T}\tau\tau}^{\mathrm{SV}}$	✓✓	−−	−−	−−	$\max(m_{\mathrm{T}}^{\mu}, m_{\mathrm{T}}^{\mathrm{e}})$	✓✓	−−	−−	−−
$\eta\ (\mathrm{jet}_1)$	✓−	−−	−−	−−	D_ζ	✓✓	−−	−−	−−
$\eta\ (\mathrm{jet}_2)$	✓−	−−	−−	−−	$p_{\mathrm{T}}^{\tau\tau\mathrm{jj}+\mathrm{miss}}$	−✓	−−	−−	−−

Table 5.2 − Parameters selected for the NN classification task used for the event categorization in each final state. The checkmarks (dashes) indicate that the variable was (not) used for the corresponding final state. The first (second) entry per column corresponds to the choice for the 2016 (2017) dataset. The individual input parameters are defined in Appendix A.1. Adapted from Ref. [12].

training, the performance of the NN is monitored by evaluating L on $25\,\%$ of randomly chosen events removed from the training sample. If the evaluation of L shows no further improvement over a sequence of 50 epochs, where one epoch is defined as one iteration over the full training sample, the training is stopped. Overfitting is prevented by applying two regularization techniques. Firstly, a Tikhonov regularization [134] term with a regularization factor of 10^{-5} is applied on all weights. Secondly, a dropout [135] layer with $30\,\%$ dropout probability is added after each hidden layer. Finally, two-fold cross-validation [136] is used for training in each final state and data-taking period. After the neural networks are trained, the classification proceeds as described below.

5.5 Event Categorization

After the events for each final state are selected according to Section 5.3 they are further categorized to increase the sensitivity of the analysis. The categories are designed to target one or multiple physical processes leading to two signal and several background categories as shown in Table 5.3. The categorization is

Process	Classes/Categories per final state			
	$e\mu$	$e\tau_h$	$\mu\tau_h$	$\tau_h\tau_h$
ggH	ggH	ggH	ggH	ggH
VBF	qqH	qqH	qqH	qqH
$Z \to \tau\tau$	ztt	ztt	ztt	ztt
QCD	qcd	qcd	qcd	qcd
$t\bar{t}$	tt	tt	tt	
$Z \to \ell\ell$		zll	zll	
W+jets	misc	wj	wj	misc
Diboson	db			
		misc	misc	
Single t	st			

Table 5.3 – Processes and event classes/categories for each $\tau\tau$ final state. Adapted from Ref. [12].

implemented for each final state and data-taking period in form of a classification task via independent multi-class neural networks. One benefit of this approach

is that not only signal and background are separated but that the background categories, which are highly enriched with the respective process, may be used to constrain these backgrounds during the statistical inference of the signal. Another benefit is that events from the respective process are cumulated at large NN output values in the respective category.

5.5.1 Categorization

An event that is propagated through the trained neural network gives rise to an output $\mathbf{y}$. Depending on the final state, $\mathbf{y}$ is a vector of five or eight values. The largest of the values is considered as indicator to what event class the propagated event belongs, which in turn is then categorized accordingly. This single value, further denoted as NN output, is used as final discriminator in each category. Since only the largest value of the $\mathbf{y}$ is taken into account the categories are mutually exclusive and the events are ordered within a given category by expected purity for increasing NN output.

Each category is populated also by events other than the true process. To quantify the power of the NN the Positive Predictive Value (PPV) for each class can be defined. This value is defined as the the number of events from a given class assigned to the correct category, divided by the total number of events of this class. The confusion matrices evaluated on the 2016 and 2017 dataset for all final states are shown in Figs. 5.7 and 5.8. The PPVs are the values in the diagonal of these matrices. For a random association of the events to any class the PPVs would be uniformly distributed values with an expectation value of $\frac{1}{8}$ ($\frac{1}{5}$ in the $\tau_{\mathrm{h}}\tau_{\mathrm{h}}$ channel), while larger values indicate that the NN is able to separate a specific class from all other classes. The off-diagonal entries provide information which classes get confused and not predicted as belonging to the corresponding category by the NN. In general, the confusion matrix of each final state is similar between the 2016 and 2017 dataset. In the $\tau_{\mathrm{h}}\tau_{\mathrm{h}}$ final state the NN is able to categorize the majority of events in their respective event class. However, there is a slight tendency of events from the `ggH` class being confused with `qqH` and vice versa. The networks trained in the $e\mu$, $e\tau_{\mathrm{h}}$ and $\mu\tau_{\mathrm{h}}$ final states are able to correctly categorize the majority of events from the `qqH`, `ztt` and `tt` event classes. It is also able to separate events from the `db` event class in $e\mu$ and from the `zll` and `wj` event classes in $e\tau_{\mathrm{h}}$ and $\mu\tau_{\mathrm{h}}$. A notable feature in the $e\mu$, $e\tau_{\mathrm{h}}$ and $\mu\tau_{\mathrm{h}}$ final states is the confusion of events belonging to `ggH` event class. About the same amount of events being predicted as `ggH` are predicted to belong to either the `qqH` or `ztt` event class.
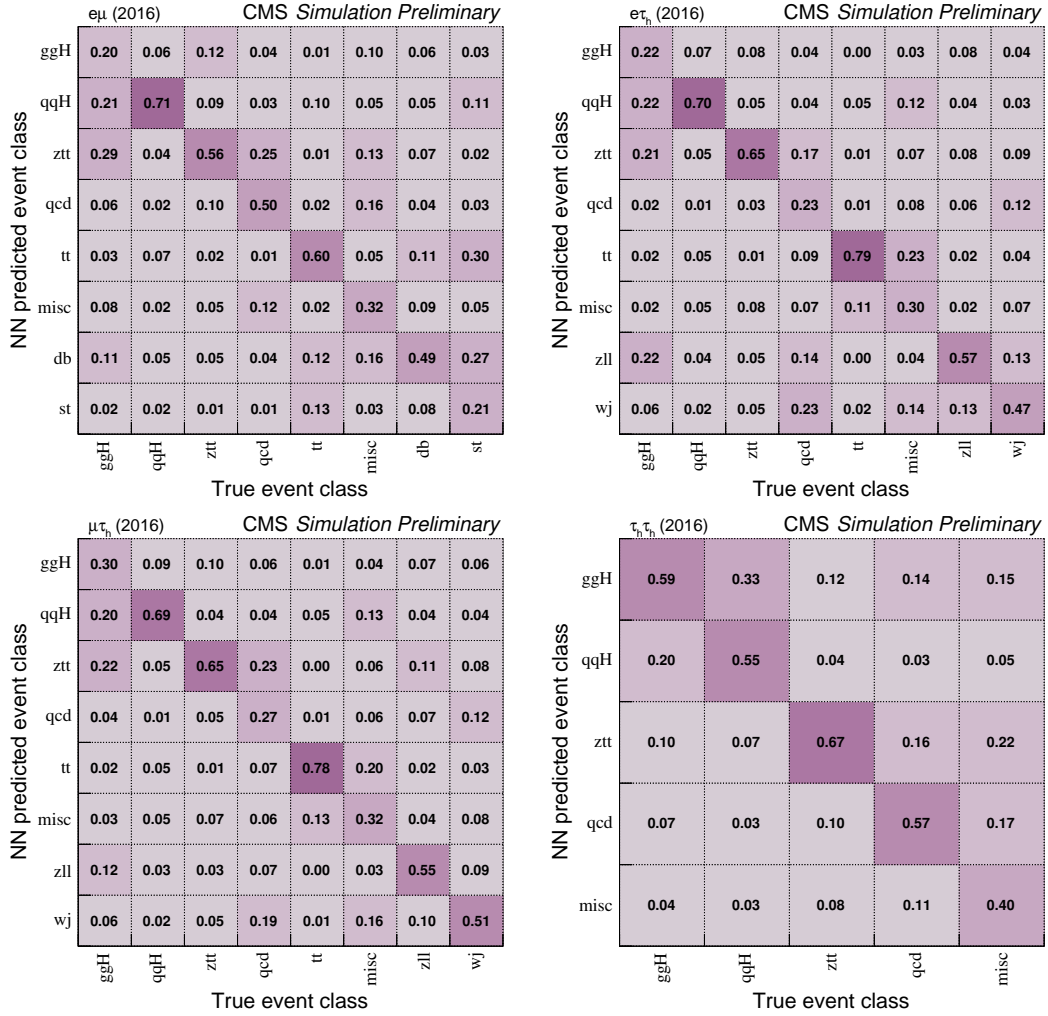
Figure 5.7 – Confusion matrices for the NN classification task, for the eμ (top left), eτ_h (top right), $\mu\tau_\mathrm{h}$ (bottom left), and $\tau_\mathrm{h}\tau_\mathrm{h}$ (bottom right) final states evaluated on the 2016 dataset. Each matrix is normalized such that all entries in each column, corresponding to a given event class, sum up to unity. All event classes enter the matrix with the same statistical weight (i.e. with uniform prevalence) [12].
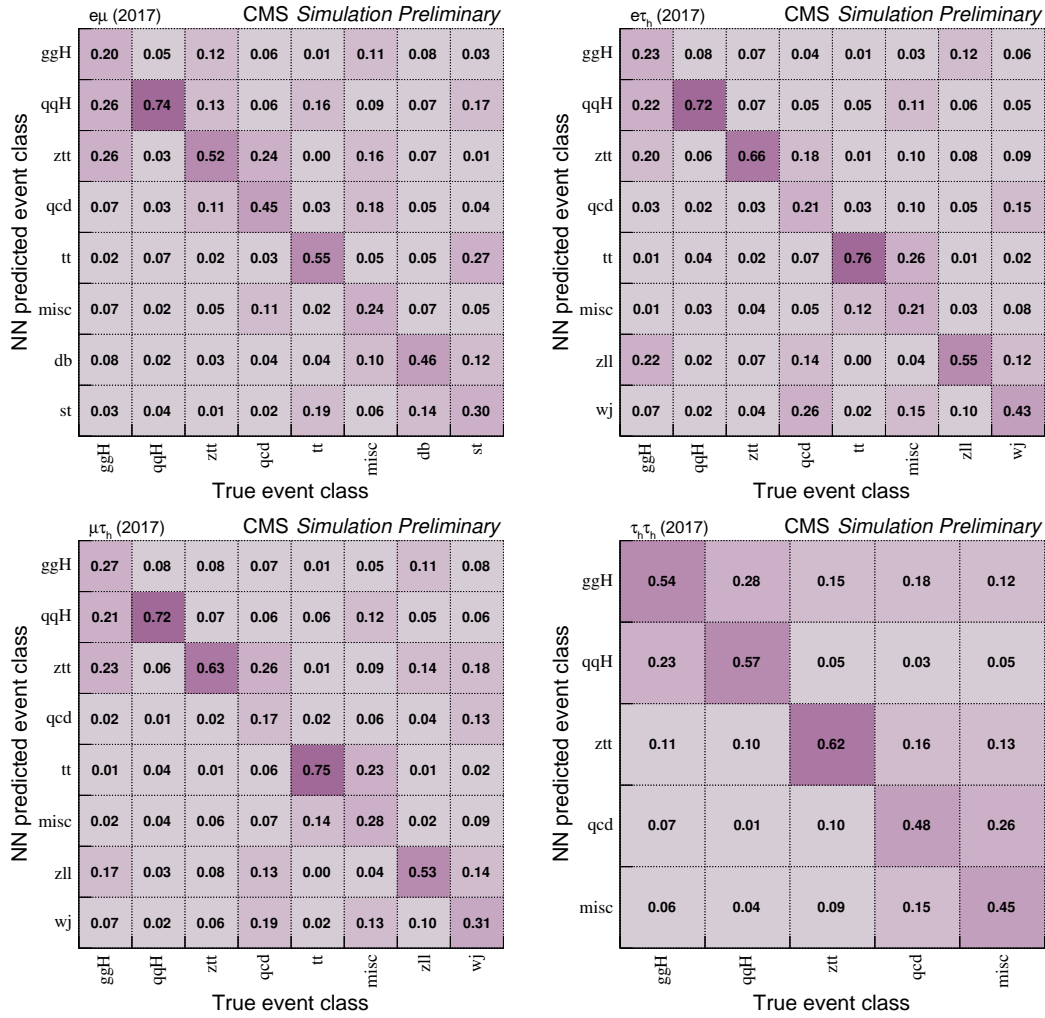
Figure 5.8 – Confusion matrices for the NN classification task, for the eμ (top left), eτ_h (top right), $\mu\tau_\mathrm{h}$ (bottom left), and $\tau_\mathrm{h}\tau_\mathrm{h}$ (bottom right) final states evaluated on the 2017 dataset. Each matrix is normalized such that all entries in each column, corresponding to a given event class, sum up to unity. All event classes enter the matrix with the same statistical weight (i.e. with uniform prevalence) [12].
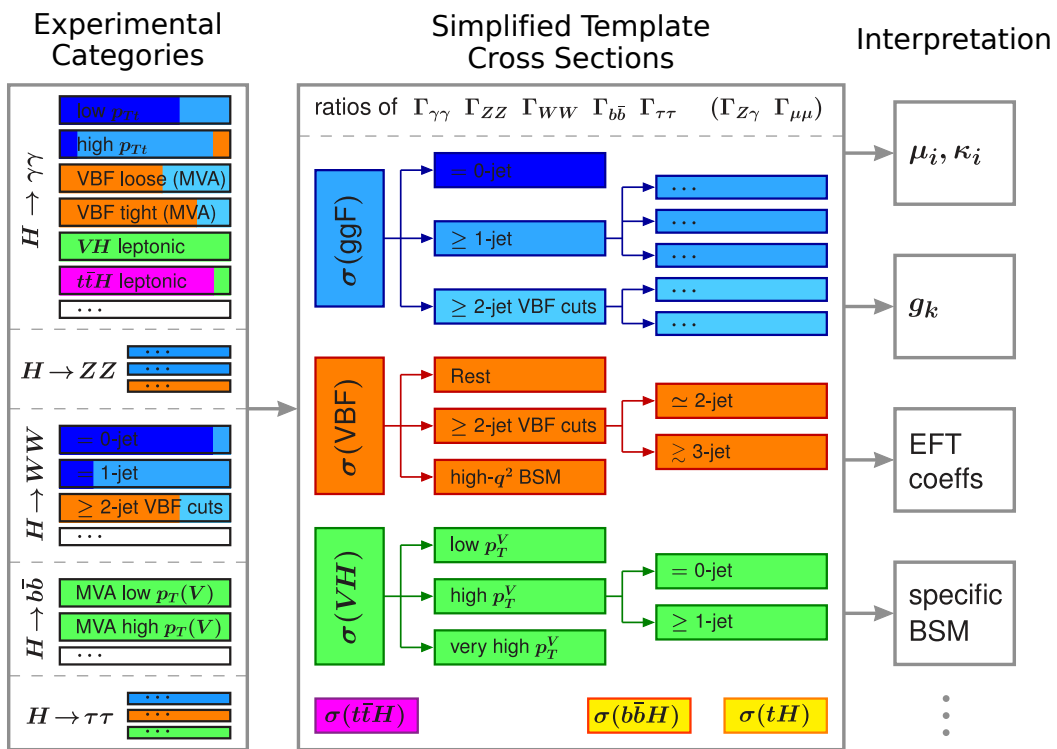
5.5.2 Simplified Template Cross-Sections

The majority of Higgs boson coupling measurements presented during Run-1 and Run-2 are in the form of signal strength μ and coupling modifier κ measurements. A natural way to evolve the signal strength measurements are simplified template cross sections, first discussed in Les Houches 2015 [137] and described in detail in Ref. [11]. The main goal of the simplified template cross section framework is to minimize the theory dependence that must be directly folded into the measurement while at the same time maximizing the sensitivity. In addition, it allows and benefits from global combination of the measurements in all decay channels and provides more finely grained measurements. Instead of signal strengths, cross sections are measured in mutually exclusive regions of fiducial phase space, called *bins*, which are specific to the different production modes. A schematic overview of the simplified template cross section framework can be found in Fig. 5.9. The

Figure 5.9 – Schematic overview of the simplified template cross section framework. Adapted from Ref. [11].

simplified template cross sections, shown at the centre of Fig. 5.9, are determined from the experimental categories by a global fit in which one or more decay channels can be combined. Together with the partial decay widths these measurements serve as input for subsequent interpretations, such as signal strength modifiers μ,

coupling scale factors κ, coefficients from effective field theories (EFT) or BSM interpretations. This approach makes the measurements more long-term useful, since the theory dependence can be decoupled from the measurement and can be dealt with at the interpretation step.

The definition of the simplified template cross sections (STXS), as described in Ref. [11], is implemented in two stages with different levels of granularity:

> **Stage 0:** Each main production mode has a single inclusive bin. The two combined production modes measured in this analysis are $gg \to H, bbH$ and $VBF+V(qq)H$. This stage corresponds most closely to the signal strength measurements performed during Run-1.

> **Stage 1:** Corresponds to the binning that should be used on an intermediate time scale by all Higgs boson measurements. The bins are defined in terms of quantities that are directly measured by the experiment, such as e.g. jet multiplicity, and a corresponding truth bin to largely avoid the theoretical uncertainties.

The presented analysis measures the inclusive cross section of the $gg \to H, bbH$ and $VBF+V(qq)H$ productions modes combined, separately (stage 0), and the stage 1 STXS. In the stage 1 classification scheme events are split into subcategories by requiring 0, 1 or ≥ 2 jets with $p_T > 30\,\text{GeV}$. Events are further split by the p_T of the Higgs boson (p_T^H), or the vector sum of $\vec{p}_T^H$ and $\vec{p}_T^{jj}$ of the system of the leading jets (p_T^{Hjj}) if more than 2 jets are found. Further distinction for the classification of events from the $VBF+V(qq)H$ process is made by the p_T of the leading jet p_T^{j1}, the invariant mass of the dijet system (m_{jj}), and the separation in η of the two leading jets ($\Delta\eta_{jj}$). Events with at least one jet and $p_T^H > 200\,\text{GeV}$ in the $gg \to H, bbH$ production mode are particularly sensitive to BSM physics. For both production modes, the subclasses with ≥ 2 jets and VBF topology distinguish between two and more than two jets. The inclusive, stage 0 and stage 1 cross section measurements are performed using the events classified by the NN to belong to the `ggH` and `qqH` categories, while the remaining categories are used to control the background processes. Figures 5.10 and 5.11 show the binning of the STXS stage 1 splitting utilized in this analysis, which is defined by quantities at generator level, for both production modes. The `ggH` and `qqH` categories are split according to the stage 1 event classes described above by using a cut-based approach. For this purpose, the requirements imposed on the quantities defined on generator level are replaced by a corresponding selection on reconstructed observables. The p_T of the Higgs boson, p_T^H, is replaced by $p_T^{\tau\tau+\text{miss}}$,
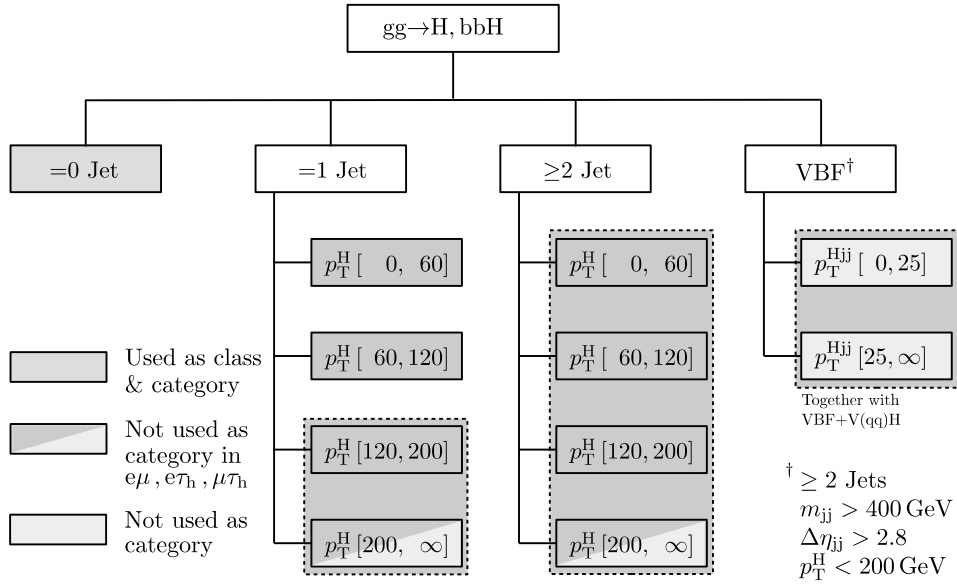
Figure 5.10 − Classification scheme as proposed by the LHC HXSWG [11] to further investigate the kinematic properties of $gg \rightarrow H, bbH$ production. All proposed classes are indicated by gray boxes. The darker colour indicates classes that are also used as categories in the analysis. The lighter coloured classes are not used as categories (in all final states). All bin boundaries are given in GeV. For the statistical inference of the signal some event classes are combined as indicated by the dashed lines [12].

jets are replaced by reconstructed jets as described in Section 4.2.4, and p_T^{Hjj} by the vector sum of the $\tau\tau$ system, dijet system and the missing transverse momentum $\vec{p}_T^{\text{miss}}$ denoted as $p_T^{\tau\tau jj+\text{miss}}$. This way, each event class leads to a corresponding category with some exceptions for the ggH category, indicated by the light gray boxes in Fig. 5.10, since they would be too weakly populated. This includes the $p_T > 200\,\text{GeV}$ categories in the $e\mu$, $e\tau_h$ and $\mu\tau_h$ channels and the ≥ 2 jet category with VBF topology. This leads to 12–14 signal categories depending on the final state for the categorization of the $gg \rightarrow H, bbH$ and $VBF+V(qq)H$ processes that are used for the signal extraction in all signal models. For all three cross section measurements the same NN training is used.

5.6 Background Estimation

There are several processes that contribute to the background in the event selection in each final state with different compositions. In all four channels the $Z \rightarrow \tau\tau$ process is a dominant background since it decays to two genuine τ leptons and
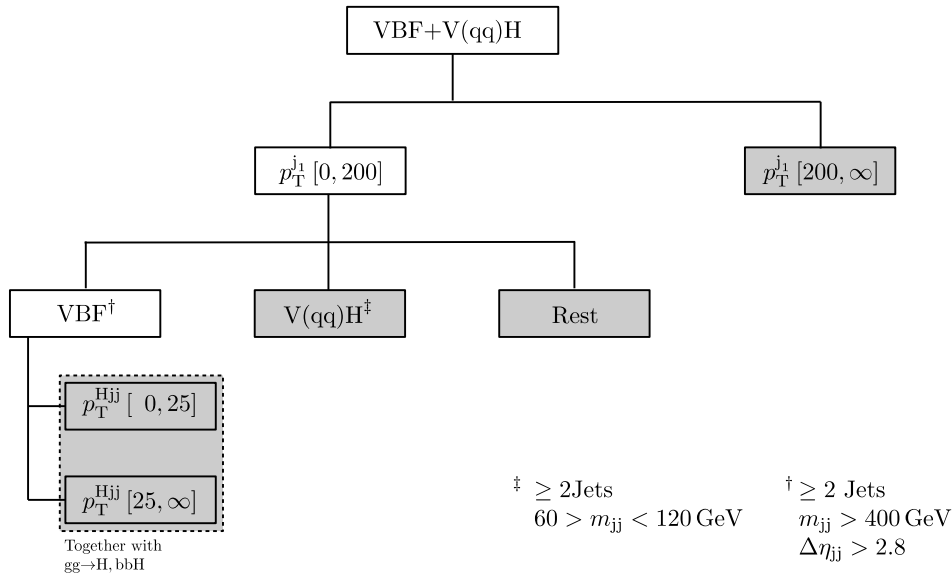
Figure 5.11 – Classification scheme as proposed by the LHC HXSWG [11] to further investigate the kinematic properties of VBF+V(qq)H production. All proposed classes are indicated by gray boxes. They are also used as categories in the analysis. All bin boundaries are given in GeV. For the statistical inference of the signal two event classes are combined as indicated by the dashed lines [12].

is therefore difficult to distinguish from the signal. To estimate the contribution from $Z \to \tau\tau$ the τ-embedding technique [102] is used. In the channels with a τ_h candidate in the final state a dominant background source is from processes where a jet is misinterpreted as hadronic τ decay. These are in particular events from W+jets and QCD multijet production but also events from $t\bar{t}$ and diboson (WW, WZ and ZZ) production enter the event selection in this way. These processes are labelled as jet $\mapsto \tau_h$ and are estimated using the F_F method [103]. The shape and normalization from QCD multijet production in the eμ channel is estimated from a control region (CR) in data. By using both the τ-embedding and the F_F method it is possible to estimate about $90\,\%$ of the backgrounds directly from collision data in the eτ_h, $\mu\tau_h$ and $\tau_h\tau_h$ channels. All processes that enter the event selection due to a lepton ℓ (with $\ell = e, \mu$) that is misidentified as a τ_h candidate are modelled with events from MC simulation. These are foremost events from $Z \to \ell\ell$ production but also from $t\bar{t}$, single t and diboson production. In the eμ channel events from W+jets are also estimated from simulation. Table 5.4 shows a list of all background processes contributing to the event selection, with the corresponding estimation technique given for each final state.

Background process	Misidentification	$e\mu$	$e\tau_h$	$\mu\tau_h$	$\tau_h\tau_h$
$Z \to \tau\tau$		EMB	EMB	EMB	EMB
W+jets		MC	F_F	F_F	F_F
QCD multijet production		CR	F_F	F_F	F_F
$Z \to \ell\ell$	Jet $\mapsto \tau_h$	MC	F_F	F_F	F_F
	$\ell \mapsto \tau_h$		MC	MC	MC
Diboson+single t [†]	Jet $\mapsto \tau_h$	MC	F_F	F_F	F_F
	$\ell \mapsto \tau_h$		MC	MC	MC
$t\bar{t}$ [†]	Jet $\mapsto \tau_h$	MC	F_F	F_F	F_F
	$\ell \mapsto \tau_h$		MC	MC	MC

[†] Decays into genuine τ leptons taken from embedded samples.

Table 5.4 – Background processes contributing to the event selection labelled with the corresponding estimation method for each final state. "EMB" is estimated with the τ-embedding technique [102], "MC" is estimated from simulation, "CR" indicates that shape and normalization are estimated from a control region in data and F_F refers to the F_F method [103]. Adapted from Ref. [12]

.

5.6.1 Tau Embedding

The τ-embedding technique [102] is a data-driven approach to estimate SM background processes decaying to two τ leptons with the most prominent being $Z \to \tau\tau$. This estimation method proceeds in four steps as sketched in Fig. 5.12. At first, $\mu\mu$ events are selected from data with a tight enough selection requirement to ensure high purity events while at the same time being loose enough to minimize biases. From the selected muons a Z boson candidate is formed by requiring two muons of opposite charge with a requirement on the invariant mass of $m_{\mu\mu} > 20\,\mathrm{GeV}$. In case of several Z boson candidates the one most compatible with the Z boson mass is chosen. After the selection, all hits in the inner tracker and muon systems as well as clusters in the calorimeters from the selected muons are removed from the reconstructed event (*cleaning*). In the next step, the energy and momentum of the selected muons is used to seed the simulation of τ lepton decays in an otherwise empty detector. In the final step, all energy deposits of the simulated τ lepton decays are combined with the cleaned original event to form a hybrid event. In this analysis, embedded samples are used to estimate, besides $Z \to \tau\tau$, the parts of $t\bar{t}$ and diboson events that decay to two genuine τ leptons.
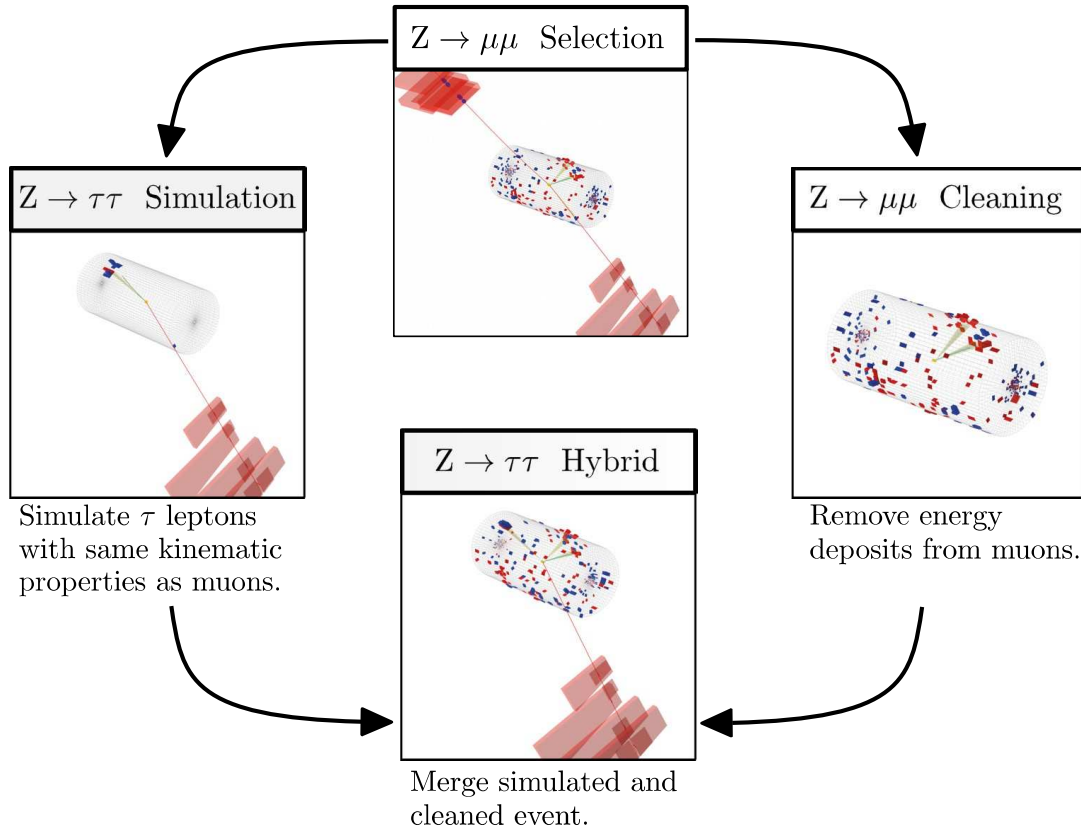
Figure 5.12 – Sketch of the four main steps of the τ-embedding technique [102].

5.6.2 Fake Factor Method

The Fake Factor (F_F) method [103] provides an estimate for backgrounds arising from quark or gluon jets that are misinterpreted as hadronic τ lepton decays. This method relies on collision data with the exception of corrections that are partially derived from simulation and are needed to remove a bias in the prediction. Each of the different processes contributing to the jet $\mapsto \tau_h$ background is composed of a specific mixture of gluon, light and heavy flavour jets. These jet flavours each have different τ_h misidentification probabilities as shown for QCD multijet production in Fig. 5.13a. As can be seen in Fig. 5.13b, the flavour dependence of the misidentification rate can be reduced by requiring the jets to pass minimal τ_h identification criteria. To get an unbiased modeling of the jet $\mapsto \tau_h$ background by the F_F method it is therefore necessary to take into account the nature of the different jet $\mapsto \tau_h$ processes and the kinematic properties of the (fake) τ_h decays.
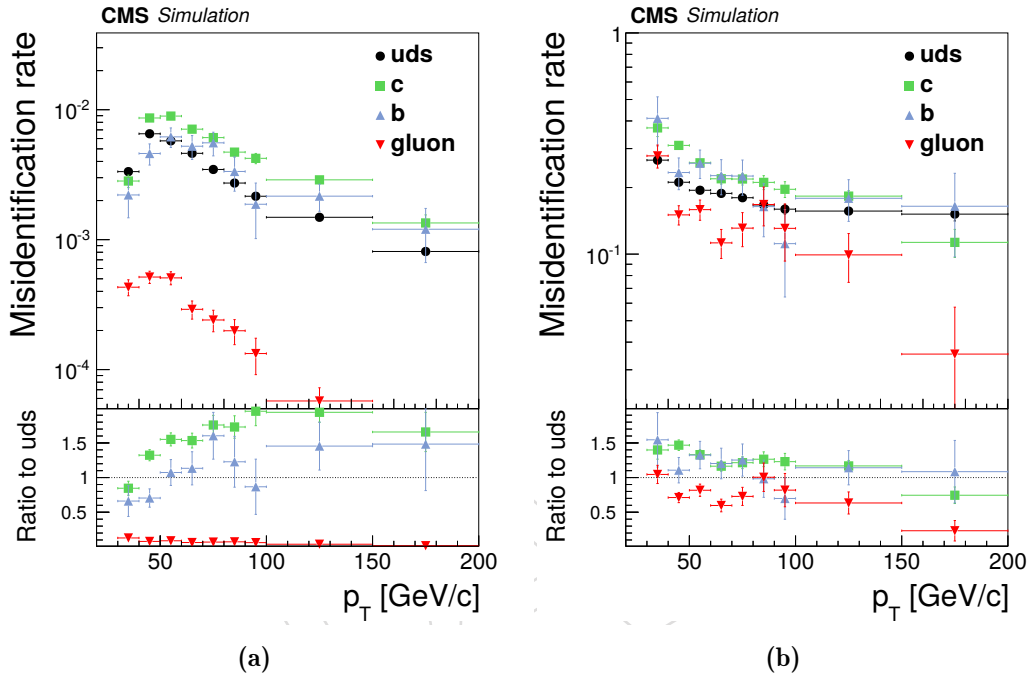
Figure 5.13 – Misidentification rate of gluon, light and heavy flavour jets in simulated QCD multijet events, for jets with $p_T > 20$ GeV and $|\eta| < 2.3$ (a) and for jets with $p_T > 20$ GeV and $|\eta| < 2.3$ passing the `VeryLoose` working point of the τ_h MVA discriminant (b). All jets are matched to a reconstructed τ_h candidate with $p_T > 18$ GeV [103].

General Idea of the Method

The general idea of the method, which is illustrated in Fig. 5.14, is to measure the number of events in a region orthogonal but not complementary to the signal region (SR). This region, referred to as application region (AR), differs from the SR only by having a modified τ_h identification (ID) requirement. The implementation used in this analysis requires the τ_h candidates in the AR to pass the `VeryLoose` but not the `Tight` working point of the τ_h MVA discriminant in the $e\tau_h$, $\mu\tau_h$ and $\tau_h\tau_h$ channels. The resulting region is very pure in events from jet $\mapsto \tau_h$ processes. Impurities from hadronic decays of genuine τ candidates are at the level of a few percent with some exceptional phase space regions such as around the Z boson mass peak where they can be up to 20 %. To estimate the jet $\mapsto \tau_h$ background in the SR from the number of events in the AR, transfer factors

$$F_F^i = \frac{N_{\text{pass}}^i}{N_{\text{fail}}^i} \tag{5.9}$$

are derived for $i = \text{QCD}, \text{W+jets}$ and $t\bar{t}$. The transfer factors F_F^i are called fake factors. They are defined as the ratio of events that pass the `Tight` working
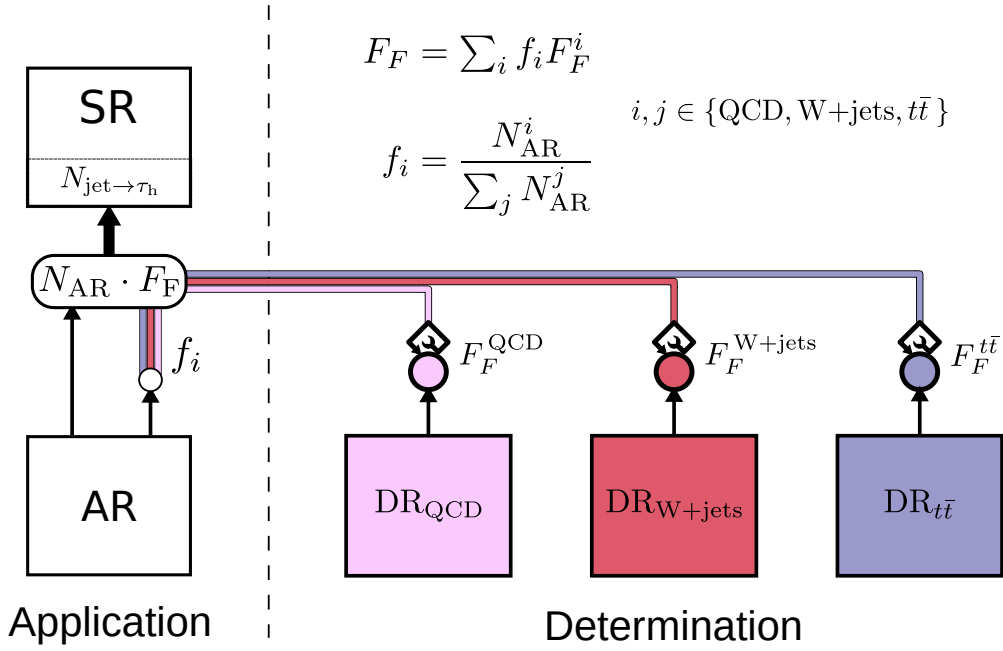
Figure 5.14 – Schematic overview of the F_F method used to estimate the jet $\mapsto \tau_\mathrm{h}$ backgrounds in the $\mathrm{e}\tau_\mathrm{h}$, $\mu\tau_\mathrm{h}$ and $\tau_\mathrm{h}\tau_\mathrm{h}$ channels. The fake factors (F_F^i) are calculated for the respective processes in determination regions (DR$_i$). In the analysis the F_F^i are applied in an application region (AR) to get the estimate for the jet $\mapsto \tau_\mathrm{h}$ background in the signal region (SR).

point (N_pass^i) and the events that pass the `VeryLoose` but fail the `Tight` working point (N_fail^i). The fake factors are derived independently for each process i in a dedicated determination region (DR$_i$), which are enriched with the respective process and defined to be orthogonal to the AR and SR. The F_F^i are determined in each DR$_i$ as a smooth function of the τ_h transverse momentum and as a function of jet multiplicity ($N_\mathrm{jet} = 0$ and $N_\mathrm{jet} \geq 1$). To finally estimate the jet $\mapsto \tau_\mathrm{h}$ background in the SR the weighted average of the F_F^i is applied to all events in the AR according to

$$F_\mathrm{F} = \sum_i f_i F_\mathrm{F}^i \qquad \text{with } f_i = \frac{N_\mathrm{AR}^i}{\sum_j N_\mathrm{AR}^j} \qquad i, j \in \left\{ \mathrm{QCD}, \mathrm{W+jets}, \mathrm{t\bar{t}} \right\}, \qquad (5.10)$$

where the weight f_i is derived from the expected number of events N_AR^i of the process i in the AR using simulation. It is possible that the jet $\mapsto \tau_\mathrm{h}$ background estimated with these raw fake factors can be imperfectly modelled in the SR in terms of shape and normalization due to:

- The underlying assumption that the fake factor is identical in the DR and SR is mildly violated. A possible cause may be a different jet flavour composition in the respective regions.

- Residual differences due to a dependency on variables that are not used to model the fake factor.

For this purpose, corrections are derived to mitigate the discrepancies.

Fake Factor Measurement and Correction: QCD multijet production

The determination region to measure the fake factors for QCD multijet production ($\mathrm{DR_{QCD}}$) for the $e\tau_h$ and $\mu\tau_h$ channels differ from the selection in the SR, as defined in Section 5.3.2, in the following way:

- The electric charges of the τ_h candidate and the light lepton are required to be the same.

- The relative lepton isolation has to be in the range $0.05 < I_{\mathrm{rel}}^{\ell} < 0.15$ (except for $e\tau_h$ in 2017, where the range is $0.02 < I_{\mathrm{rel}}^{e} < 0.15$).

These selection cuts are chosen to enhance QCD multijet events with respect to events from other processes. Poorly isolated leptons which have a relative isolation larger than the upper boundary are rejected to avoid a bias in the fake factor measurement.

For the $\tau_h\tau_h$ channel the requirement on the two τ_h candidates in the DR is that they need to have the same electric charge. In this final state the contribution of jet $\mapsto \tau_h$ processes other than QCD multijet production is very small. As an approximation, the fake factors derived in the $\mathrm{DR_{QCD}}$ are applied to all events in the AR for this channel.

Figure 5.15 shows the uncorrected QCD multijet fake factor as measured for the 2017 dataset. The corresponding corrections are shown in Fig. 5.16. The uncorrected fake factors and the corresponding corrections for the 2016 dataset as well as the corrections for the 2017 dataset in the $e\tau_h$ channel can be found in Appendices A.2 and A.5.

Corrections: The τ_h ID is correlated with some of the light lepton properties in the event. The resulting dependency is therefore corrected as a function of I_{rel}^{ℓ} in order to extrapolate the fake factor to the SR. Additionally, a small difference is observed depending on whether the τ_h and the light lepton have the same- or opposite-sign electric charge. Based on these observations three corrections are derived:
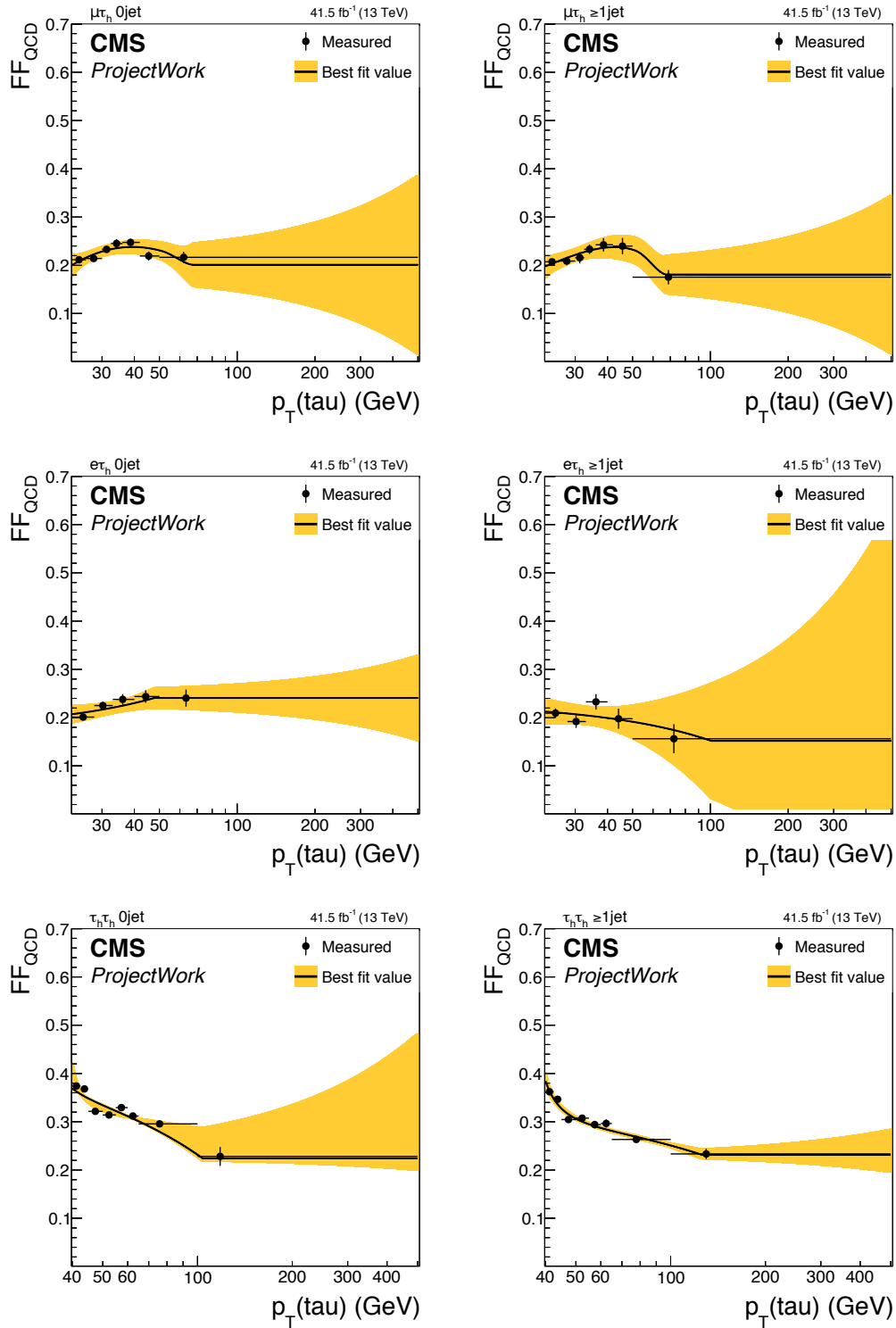
Figure 5.15 − The uncorrected fake factor for the 2017 dataset as measured in the DR_{QCD}. Shown are the measurements for $\mu\tau_h$ (top row), $e\tau_h$ (centre row) and $\tau_h\tau_h$ (bottom row). The columns are for different jet multiplicities, as indicated in the plots.

Non-closure correction

The non-closure correction is derived in the DR_{QCD} as a function of m_{vis}. The jet $\mapsto \tau_h$ background in this region is estimated using the raw fake factors and compared to the observed jet $\mapsto \tau_h$ background distribution. The ratio of the two distributions is smoothed with a Gaussian kernel with variable width to limit statistical fluctuations and the result is used as a correction.

Same-sign (SS) to the opposite-sign (OS) electric charge extrapolation

The extrapolation to the OS region is assumed to be independent of the isolation of the light lepton or the second τ_h. The correction is derived in a region with I_{rel}^{ℓ} between the upper limit, as required in the DR_{QCD}, and this upper limit plus 0.1 for $e\tau_h$ and $\mu\tau_h$, or where the second τ_h fulfills the `VeryLoose` but fails the nominal τ_h ID requirement in $\tau_h\tau_h$. In this region with a same-sign requirement, raw fake factors are derived and corrected for non-closure as a function of m_{vis}. The obtained fake factors are then applied in the same region but requiring the leptons to have opposite-sign charges. The estimated background is compared with the observed background and the ratio as a function of m_{vis} is the resulting correction function.

Correction depending on $I_{rel}^{e\,(\mu)}$ or on the p_T of the other τ_h candidate

The third correction, which depends on I_{rel}^{ℓ} in the $e\tau_h$ and $\mu\tau_h$ channel and on the p_T of the other τ_h candidate in the $\tau_h\tau_h$ channel, is derived in the SS region. Since this correction is later applied in the OS region the underlying assumption is that there is no significant correlation between the dependency of the fake factor on the relative electric charges and I_{rel}^{ℓ} (p_T).

Fake Factor Measurement and Correction: W+jets

The determination region to measure the fake factors for W+jets (DR_{W+jets}) for the $e\tau_h$ and $\mu\tau_h$ channels differ from the selection in the SR, as defined in Section 5.3.2, in the following way:

- The transverse mass m_T^{ℓ}, calculated from the p_T of the light lepton and p_T^{miss} as defined in Eq. (5.1), has to be above $70\,\text{GeV}$.

- There is no b-tagged jet allowed in the event. There is no requirement on the number of b-tagged jets in the SR.

The contamination from other background processes is small since large m_T values are mostly populated by W+jets and $t\bar{t}$ events, with the later being reduced by the b-tag veto. Since the m_T requirement is kinematically correlated to the fake
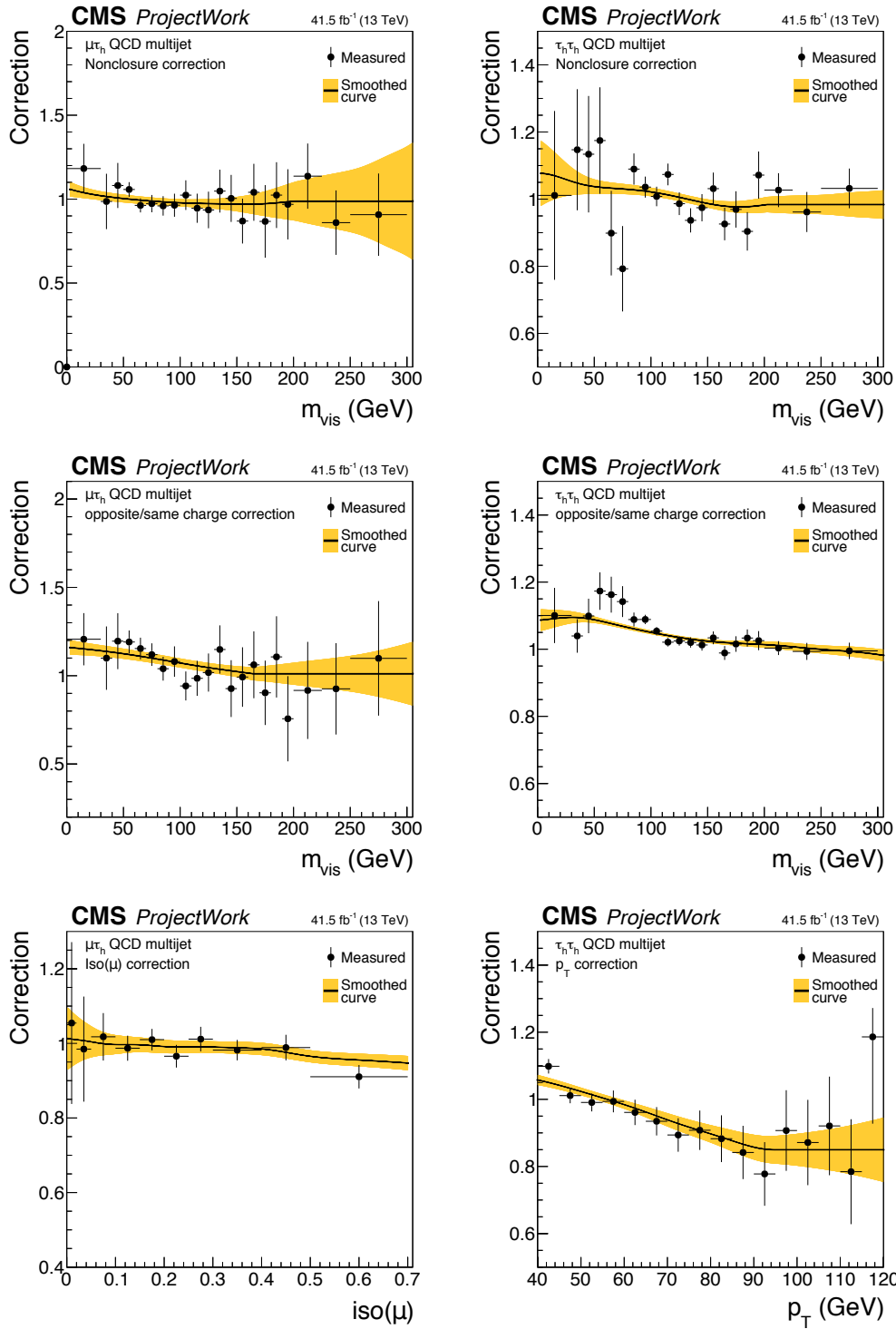
Figure 5.16 – Corrections applied on the QCD multijet fake factor measured for the $\mu\tau_\mathrm{h}$ and $\tau_\mathrm{h}\tau_\mathrm{h}$ final states on the 2017 dataset. Shown are the non-closure correction as function of m_vis (top row), the correction accounting for the same-sign to opposite-sign extrapolation (centre row) and the corrections as function of the muon isolation and the p_T of the second τ_h candidate (bottom row).

factor a correction is needed.

Figure 5.17 shows the uncorrected W+jets fake factor and the corresponding corrections as measured for the 2017 dataset. The fake factors are shown for the case with no jets in the event selection. The uncorrected fake factors and the corresponding corrections for the 2016 dataset as well as the uncorrected fake factors for the 2017 dataset for the case with at least one jet in the event selection can be found in Appendices A.2–A.4.

Corrections: By requiring a minimum p_T on the light lepton, events with a hard recoil of the W boson are favoured for low m_T, which is only possible in the presence of one or more hard jets. Since the τ_h ID is strongly correlated to the difference of jet p_T and its associated τ_h p_T, this leads to a bias towards less isolated τ_h candidates at low transverse momentum and therefore to an anti-correlation of the transverse mass with the τ_h ID for the selected W+jets events. In total, two types of corrections are derived:

Non-closure correction
The non-closure correction is derived in the $\mathrm{DR_{W+jets}}$ as a function of m_vis. The jet $\mapsto \tau_\mathrm{h}$ background in this region is estimated using the raw fake factors and compared to the observed jet $\mapsto \tau_\mathrm{h}$ background distribution. The ratio of the two distributions is smoothed with a Gaussian kernel with variable width and used as the resulting correction.

Correction depending on the transverse mass
The correction depending on m_T is derived from simulated W+jets events. The raw fake factors and the non-closure correction, as described above, are remeasured from simulation and applied to estimate the jet $\mapsto \tau_\mathrm{h}$ background arising from W+jets. The estimate is compared to the actual simulated jet $\mapsto \tau_\mathrm{h}$ background and the ratio of the two distributions as a function of m_T is used as correction. To check if this correction derived from simulation can be applied on data, fake factors derived from $\mathrm{Z} \to \mu\mu$+jets events in data and simulation are compared and found to be compatible within uncertainties.

Fake Factor Measurement and Correction: $\mathrm{t\bar{t}}$

There is no suitable pure $\mathrm{t\bar{t}}$ determination region with a sufficient event count, that populates a similar phase space as the SR in terms of jet flavour composition.
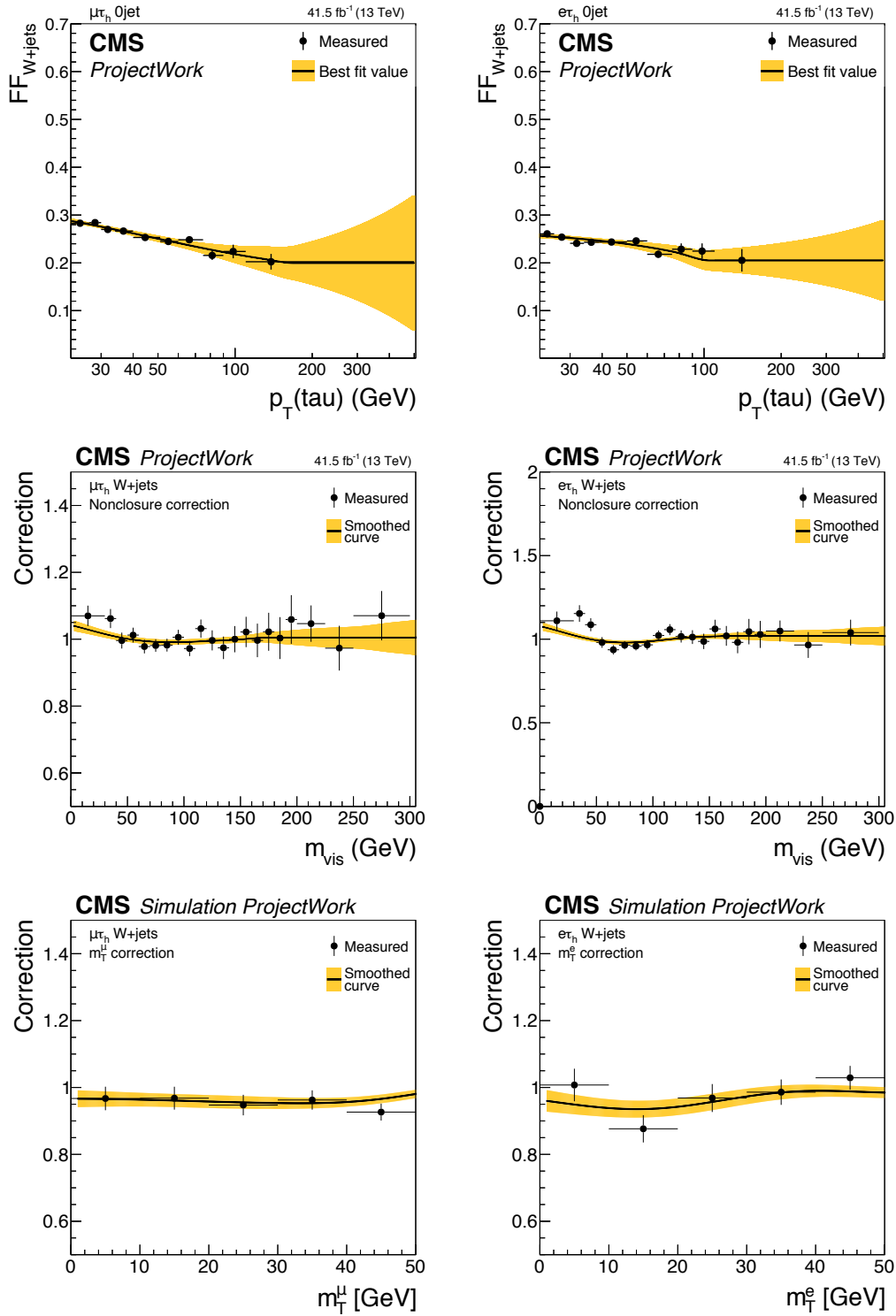
Figure 5.17 − Shown are the uncorrected W+jets fake factors measured for $\mu\tau_h$ and eτ_h for the case with no jets in the event selection (top row), the non-closure correction as function of m_{vis} (centre row), and the correction depending on the transverse mass (bottom row) for the 2017 dataset

Therefore, the $t\bar{t}$ fake factors are derived from simulation after the selection in Section 5.3.2 is applied and are then corrected using a dedicated $t\bar{t}$ CR in data.

Figure 5.18 shows the uncorrected $t\bar{t}$ fake factor and the corresponding corrections as measured for the 2017 dataset. The uncorrected fake factors and the corresponding corrections for the 2016 dataset can be found in Appendices A.2 and A.3.

Corrections: The $t\bar{t}$ CR is characterized by the same event selection as defined in Section 5.3.2 except for the following requirements:

- At least one b-tagged jet is required in the event.

- There is at least one isolated electron and one isolated muon in the event.

These requirements are needed to get a relatively pure sample. However, requiring both an electron and a muon leads to a bias: the τ_h candidate in the $t\bar{t}$ CR is typically a jet originating from a B meson since in most events both W bosons decay leptonically, whereas a significant fraction in the SR are light jets from hadronic W boson decays. For these reasons, the $t\bar{t}$ CR is only used to correct the fake factor estimate from simulation. The correction factor is the ratio of the fake factors in the $t\bar{t}$ CR in data and simulation and does not depend on p_T within uncertainties. Therefore, a common factor is used, as shown in Table 5.5.

	2016	2017
$e\tau_h$	0.80	0.97
$\mu\tau_h$	0.88	0.98

Table 5.5 – Data-to-simulation ratio for the $t\bar{t}$ fake factors in $e\tau_h$ and $\mu\tau_h$.

A non-closure correction is derived as a function of m_{vis}. For this purpose, the raw fake factors derived from simulation are applied to estimate the $t\bar{t}$ background in the SR. The result is compared to the simulated background in the SR and the ratio is used as correction.

Composition of the $jet \mapsto \tau_h$ Background in the AR

The fake factor applied to a given event in the AR is a weighted average of the F_F^i measured in the respective DR_i according to Eq. (5.10). The weight f_i is the expected fraction of events where a jet is misidentified as hadronic τ decay
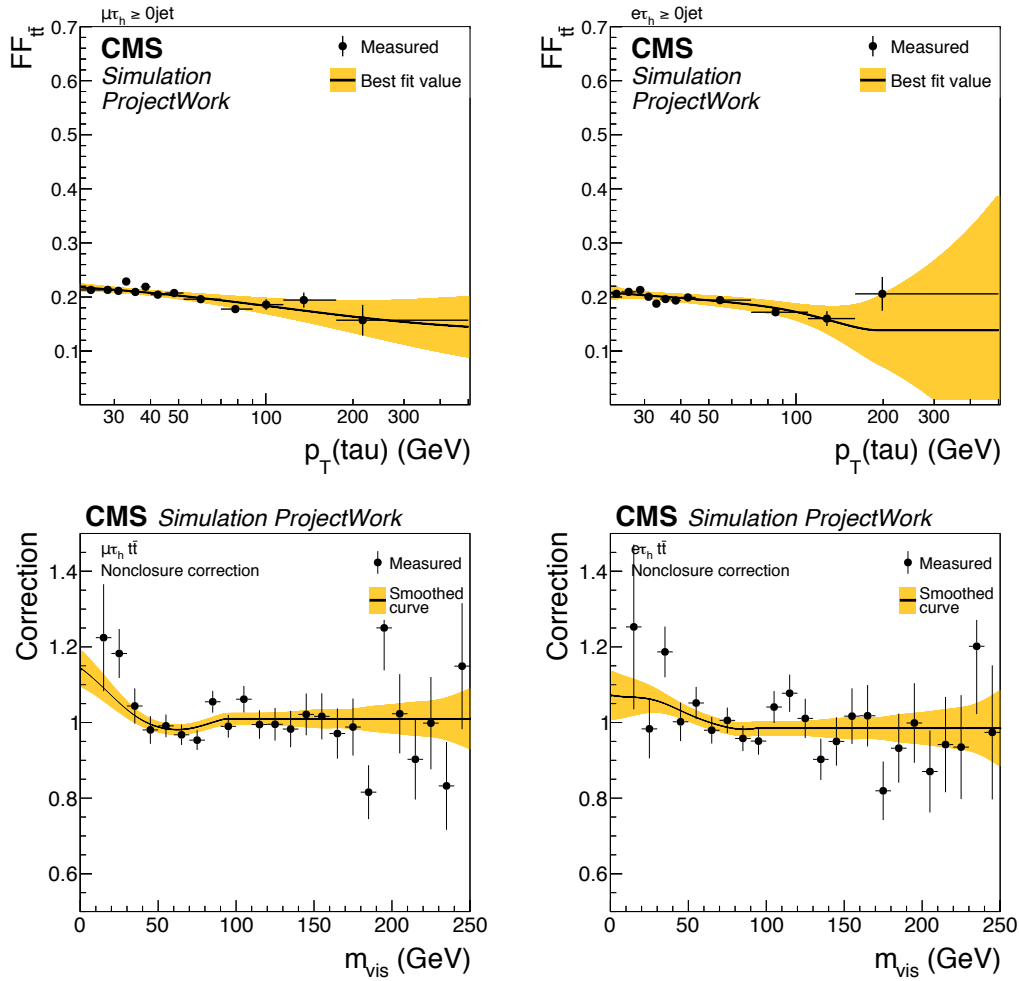
Figure 5.18 – Shown are the uncorrected $t\bar{t}$ fake factors measured for $\mu\tau_h$ and $e\tau_h$ (top row) and the simulation based non-closure correction (bottom row) for the simulated samples modelled after the 2017 dataset.

from $t\bar{t}$, W+jets and QCD multijet production in the AR. Only for the purpose of estimating these fractions, the expected number of events from W+jets and $t\bar{t}$ are taken from simulation. The estimate for QCD multijet production is obtained in the AR after subtracting all other backgrounds. The subdominant contributions from Z $\rightarrow \ell\ell$, single t and diboson production are subsumed in the estimate of the W+jets fraction. The fractions are measured for each event category listed in Table 5.3 as a function of the NN output and the jet multiplicity. To give an example: Assuming that in a given bin of the NN output the expected composition of events is 70 % W+jets and Z+jets, 20 % QCD multijet production, 5 % $t\bar{t}$ and 5 % from genuine hadronic τ decays or misidentified light leptons (mostly Z $\rightarrow \tau\tau$) then the applied fake factor is

$$ F_{\mathrm{F}} = \frac{0.7\, F_{\mathrm{F}}^{\mathrm{W+jets}} + 0.2\, F_{\mathrm{F}}^{\mathrm{QCD}} + 0.05\, F_{\mathrm{F}}^{t\bar{t}}}{0.7 + 0.2 + 0.05}. $$

This weighted fake factor is calculated and applied for all events in the AR. From the resulting distribution, the expected contribution from events with genuine hadronic τ decays or misidentified light leptons is subtracted using τ-embedding samples or simulation. This is done by considering such events in the AR, reweighting them with the appropriate fake factors and subtracting the resulting distribution from the previous fake factor estimate.

The treatment is slightly different in the $\tau_{\mathrm{h}}\tau_{\mathrm{h}}$ channel. Here, the main challenge is the ambiguity of potentially having two jets misidentified as τ_{h} in the event. Both τ_{h} candidates in QCD multijet events are due to misidentification in all but a negligible fraction of events. Therefore, each event is considered under two mutually exclusive hypotheses. The event is selected for the AR if the first τ_{h} candidate passes the nominal τ_{h} ID requirement, and the second passes the `VeryLoose` τ_{h} ID but fails the nominal requirement, or vice versa. In both cases a weight of 0.5 is applied to take the combinatorial factor into account. For the other jet $\mapsto \tau_{\mathrm{h}}$ background processes typically one of the τ_{h} candidates is identified correctly and the other is a misidentified jet. The correctly identified τ_{h} passes the nominal τ_{h} ID requirement and the misidentified jet passes the `VeryLoose` but fails the nominal requirement in the vast majority of the selected events in the AR. These events are taken into account with a weight of 1.0 since there is no significant combinatorial effect.

5.6.3 Estimation of QCD Multijet Production in eμ

The QCD multijet background in the eμ final state is estimated from a control region with the same selection as in the signal region (SR) but by requiring the electron and muon to have the same electric charge. The yield in the SR is estimated by applying same-to-opposite charge transfer factors on the data in the control region from which any non-QCD contribution is subtracted using simulation.

Figure 5.19 shows a simplified sketch of the SR and the three side-band regions used to measure the equal to opposite charge transfer factors and to validate the method. The region $\mathrm{DR_{SR}}$ is used to measure the transfer factors for the SR
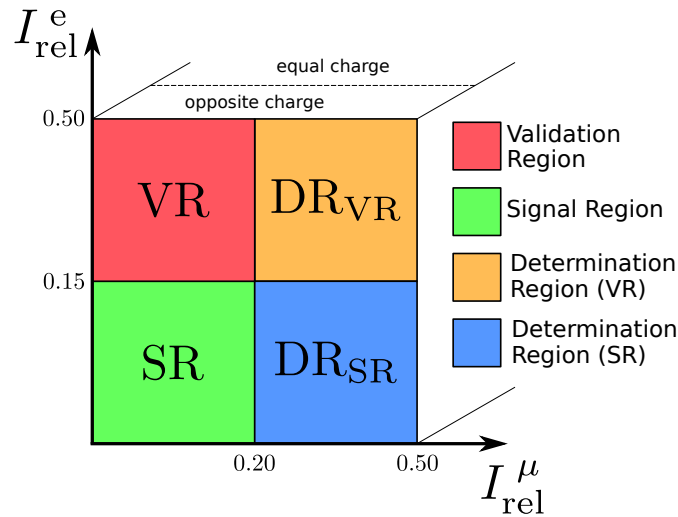
Figure 5.19 − Sketch of the different regions used to estimate and validate the QCD multijet background in the eμ channel as explained in the text.

and contains all events where the relative muon isolation is in the range $0.20 < I_{\mathrm{rel}}^{\mu} < 0.50$ and the electron passes the nominal isolation requirement of $I_{\mathrm{rel}}^{\mathrm{e}} < 0.15$. The method is tested in a validation region (VR) which contains all events where the muon passes the nominal isolation requirement of $I_{\mathrm{rel}}^{\mu} < 0.2$ and the electron passes the relaxed requirement of $0.15 < I_{\mathrm{rel}}^{\mathrm{e}} < 0.50$. The yield of the QCD multijet background in the VR is estimated using the same procedure as for the SR except for the transfer factors, which are measured in a second determination region, $\mathrm{DR_{VR}}$, where both the electron and the muon pass a relaxed isolation requirement. The transfer factors are measured as a function of the number of jets, the transverse momentum of the electron and the muon and the spatial separation of the electron and the muon in the $\eta \times \phi$ plane.

5.7 Data-to-Simulation Corrections

Pileup reweighting

As described in Section 4.1 simulated samples are generated with pileup interactions that are added to the hard-scattering process. These samples are usually generated before or during data-taking and therefore with pileup conditions that differ from the conditions in data. Furthermore, pileup conditions may change throughout an extensive data-taking period due to changing machine conditions as indicated e.g for 2017 by the double-peak structure corresponding to two subperiods with different conditions, see Fig. 5.20. In order to align the pileup distribution in the simulated samples with the distribution observed in data a reweighting technique is used.
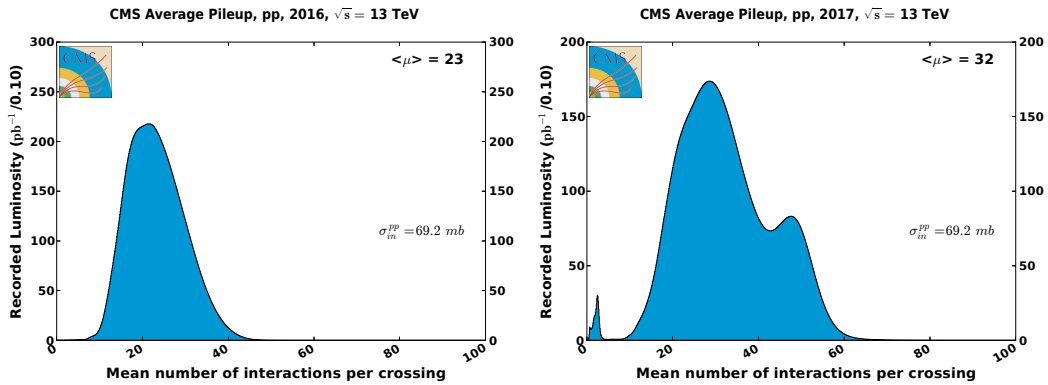
Figure 5.20 – Pileup conditions during the 2016 (left) and 2017 (right) data-taking period at CMS [138].

EE noise

During the 2017 data-taking period it was observed that noise in the ECAL endcaps at high $|\eta|$ values leads to an increase in the number of reconstructed jets and to a significant tail in the $p_{\mathrm{T}}^{\mathrm{miss}}$ distribution. To mitigate the effect, jets with $p_{\mathrm{T}} < 50\,\mathrm{GeV}$ in the region $2.65 < |\eta| < 3.139$ are removed from the event and also from the calculation of $p_{\mathrm{T}}^{\mathrm{miss}}$. This removal leads to an efficiency loss in the VBF signal of about $4\,\%$.

Tracking efficiency

Scale factors are applied to muons for 2016 and electrons for 2017 in simulated samples to take different track reconstruction efficiencies in data and simulation

into account. Furthermore, scale factors are applied for τ_h candidates in the 2016 and 2017 embedded samples.

Electron, muon and τ identification, isolation and trigger efficiency

Efficiencies for identification, isolation and the respective trigger items of electron, muon and τ_h candidates are measured in data, simulation and embedded samples. In data and simulation, events from $Z \rightarrow \mathrm{ee}$ ($Z \rightarrow \mu\mu$) are used to obtain the efficiencies for electrons (muons). For embedding, events from e (μ)-embedded event samples are used, which is explained in detail in Ref. [102]. The electron (muon) efficiencies are measured via a tag-and-probe method in bins of lepton p_T and $|\eta|$. For τ_h candidates, the identification efficiencies are measured as an inclusive factor in $Z \rightarrow \tau\tau \rightarrow \mu\tau_\mathrm{h}$ events using a tag-and-probe method. The τ_h trigger efficiencies are measured with the same type of events and method but in bins of p_T, $|\eta|$ and τ_h decay mode for 2016 and p_T, $|\eta|$ and ϕ for 2017. The ratio of the efficiencies measured in data and the efficiencies measured in simulation are applied as correction for simulated events.

$\mathrm{e} \mapsto \tau_\mathrm{h}$ and $\mu \mapsto \tau_\mathrm{h}$ misidentification rate

The background from $Z \rightarrow \ell\ell$ ($\ell = \mathrm{e}, \mu$) events where one lepton is misidentified as τ_h candidate is largely reduced by the application of the anti-electron and anti-muon discriminators. The misidentification rates for $\mathrm{e} \mapsto \tau_\mathrm{h}$ and $\mu \mapsto \tau_\mathrm{h}$, after applying the anti-lepton discriminators are measured in $Z \rightarrow \mathrm{ee}$ and $Z \rightarrow \mu\mu$ events in data and simulation using a tag-and-probe method. Scale factors derived from these measurements are applied to events in simulation where a generated electron or muon is reconstructed as hadronic tau decay.

Correction to the τ_h energy scale

Two types of energy scales are applied to simulated τ_h events depending on whether the τ_h candidate is reconstructed from a genuine τ_h or if an electron or muon is misidentified as τ_h. For genuine hadronic τ decays, the correction to the energy scale is defined as the difference between the generator level energy of the visible τ decay products and the average reconstructed τ_h energy. The correction to the energy scale of misidentified leptons is defined as the difference in the position of the Z boson mass peak in $Z \rightarrow \ell\ell$ events. The corrections are measured as a function of the decay mode of the reconstructed τ_h candidate. The respective numbers are listed in Table 5.6. No correction is applied to muons that

are misidentified as τ_h candidate for the 2017 dataset since the corrections are found to be negligible.

Correction for 2016 in %	1-prong	1-prong $+ 1\,\pi^0$	3-prong
$\tau \to \tau_h$	-0.5 ± 1.2	1.1 ± 1.2	0.6 ± 1.2
$e \mapsto \tau_h$	0.0 ± 3.0	9.5 ± 3.0	—
$\mu \mapsto \tau_h$	-0.2 ± 1.5	1.5 ± 1.5	—

Correction for 2017 in %	1-prong	1-prong $+ 1\,\pi^0$	3-prong
$\tau \to \tau_h$	0.7 ± 0.8	-0.2 ± 0.8	0.1 ± 0.9
$e \mapsto \tau_h$	0.3 ± 0.7	3.6 ± 0.7	—
$\mu \mapsto \tau_h$	—	—	—

Table 5.6 – Correction to the τ_h energy scale for the 2016 and 2017 dataset. The correction is derived for genuine τ_h decays, and events where an electron or muon is misidentified as τ_h candidate.

Recoil corrections

As described in Section 4.2.5, recoil corrections are applied to $\vec{p}_T^{\,\mathrm{miss}}$ to account for differences in resolution and response between data and simulation. The corrections are applied to simulated Drell-Yan, W+jets and Higgs production events. Note that no recoil corrections are applied to embedded samples.

Drell-Yan shape reweighting

To model the Drell-Yan background events that are not estimated by the embedded samples, simulated samples at LO precision are used. Correction factors depending on the Z boson mass and transverse momentum are applied to account for deficiencies in the modelling. The correction factors are derived in a Z $\to \mu\mu$ control region such that the transverse momentum and invariant mass distribution of the Z boson agrees with the one found in data.

Top quark p_T reweighting

To improve the modelling of the $t\bar{t}$ background a reweighting based on the p_T spectrum of the top quarks is applied. The reweighting derived during Run-1 [139]

is used. For the reweighting each $t\bar{t}$ event is assigned the weight $w_{t\bar{t}}$ defined as

$$w_{t\bar{t}} = \sqrt{e^{0.156-0.00137 \cdot p_{\mathrm{T}}(t)} \cdot e^{0.156-0.00137 \cdot p_{\mathrm{T}}(\bar{t})}}. \tag{5.11}$$

Prefiring L1 Trigger

During the 2016 and 2017 data taking a timing drift in the forward ECAL led to L1 trigger objects being associated with the previous event. The trigger logic only allows one of three consecutive events to be triggered. Therefore, this *prefiring* leads to a lower than expected data taking efficiency in event topologies involving several forward jets. For this analysis the decrease in efficiency is most prominent for VBF for which it is at the percent level. For this purpose, a correction based on the prefire probabilities of all jets is derived for 2016 and 2017 and applied on an event basis to the VBF signal.

Gluon fusion reweighting

The simulated ggH sample is reweighted to NNLO order precision using the NNLOPS generator [140]. The reweighting improves the description of the Higgs boson p_{T} spectrum and of the jets, which is of particular importance for the measurement in the context of STXS in which the signal samples are split in bins of both quantities [11].

5.8 Uncertainty Model

The sources can be divided into normalization- and shape-altering uncertainties. Normalization uncertainties only affect the yield of the NN output distributions while shape-altering uncertainties allow correlated changes across bins to alter the shape of the distributions. Furthermore, each bin of each background template is allowed to vary within its statistical uncertainty, which is taken into account by additional uncertainties that are uncorrelated across the bins of the input distribution. Systematic uncertainties are implemented in the form of nuisance paramters in the likelihood [141], which is further constrained by the fit to the observed data. A summary of the dominant systematic uncertainties is given in Table 5.7.

5.8.1 Normalization Uncertainties

Gluon Fusion STXS uncertainty scheme A dedicated uncertainty scheme, as discussed in Ref. [11], is used to account for migration of events from the

ggH to qqH categories, as well as the modelling of the number of jets in the event and the p_T^H distributions.

Luminosity

The uncertainties in the integrated luminosity are treated as uncorrelated across the two data-taking periods. They amount to 2.5 % in 2016 [142] and 2.3 % in 2017 [143].

Background cross section

Uncertainties in the background cross section are treated as fully correlated across the two data-taking periods. The uncertainties amount to 4 % for $Z \to \ell\ell$ [144], 6 % for $t\bar{t}$ and 5 % for single t and diboson production [145, 146]. To account for the extrapolation to signal and background categories additional uncertainties are applied on all simulated processes.

Electron and Muon identification and isolation efficiency

Uncertainties in the identification and isolation efficiencies for electrons and muons are treated as fully correlated across the 2016 and 2017 datasets. They amount to 2 % adding the individual contributions in quadrature.

τ_h identification efficiency

The uncertainties for the τ_h identification efficiency is split into a correlated and uncorrelated part across the final states to account for the different MVA discriminant working points. The uncertainty amounts to 5 % (3 %) per τ_h candidate on the 2016 (2017) dataset.

Trigger efficiencies

Uncertainties in the trigger efficiencies are uncorrelated for all trigger parts related to electrons and muons, and partially correlated across 2016 and 2017 for the τ_h parts of the trigger. The uncertainties amount to 2 % for electrons and muons, and 5 % for τ_h candidates. They are summed for all combined trigger parts linearly if the same object appears twice and quadratically otherwise.

e $\mapsto \tau_\mathrm{h}$ and $\mu \mapsto \tau_\mathrm{h}$ misidentification rate

Uncertainties in the e $\mapsto \tau_\mathrm{h}$ and $\mu \mapsto \tau_\mathrm{h}$ misidentification rate are partially correlated across the two data-taking periods. They are applied in the eτ_h and $\mu\tau_\mathrm{h}$ channels on the corresponding simulated $Z \to \ell\ell$ ($\ell = \mathrm{e}, \mu$) events. The uncertainty amounts to 16 % in the e $\mapsto \tau_\mathrm{h}$ misidentification rate on both datasets, and to 27 % (26 %) on the 2016 (2017) dataset for the $\mu \mapsto \tau_\mathrm{h}$ misidentification rate.

QCD sideband extrapolation
Uncertainties in the same-to-opposite charge transition factors for the estimation of the QCD multijet background in the eμ channel are treated as fully correlated across the 2016 and 2017 dataset. These uncertainties on the slope and intercept of the transfer factors are obtained from the fit that is used to extract the numbers.

5.8.2 Shape Uncertainties

τ_h energy scale
The uncertainties in the energy scale for genuine τ_h decays are partially correlated across the two data-taking periods. Independent uncertainties are applied for the 1-prong, 1-prong+π^0 and 3-prong decay modes of the τ_h candidate. They amount to 1.2 % on the 2016, and 0.8 % to 0.9 % on the 2017 dataset. A similar uncertainty, which is uncorrelated between the datasets, is applied to the p_T of electrons or muons that are misidentified as τ_h candidate. In the eτ_h channel an uncertainty of 3 % (0.7 %), and in the $\mu\tau_h$ channel an uncertainty of 1.5 % (2 %) is applied for 2016 (2017).

Electron energy scale
A fully uncorrelated uncertainty on the electron energy scale is applied in eμ and eτ_h final states. The uncertainty is 1 % in the barrel region and 2.5 % in the endcap.

Jet energy scale and b tagging efficiency
The uncertainty in the jet energy scale is treated as correlated across all final states and partially correlated across the two datasets. It affects all jet related quantities used for the event categorization as discussed in Section 5.5, which also includes the NN output. Furthermore, fully uncorrelated uncertainties on the different b tagging discriminator requirements used on the 2016 and 2017 dataset are applied.

Response and resolution of the p_T^{miss}
Uncertainties in the response and resolution of the p_T^{miss} are uncorrelated across the 2016 and 2017 datasets. They are applied to Z $\rightarrow \ell\ell$ and W+jets production as well as all signal processes and derived as part of the determination of the recoil corrections. For backgrounds from simulation which do not have recoil corrections applied, jet energy scale and unclustered energy scale variations are propagated to the p_T^{miss}. They are treated fully correlated

across the 2016 and 2017 datasets and affect the diboson, $t\bar{t}$, and single t production.

Top quark p_{T} and Drell-Yan shape reweighting
The uncertainties on both reweighting procedures, as described in Section 5.7, are treated as correlated across the two datasets. The maximum shape variations are obtained by not applying the weighting or by applying it twice.

5.8.3 Embedding-specific Uncertainties

Most of the event content and the normalization are estimated from data when using the τ-embedding technique. Nevertheless, the τ decays are modelled using simulation and are therefore subject to the uncertainties related to the selected τ candidate, with the same magnitude and correlation between the 2016 and 2017 datasets as described above. Since the conditions of the simulation of the full event are different compared to the simulation of the τ decay used for embedding the corresponding uncertainties are usually treated as uncorrelated. While this is true for the identification and isolation efficiencies of electrons, muons and τ_{h} decays, the τ_{h} energy scale, and the trigger efficiencies, there are several exceptions and additional uncertainties:

Dimuon trigger efficiency
To account for the efficiency of the $\mu\mu$ trigger used to select the muons that are replaced by simulated τ candidates an additional uncertainty of 4 % is applied.

τ_{h} identification efficiency and energy scale
The uncertainties in the τ_{h} identification efficiency, and in the τ_{h} energy scale are partially correlated with the corresponding uncertainties applied to the simulation. To account for differences in the tracking efficiency for the τ candidates, which are simulated in an otherwise empty detector, an additional uncertainty of 2 % is applied on the τ_{h} energy scale.

Fraction of $t\bar{t}$ events
The τ-embedding technique also estimates $t\bar{t}$ events with genuine τ decays. Therefore, an uncertainty in the fraction of $t\bar{t}$ events is applied, which is uncorrelated across the datasets. This is done by adding and subtracting 10 % of the simulated $t\bar{t}$ yield with genuine τ decays from the embedded event sample.

5.8.4 Fake-Factor-specific Uncertainties

Additional uncertainties related to the F_F method are applied in the $e\tau_h$, $\mu\tau_h$ and $\tau_h\tau_h$ channels. These uncertainties comprise statistical uncertainties on the F_F measurement as well as uncertainties on the applied corrections and on the background composition in the AR. The statistical uncertainties are uncorrelated and the systematic uncertainties are partially correlated across the two data-taking periods.

Statistical uncertainties
Statistical uncertainties in the measurement of the F_F^i in the DR_i obtained from the uncertainty of the fit. They are uncorrelated except for the uncertainty on $F_F^{t\bar{t}}$, which is correlated between the $e\tau_h$ and $\mu\tau_h$ final state. They range between 4–8 % in the $e\tau_h$, between 3–5 % in the $\mu\tau_h$ and between 2–4 % in the $\tau_h\tau_h$ channel.

Corrections of the QCD multijet fake factor
The following uncertainties are added quadratically, leading to one additional nuisance parameter per final state:

- Uncertainty of the non-closure correction, as a function of m_{vis}.

- Uncertainty on the correction function for I_{rel}^{ℓ} in the $e\tau_h$ and $\mu\tau_h$ channel or on the p_T of the other τ_h candidate in the $\tau_h\tau_h$ channel.

- Uncertainty on the correction function for the extrapolation from the same-to-opposite charge region.

Corrections of the W+jets fake factor
The following uncertainties are added quadratically, leading to one additional nuisance parameter in the $e\tau_h$ and $\mu\tau_h$ channel:

- Uncertainty of the non-closure correction, as a function of m_{vis}.

- Uncertainty on the correction depending on the transverse mass.

Corrections of the $t\bar{t}$ fake factor
The following uncertainties are added quadratically, leading to one additional nuisance parameter in the $e\tau_h$ and $\mu\tau_h$ channel:

- Uncertainty of the non-closure correction, as a function of m_{vis}.

- Uncertainty on the data-to-simulation correction factors.

Background composition in the $\tau_\mathrm{h}\tau_\mathrm{h}$ AR
The systematic uncertainty on the fraction of W/Z+jets and $t\bar{t}$ events with one misidentified τ_h candidate in the AR, adding two nuisance parameters for the $\tau_\mathrm{h}\tau_\mathrm{h}$ channel.

Subtraction of genuine τ background in the AR
The systematic uncertainty due to the subtraction of genuine τ backgrounds in the AR, which is correlated across categories.

5.9 Statistical Inference of the Signal

The statistical inference of the signal is based on the NN output distributions obtained from the categorization described in Section 5.5. This is done by performing a simultaneous binned maximum likelihood fit to the distributions in all event categories and final states. The resulting fit is based on the signal categories defined in Section 5.5.2 and on 21 background categories, as given in Table 5.3. These high-purity background categories are used to constrain the background processes in the signal categories. The NN output distributions on the 2017 dataset for each background category and final state are shown in Figs. 5.21–5.23. The NN output distributions on the 2016 dataset for each background category can be found in Appendix A.6. The stack of filled histograms in the top panel of each subfigure represents the expectation of the background estimation model as described in Section 5.6. Each filled histogram corresponds to the estimated template distribution for a given process. The template estimated with the τ-embedding technique is labelled as "$\mu \to \tau$-embedded". It subsumes genuine τ decays from diboson, single t and $t\bar{t}$ events. The template labelled as "Jet $\to \tau_\mathrm{h}$" comprises events from W+jets, $t\bar{t}$ and QCD multijet production with at least one jet misidentified as τ_h, which are estimated using the F_F method. In the eμ final state the contribution from QCD multijet production is estimated from a control region in data and the respective template is labelled as "QCD multijet". All other templates are estimated using simulation. The bottom panel shows the ratio of the observed number of events to the expected background. The shaded band in the bottom panel and associated to the stacked histograms in the top panel corresponds to the combination of all shape altering and normalization uncertainties in all background processes. The vertical bars on the data points represent the statistical uncertainties associated with the observed number of events.

Uncertainty	eμ	eτ_h	$\mu\tau_\mathrm{h}$	$\tau_\mathrm{h}\tau_\mathrm{h}$	Process	Variation	Impact on μ
ggH uncertainties	✓	✓	✓	✓	signal	5–38%	4.9%
Background yields	✓	✓	✓	✓	MC	4–6%	2.2%
Signal pdf/scale	✓	✓	✓	✓	signal	0.9–1.9%	1.4%
Luminosity (2016)	✓	✓	✓	✓	MC	2.5%	0.7%
(2017)	✓	✓	✓	✓		2.3%	1%
Dimuon trigger	✓	✓	✓	✓	EMB	4%	0.3%
τ_h-ID (2016)	—	✓	✓	✓	MC	5%	1.3%
(2017)	—	✓	✓	✓		3%	1.3%
e $\rightarrow \tau_\mathrm{h}$ mis-ID	—	✓	—	—	Z $\rightarrow$ ee	16%	0.5%
$\mu \rightarrow \tau_\mathrm{h}$ mis-ID	—	—	✓	—	Z $\rightarrow \mu\mu$	26%	0.5%
τ_h-trigger	—	—	—	✓	MC	5%	1.1%
Sideband extrapolation	✓	—	—	—	QCD	1–37%	0.7%
e/μ-trigger, ID, isolation	✓	✓	✓	—	MC	2%	1.4%
Z/γ^* reweighting	✓	✓	✓	✓	Z $\rightarrow \ell\ell$	100%	0.4%
t quark p_T reweighting	✓	✓	✓	✓	t$\bar{\mathrm{t}}$	100%	0.4%
Jet energy scale	✓	✓	✓	✓	MC	1–6%	2.8%
b tagging efficiency	✓	✓	✓	✓	MC	1–5%	0.7%
$p_\mathrm{T}^\mathrm{miss}$ response/resolution	✓	✓	✓	✓	MC	1–5%	2.6%
t$\bar{\mathrm{t}}$ in embedded events	✓	✓	✓	✓	EMB	10%	0.9%
τ_h energy scale (2016)	—	✓	✓	✓	MC	1.2%	3.4%
(2017)	—	✓	✓	✓		0.8–0.9%	3.5%
e $\rightarrow \tau_\mathrm{h}$ energy scale (2016)	—	✓	—	—	Z $\rightarrow$ ee	3%	0.4%
(2017)	—	✓	—	—		0.7%	0.4%
$\mu \rightarrow \tau_\mathrm{h}$ energy scale (2016)	—	—	✓	—	Z $\rightarrow \mu\mu$	1.5%	0.9%
(2017)	—	—	✓	—		2%	0.3%
$F_\mathrm{F}^{~i}$ statistical uncertainty		✓	—	—		4–8%	2.2%
		—	✓	—	F_F	3–5%	2.5%
		—	—	✓		2–4%	1.9%
		✓	—	—		4–6%	0.9%
$F_\mathrm{F}^{~i}$ corrections	—	—	✓	—	F_F	4–7%	2%
		—	—	✓		7–8%	2.1%
e energy scale	✓	✓	–	–	MC	1–2.5%	0.3%

Table 5.7 – Overview of the dominant systematic uncertainties used in the likelihood model for the statistical inference of the signal. The expression "MC" refers to all processes that are modelled using simulation, the expression "F_F" refers to all backgrounds that are estimated with the F_F method and "EMB" refers to all backgrounds estimated with the embedding method. Detailed descriptions are given in the text. The first block contains uncertainties which only affect the yield while uncertainties in the second block may also affect the shape of the discriminant distributions. Adapted from Ref. [12].
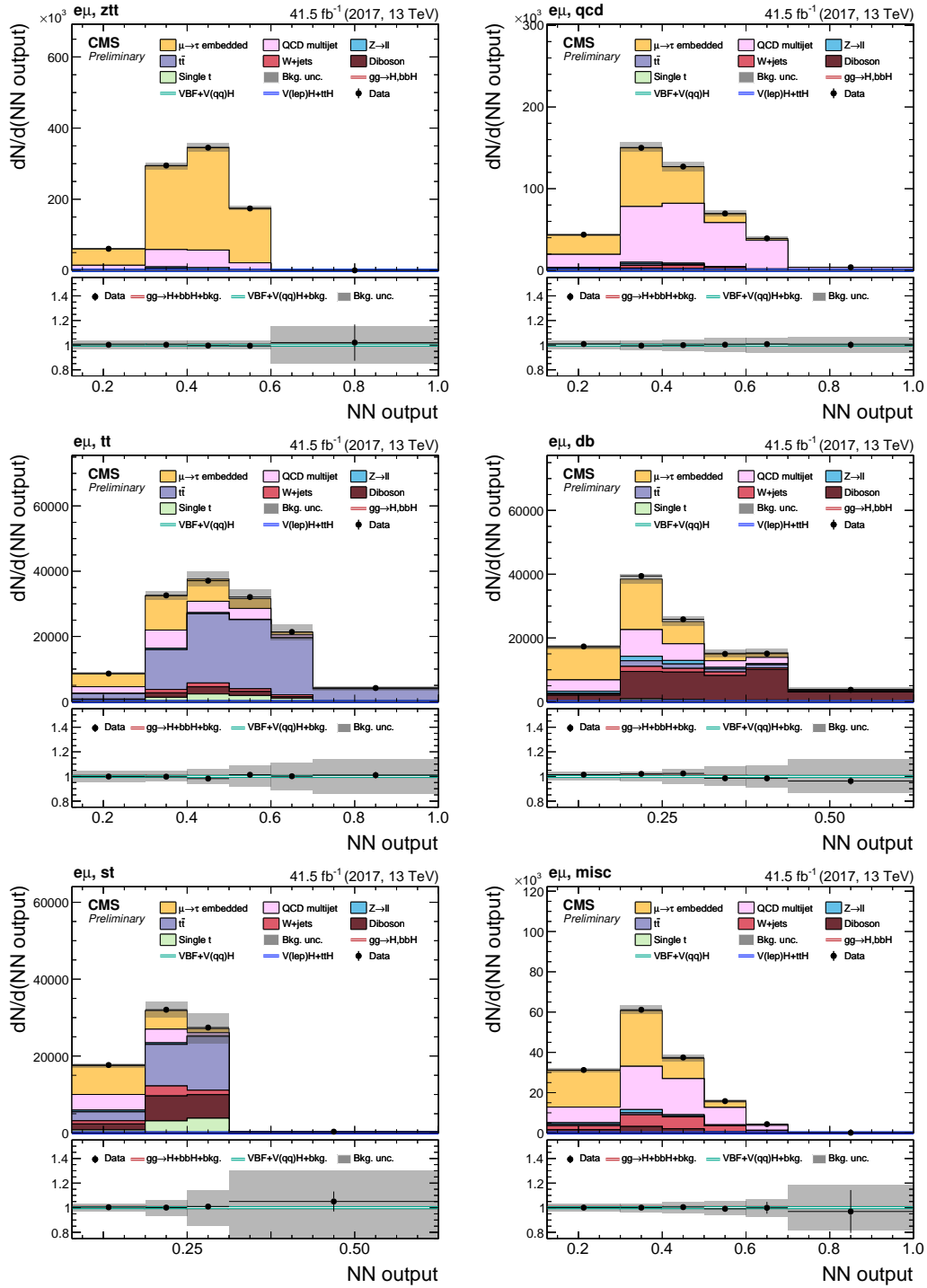
Figure 5.21 – Distributions of the NN output in the `ztt` (top left), `qcd` (top right), `tt` (centre left), `db` (centre right), `st` (bottom left), and `misc` (bottom right) event categories, in the eμ final state. All distributions and uncertainties are shown after the fit that has been applied to the data [12].
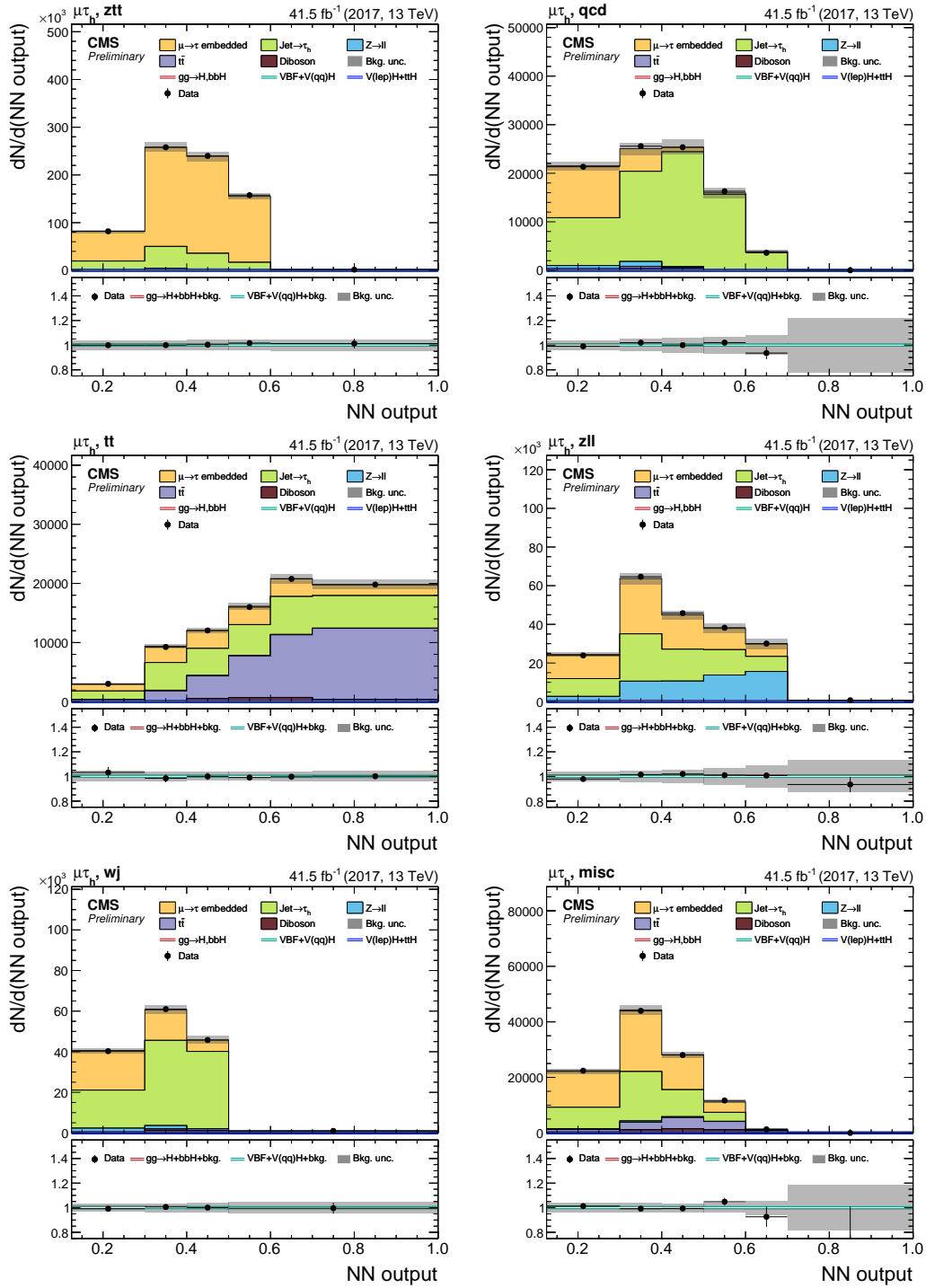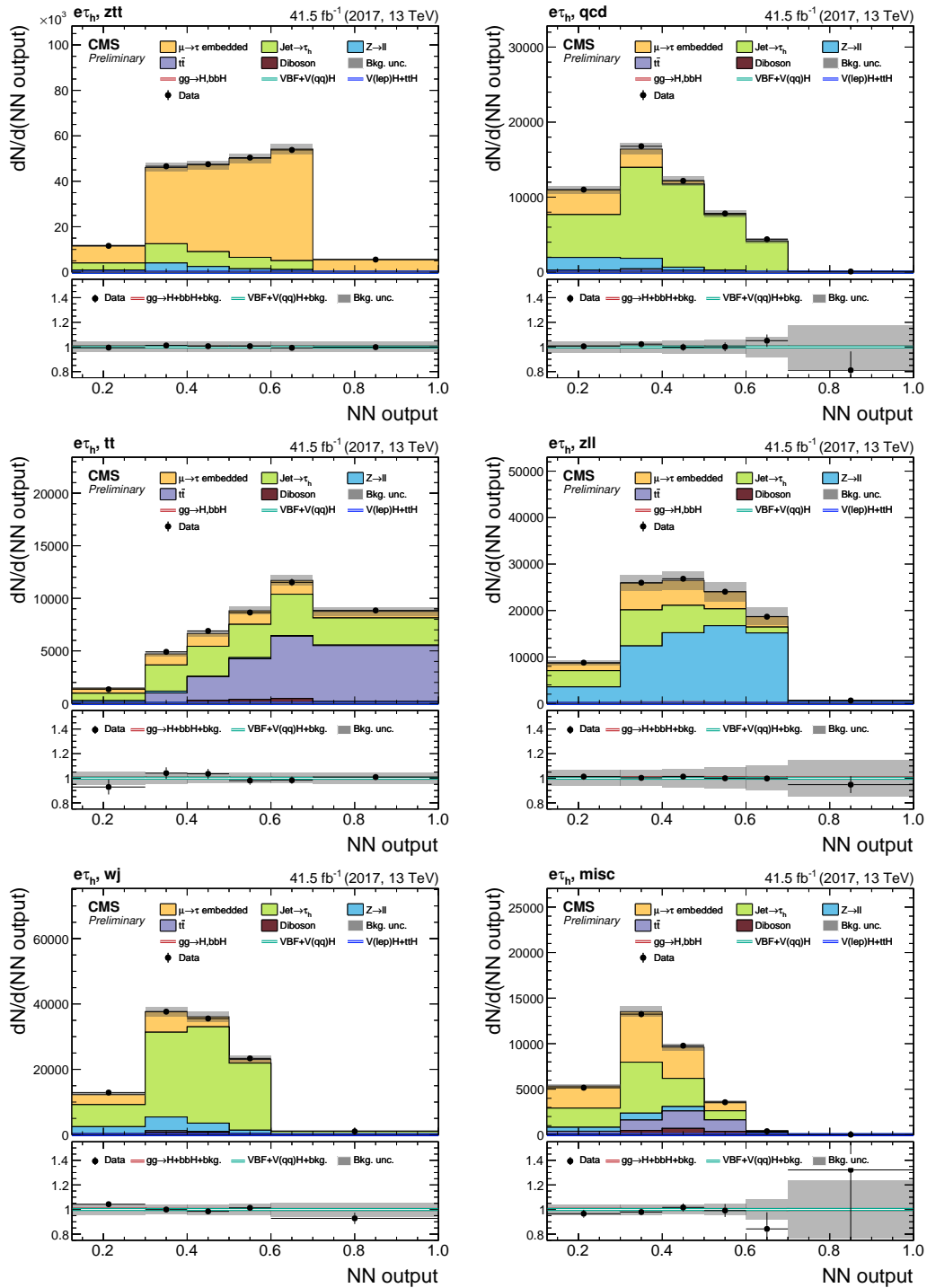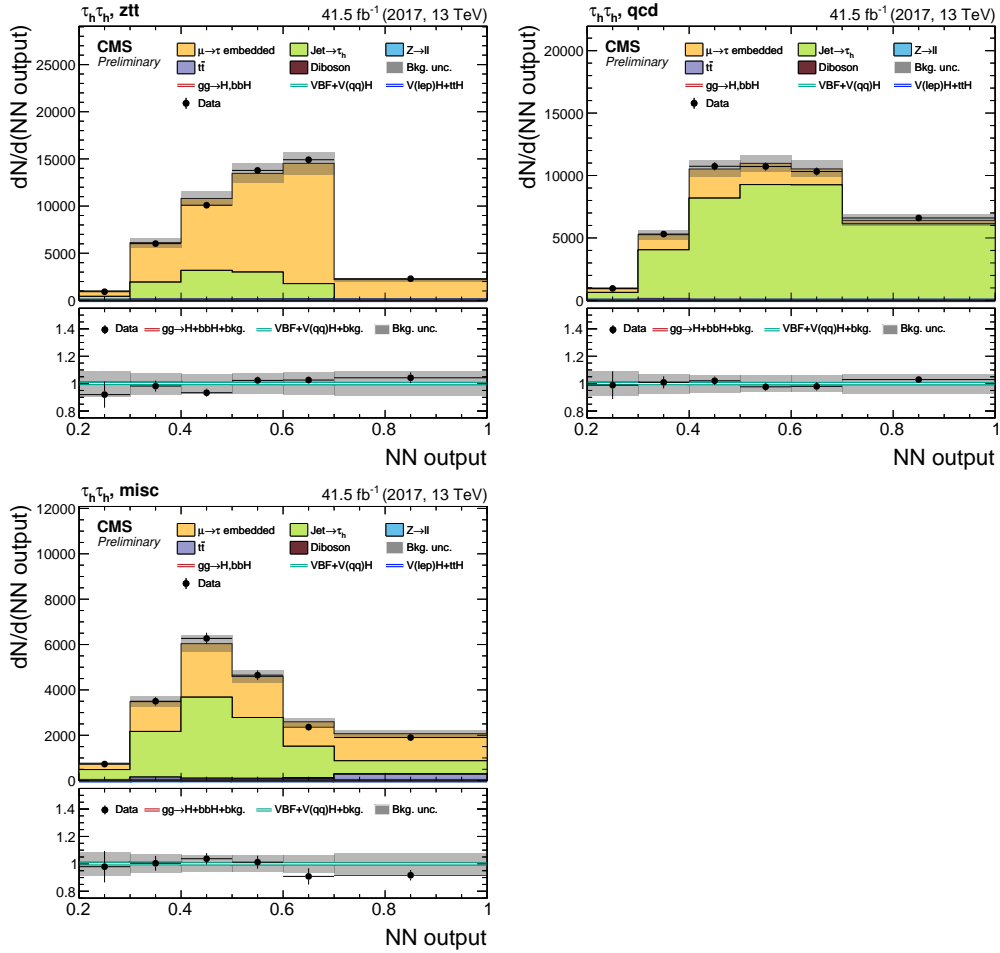
Figure 5.22 – Distributions of the NN output in the `ztt` (top left), `qcd` (top right), `tt` (centre left), `zll` (centre right), `wj` (bottom left), and `misc` (bottom right) event categories, in the $\mu\tau_h$ final state. All distributions and uncertainties are shown after the fit that has been applied to the data [12].

Figure 5.23 – Distributions of the NN output in the `ztt` (top left), `qcd` (top right), `tt` (centre left), `zll` (centre right), `wj` (bottom left), and `misc` (bottom right) event categories, in the eτ_h final state. All distributions and uncertainties are shown after the fit that has been applied to the data [12].

Figure 5.24 – Distributions of the NN output in the `ztt` (top left), `qcd` (top right), `misc` (bottom left) event categories, in the $\tau_h\tau_h$ final state. All distributions and uncertainties are shown after the fit that has been applied to the data [12].

For the statistical inference of the signal three models, each with a different number of parameters of interest (POIs), are considered, as discussed in Section 5.5.2. In the first model the Higgs boson cross section is measured inclusively (one single POI). In the second model the $gg \rightarrow H$, bbH and VBF+V(qq)H production modes are measured separately in a simultaneous fit (two POIs). For this model the NN output distributions for the 2017 dataset for the ggH and qqH categories for all final states are shown in Figs. 5.25 and 5.26. The NN output distributions for the 2016 dataset for each signal category can be found in Appendix A.7. In the third model the Higgs boson production is measured in the context of the STXS classification scheme described in Section 5.5.2 (in total 9 POIs). This classification scheme leads to 12–14 signal categories per final state. The distributions of the NN output split into the stage-1 signal categories used as input to the fit are shown in Appendix A.8. The same NN trainings are used as for the more inclusive measurements and the NN output is fitted in all signal categories. Contributions from ttH and V($\rightarrow$ lep)H are small and therefore included as background with the cross section set to the SM expectation.

The likelihood function used to extract the signal has the following form:

$$\mathcal{L}\Big(\{k_i\}\big|\mu_\alpha s_\alpha(\theta) + b_\beta(\theta)\Big) = \prod_{i,\alpha,\beta} \mathcal{P}\Big(k_i\big|\mu_\alpha s_{\alpha,i}(\theta) + b_{\beta,i}(\theta)\Big) \prod_j \mathcal{C}_j\Big(\hat{\theta}_j\big|\theta_j\Big), \quad (5.12)$$

where the function $\mathcal{P}$ corresponds to a Poisson distribution, $\mathcal{C}_j$ to the probability density function used to implement the uncertainty related to the nuisance parameter θ_j, and $\hat{\theta}_j$ to the estimate for θ_j from the fit to data. The function $\mathcal{C}_j$ is chosen to be a lognormal distribution for all normalization uncertainties and a vertical morphing algorithm [141] is used to implement shape altering uncertainties. The i labels all bins of the input distributions with event number k_i in all final states and categories and j all nuisance parameters θ. All signal classes are labelled as α and all background classes as β. The term $b_{\beta,i}$ corresponds to the prediction for a single background class in bin i, and μ_α to a scaling parameter for a given signal $s_{\alpha,i}$. The scaling parameters μ_α correspond to the POIs listed above.

5.10 Results

Assuming an overall scaling of all signal production modes with respect to the SM, except for ttH and V($\rightarrow$ lep)H, the fit with one single POI leads to a value of $\mu_{\mathrm{incl}} = 0.75 \pm 0.18$, which is 1.4σ below 1. Within the statistical precision of the measurement, it is consistent with the expectation of the SM, with a p-value of
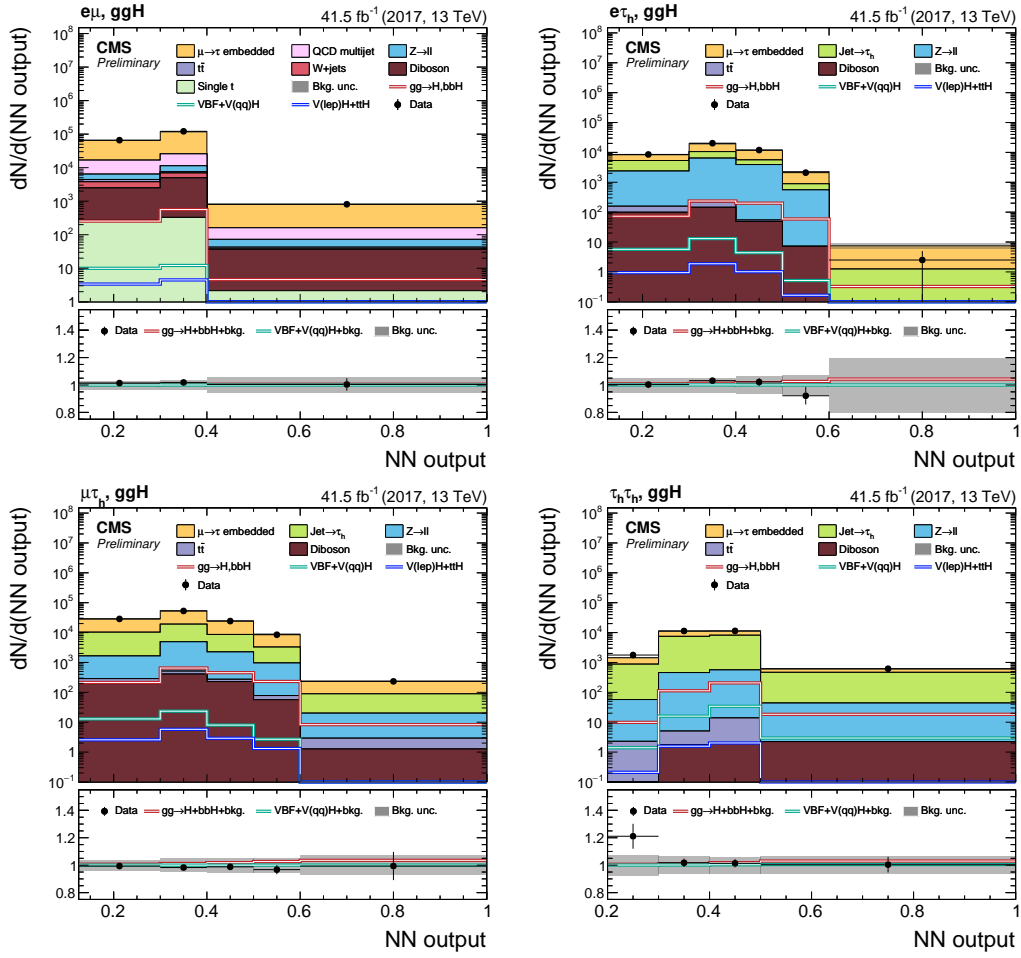
Figure 5.25 – Distributions of the NN output in the **ggH** event category, in the eμ (top left), eτ_h (top right), $\mu\tau_\mathrm{h}$ (bottom left), and $\tau_\mathrm{h}\tau_\mathrm{h}$ (bottom right) final states, on the 2017 dataset. All distributions and uncertainties are shown after the fit that has been applied to the data [12].
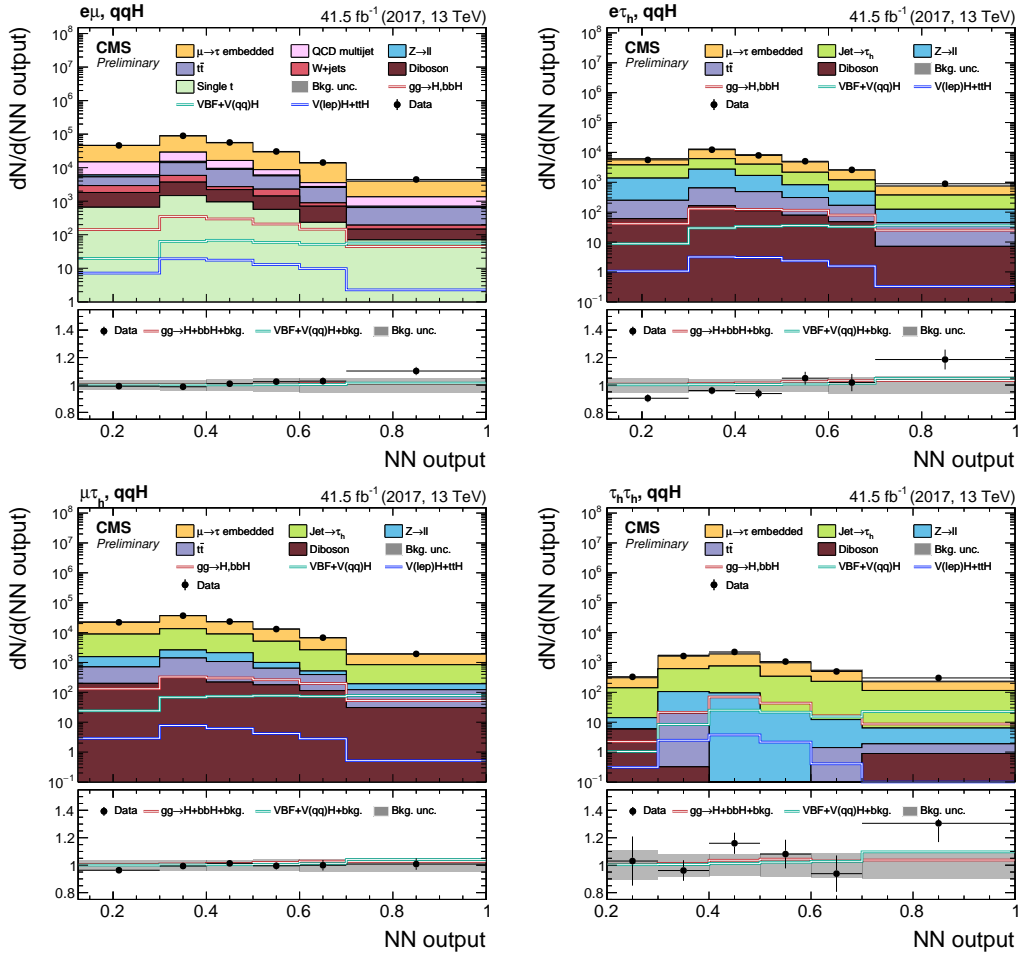
Figure 5.26 – Distributions of the NN output in the qqH event category, in the eμ (top left), eτ_h (top right), $\mu\tau_\mathrm{h}$ (bottom left), and $\tau_\mathrm{h}\tau_\mathrm{h}$ (bottom right) final states. All distributions and uncertainties are shown after the fit that has been applied to the data [12].

11 %. The observed signal strength, when restricted to the 2016 dataset, has been cross checked with the result of the previous H $\to \tau\tau$ analysis given in Ref. [10]. The previous analysis was optimized for the inclusive significance and used cut-based event categories and 1- or 2-dimensional discriminators. Only little overlap has been observed between the sensitive regions of this and the previous analysis. Both results are statistically compatible within about 1.5σ and the differences can be explained by statistical fluctuations.
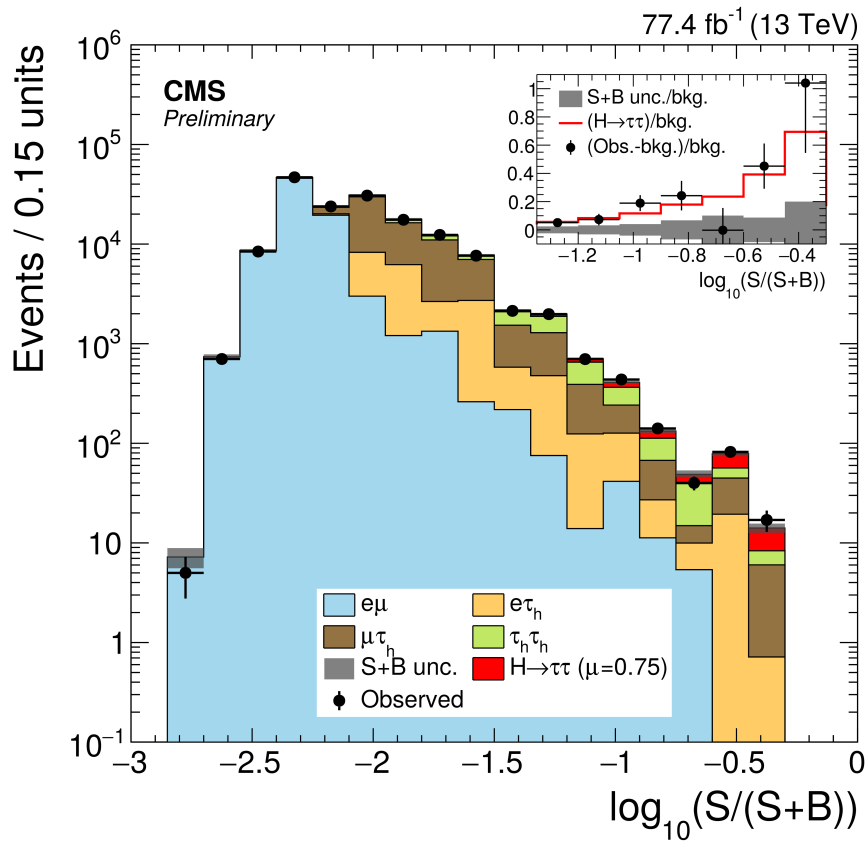

Figure 5.27 – Distribution of the signal purity (after the fit to the data), evaluated in the bins of the signal categories used to extract the results. The signal contribution and the background contributions from each decay channel are stacked on top of each other. The inset shows the relative excess with respect to the estimated background [12].

Figure 5.27 shows the observed Higgs boson signal, with the bins of all signal categories reordered according to their signal purity. The background contributions are split by final state and the expected signal contribution is scaled to the measured signal strength of $\mu_{\mathrm{incl}} = 0.75$. An excess of observed events with respect to

the background expectation is visible in the most sensitive bins and the ratio is shown in the inset on the top right. The observed (expected) significance of the signal is found to be 4.7σ (6.6σ).

The observed signal in the $e\mu$, $e\tau_h$, $\mu\tau_h$ and $\tau_h\tau_h$ final states as well as the inclusive measurement is shown in Fig. 5.28. The results obtained in the $e\mu$ and $\tau_h\tau_h$ final states are compatible with the SM expectation within 1σ. The $e\tau_h$ and $\mu\tau_h$ final states are compatible with the SM expectation within about 2σ.
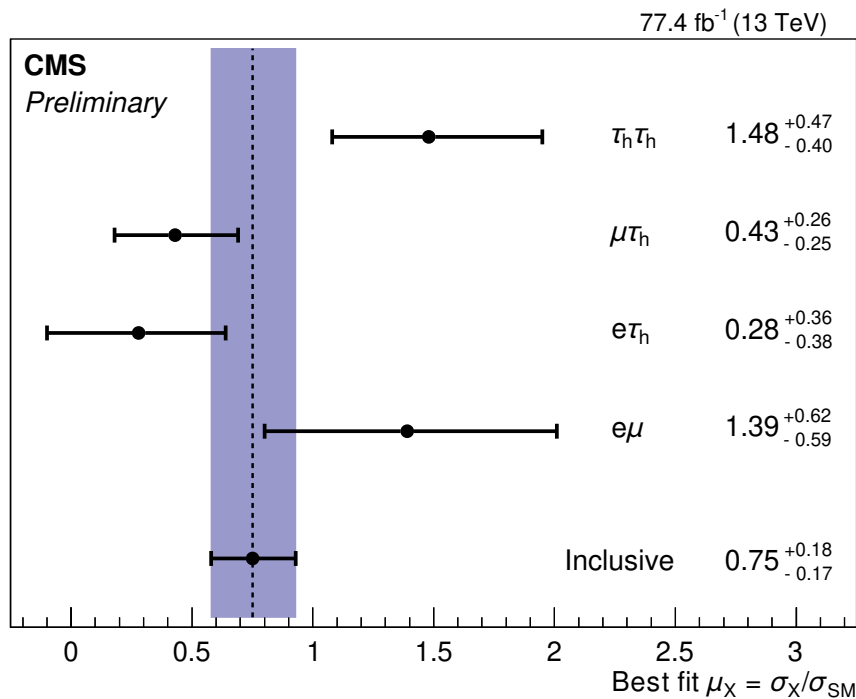
Figure 5.28 – Results of the maximum likelihood fit to the data, for a model with one signal strength parameter, as discussed in Section 5.10, measured inclusively and split by final state. The ttH and $V(\rightarrow \mathrm{lep})H$ processes have a negligible effect on the shown results and are not included as signal [12].

For the statistical model with two POIs the signal strength by production mode is measured and a two-dimensional scan of the likelihood is performed. This results in $\mu(gg \rightarrow H, bbH) = 0.36 \pm 0.37$ and $\mu(\mathrm{VBF}+V(qq)H) = 1.03 \pm 0.30$ with a linear correlation of -0.44 between these two POIs. In Table 5.8 the result of the fit for one single POI as well as for two POIs in terms of expected and observed cross sections are given. The result of the fit with two POIs shows that the measurements are in good agreement with the SM expectation, however, for $gg \rightarrow H, bbH$ a low signal yield is observed. The theoretical uncertainties on the

$\sigma\,\mathcal{B}(\mathrm{H} \to \tau\tau)$ (pb)	Expected	Observed
Inclusive	$3.40 \pm 0.20\,(\mathrm{theo})$	$2.56 \pm 0.48\,(\mathrm{stat}) \pm 0.34\,(\mathrm{syst})$
$\mathrm{gg} \to \mathrm{H}, \mathrm{bbH}$	$3.07 \pm 0.19\,(\mathrm{theo})$	$1.11 \pm 0.81\,(\mathrm{stat}) \pm 0.78\,(\mathrm{syst})$
$\mathrm{VBF}+\mathrm{V}(qq)\mathrm{H}$	$0.33 \pm 0.01\,(\mathrm{theo})$	$0.34 \pm 0.08\,(\mathrm{stat}) \pm 0.09\,(\mathrm{syst})$

Signal strength μ_{X}	
Inclusive	$0.75 \pm 0.14\,(\mathrm{stat}) \pm 0.10\,(\mathrm{syst})$
$\mathrm{gg} \to \mathrm{H}, \mathrm{bbH}$	$0.36 \pm 0.26\,(\mathrm{stat}) \pm 0.25\,(\mathrm{syst})$
$\mathrm{VBF}+\mathrm{V}(qq)\mathrm{H}$	$1.03 \pm 0.26\,(\mathrm{stat}) \pm 0.14\,(\mathrm{syst})$

Table 5.8 – Observed and expected cross sections for inclusive H boson production and for the production via the $\mathrm{gg} \to \mathrm{H}, \mathrm{bbH}$ and $\mathrm{VBF}+\mathrm{V}(qq)\mathrm{H}$ processes. The measured linear correlation between the $\mathrm{gg} \to \mathrm{H}, \mathrm{bbH}$ and the $\mathrm{VBF}+\mathrm{V}(qq)\mathrm{H}$ cross sections is -0.44. The uncertainties on the expected cross sections are discussed in the text [12].

expected cross sections in Table 5.8 include uncertainties on the parton distribution functions and α_S, and variations by the renormalization and factorization scales by factors 0.5 and 2 for each process added in quadrature. For each individual process the uncertainties have been weighted according to their combined expected cross section contribution in this calculation. The resulting fractions are 71 % of VBF on the expected $\mathrm{VBF}+\mathrm{V}(qq)\mathrm{H}$ cross section and 99 % of ggH on the expected $\mathrm{gg} \to \mathrm{H}, \mathrm{bbH}$ cross section.

The two-dimensional scan of the likelihood function is shown in Fig. 5.29 (left). Additionally, a two-dimensional scan of the likelihood function in terms of the coupling modifiers κ_{V} and κ_{F} is shown in Fig. 5.29 (right) for which the contribution of $\mathrm{H} \to \mathrm{WW}$ to the $\mathrm{e}\mu$ final state and the ttH and $\mathrm{V}(\to \mathrm{lep})\mathrm{H}$ processes are treated as a signal processes. The 68 % and 95 % CL of the SM expectation are shown by the dashed contour lines and the observed contours correspond to the shaded areas. The scan of the likelihood function as well as the scan in terms of coupling modifiers agree within 2σ with the SM expectation. For the contour in terms of coupling modifiers the contribution of events from the Higgs boson decay into two W bosons ($\mathrm{H} \to \mathrm{WW}$) to the $\mathrm{e}\mu$ final state, as well as the ttH and $\mathrm{V}(\to \mathrm{lep})\mathrm{H}$ processes are also treated as signal.
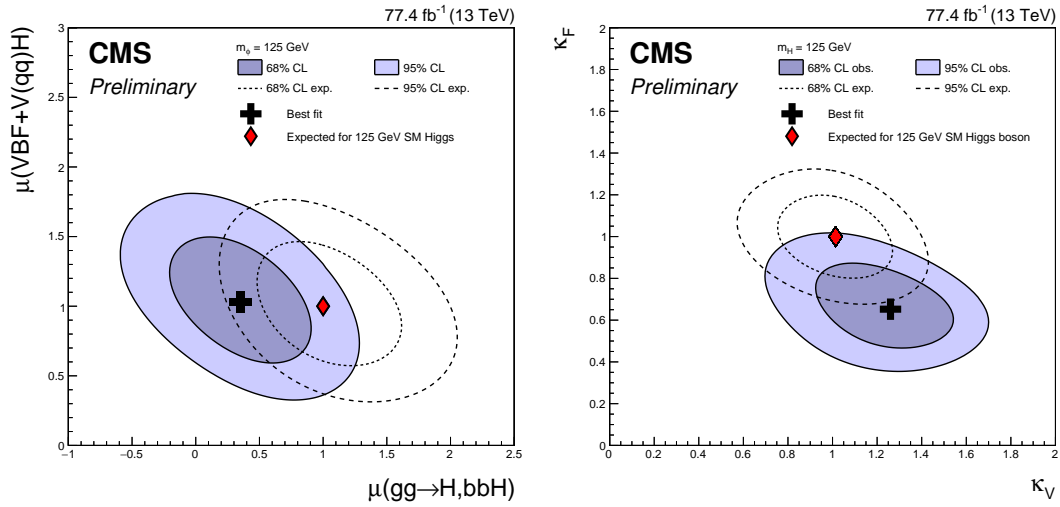
Figure 5.29 – Two-dimensional contours corresponding to the 68% and 95% confidence levels (CL) for the signal strength of the gg → H, bbH and VBF+V(qq)H production modes (left) and for the coupling strength modifiers κ_{V} and κ_{F} (right). All nuisance parameters are fit at each point evaluated in the two-dimensional parameter space. For the scan of κ_{V} and κ_{F}, the contribution of H → WW to the eμ final state and the ttH and V(→ lep)H processes are treated as a signal processes [12].

Figure 5.30 summarizes the results for the statistical model with 9 POIs. The signal strength parameters of VBF+V(qq)H are shown at the bottom of Figure 5.30. The "VBF topology" is the most sensitive POI and in general there is good agreement with the SM expectation. At the top of Figure 5.30 the signal strength parameters of gg → H, bbH are shown. The low observed signal yield of the inclusive gg → H, bbH measurement can be associated to the kinematic regions with 0 or 1 jet and low Higgs boson p_{T}.

The correlation matrix for the fit with 9 POIs is shown in Fig. 5.31. The "≥ 2 Jet" shows clear anti-correlation with the "Rest", "$p_{\mathrm{T}}^{\mathrm{j1}} > 200\,\mathrm{GeV}$" and "V($qq$)H Topology" subcategories, which for their part are correlated.
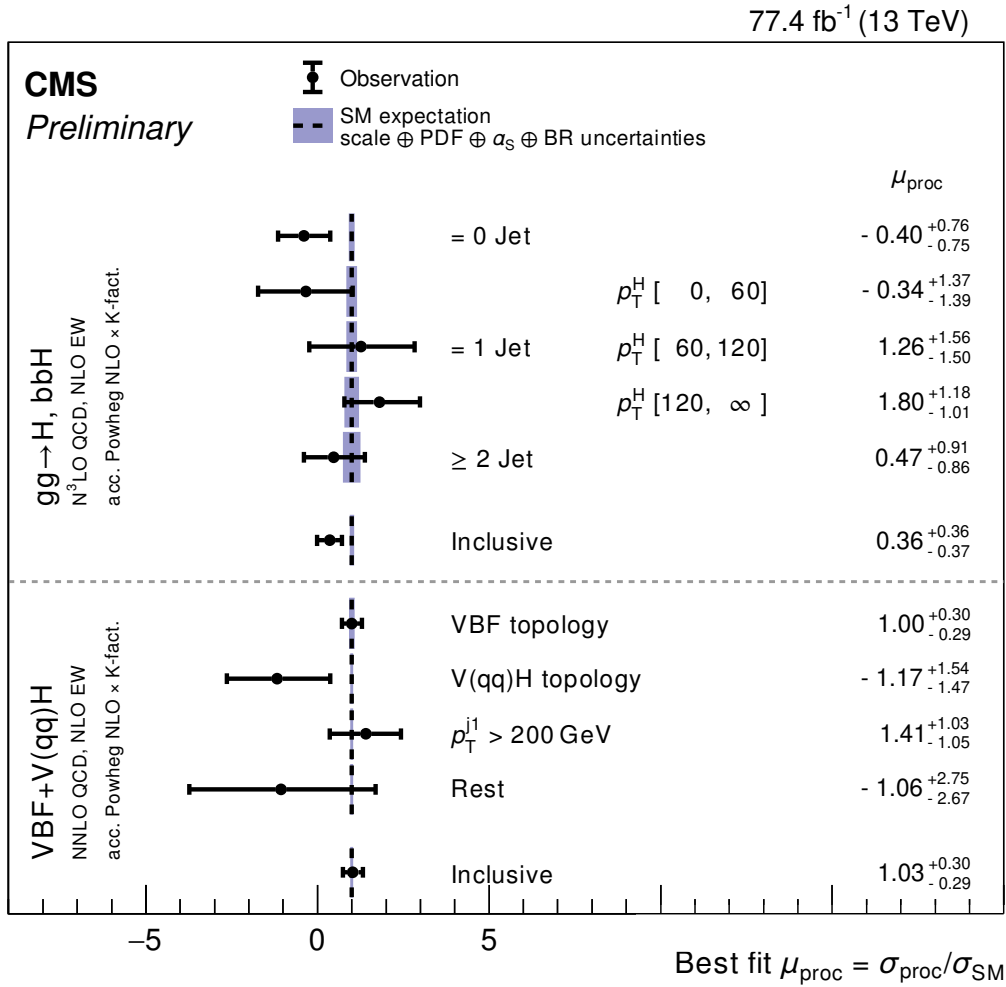
Figure 5.30 – Results of the maximum likelihood fit to the data, for a model with 9 signal strength parameters, as discussed in Section 5.10. Also shown are the results from an inclusive fit with separate signal strength parameters for the $gg \rightarrow H$, bbH and VBF+V(qq)H production processes. The ttH and VH production modes with leptonic V decays have a negligible effect on the shown results and are not included as signal [12].
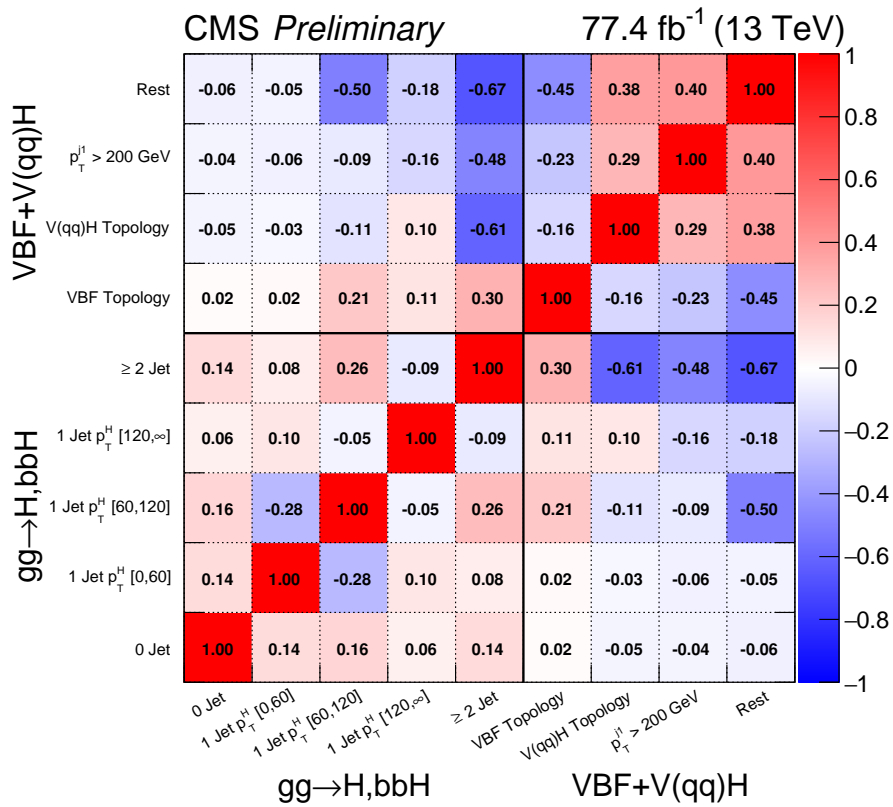
Figure 5.31 – Correlation matrix of the maximum likelihood fit with 9 signal strength parameters, as discussed in Section 5.10. The corresponding fit results are shown in Figure 5.30 [12].

SEARCH FOR ADDITIONAL NEUTRAL MSSM HIGGS BOSONS IN THE $\tau\tau$ FINAL STATE

*"When bad habits become a habit,
you have to turn over a new leaf."*

– THE 13 $\frac{1}{2}$ LIVES OF CAPTAIN BLUEBEAR

In this chapter, a search for additional neutral Higgs bosons in the $\tau\tau$ final state is presented. The measurement is based on proton-proton collisions recorded in 2016 by CMS at a centre-of-mass energy of $13\,\text{TeV}$, corresponding to an integrated luminosity of $35.9\,\text{fb}^{-1}$. In this search four different $\tau\tau$ final states, $\text{e}\mu$, $\text{e}\tau_\text{h}$, $\mu\tau_\text{h}$ and $\tau_\text{h}\tau_\text{h}$, are studied.

The author performed the analysis in the $\text{e}\tau_\text{h}$, $\mu\tau_\text{h}$ and $\tau_\text{h}\tau_\text{h}$ final states and estimated the backgrounds arising from quark or gluon jets that are misinterpreted as hadronic τ lepton decays with the fake factor method. The presented work has been published in Ref. [13].

This analysis has been performed before the measurement presented in Chapter 5. As a consequence, differences, especially concerning the fake factor method, do not describe improvements but outline the differences of the implementation that was used for the search of additional neutral MSSM Higgs bosons with respect to the description in Chapter 5.

6.1 Introduction

As described in Section 2.2, the Higgs sector in the MSSM contains two Higgs doublets, which leads to five physical Higgs bosons. At leading order, the couplings of the heavy scalar H and the pseudoscalar A to down-type fermions are

enhanced by $\tan\beta$ with respect to the SM Higgs boson. The higher experimental sensitivity of the Higgs boson decay to τ leptons with respect to the decay to b quarks and higher rate compared to the decay to muons renders this decay mode the most suitable for the search for additional Higgs bosons at intermediate and high values of $\tan\beta$.

Searches for additional heavy neutral Higgs bosons in the context of the MSSM have been conducted in e^+e^- collisions at LEP [147] and proton-antiproton collisions at the Tevatron [148–151]. The ATLAS and CMS collaborations performed searches for the Higgs boson decay to b quarks [152, 153], to muons [154, 155] and to a pair of τ leptons [154, 156–161] using data collected during LHC Run 1. Figure 6.1 shows a summary of selected analyses carried out by CMS in the MSSM $m_\mathrm{h}^\mathrm{mod+}$ scenario and the hMSSM scenario. Searches where an MSSM Higgs boson decays into a pair of τ leptons cover the largest area in the $\tan\beta$–m_A phase space. In the $m_\mathrm{h}^\mathrm{mod+}$ scenario at low values of m_A the $\mathrm{H}^\pm \to \tau\nu$ search excludes all $\tan\beta$ values. For the hMSSM the low mass region is excluded by the 125 GeV Higgs boson measurements.
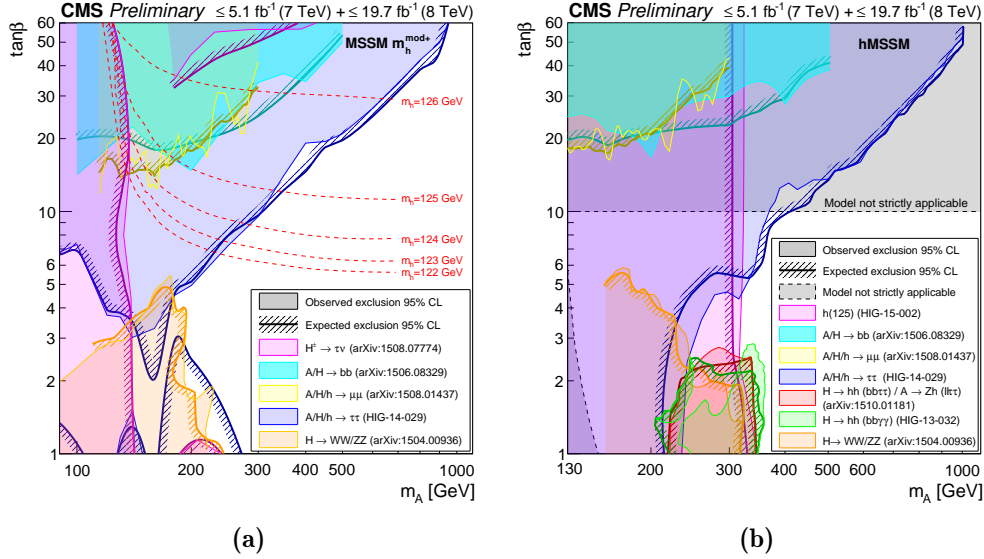
Figure 6.1 – Summary of selected CMS analyses that have been performed on LHC data collected until the end of 2012 in the MSSM $m_\mathrm{h}^\mathrm{mod+}$ scenario (a) and hMSSM (b) [162].

The analysis presented in this chapter is conducted at a centre-of-mass energy of 13 TeV, which enables the search to probe higher Higgs boson masses compared to the searches performed at 7 and 8 TeV during Run 1.

In Section 6.3 the event selection in the respective $H \to \tau\tau$ final states is introduced. The categorization used to increase the sensitivity of the search and the methods used for background estimation are described in Sections 6.4 and 6.5. The data-to-simulation corrections, the uncertainty model and the procedure to extract a signal are introduced in Sections 6.6–6.8. Finally, in Section 6.9, the results of the measurements are presented.

6.2 Dataset and Samples

Several MC event generators are used to generate signal and background events. The simulation of the ggF signal process is done at LO precision using PYTHIA 8.212 [163]. The modelling of the Higgs boson p_T spectrum is improved by reweighting it to NLO precision using POWHEG [106–111]. This is done by separately calculating the p_T spectra for the contributions from the t quark, b quark and the tb interference term to the initial loop, as discussed in [164–166], and combining the individual contributions according to their contribution to the total cross section for a given set of model parameters in the interpretation step. The production of a Higgs boson in association with b quarks is simulated using MADGRAPH5_aMC@NLO [167] at NLO precision. The production of an Higgs boson with 125 GeV via gluon fusion, vector boson fusion (VBF), or in association with a vector boson (VH, with V either Z or W) is simulated using POWHEG. Single t and $t\bar{t}$ background samples are generated at NLO precision using POWHEG 2.0. The diboson production processes are simulated using MADGRAPH5_aMC@NLO [112, 113] at NLO precision. W+jets events and events from Drell-Yan in the ee, $\mu\mu$ and $\tau\tau$ final state are generated at LO precision with MADGRAPH5_aMC@NLO. To increase the number of simulated events in regions of high signal purity additional samples with up to four outgoing jets are generated for the W+jets and Drell-Yan processes.

The NNPDF 3.0 [115] parton density functions with the underlaying event tune CUETP8M1 [116] are used for the generation of all signal and background processes. Hadronic showering and hadronization, τ lepton decays as well as additional interactions according to the expected pile-up distribution are modelled using PYTHIA [114]. All generated events are passed through a simulation of the CMS detector based on GEANT4 [83]. Drell-Yan, W+jets, $t\bar{t}$ and single t events in the tW-channel are normalized to their respective cross-section calculated at NNLO precision [117–119]. Diboson events and single t events in the t-channel are normalized to their cross section at NLO precision or higher [119–121].

6.3 Event Selection

A similar strategy as described in Section 5.3 is used for the event selection from data recorded during the 2016 data-taking period by CMS and utilized in this analysis. The purpose of the selection is to identify events in the $e\mu$, $e\tau_h$, $\mu\tau_h$ and $\tau_h\tau_h$ final states. This includes the vetoes introduced to reduce events with extra electrons and muons, and events from $Z/\gamma^* \to ee$ and $Z/\gamma^* \to \mu\mu$. The specific selection criteria for the individual leptons in each final state are described in the following.

6.3.1 Event Selection in the $e\mu$ Final State

In the $e\mu$ final state events are selected using two trigger items that require the presence of an electron and a muon. The electron is required to have $p_T > 13\,\text{GeV}$ and $|\eta| < 2.5$ and the muon $p_T > 10\,\text{GeV}$ and $|\eta| < 2.4$. To avoid events in the lower part of the trigger turn-on curve being selected when the event only passes one trigger the higher-p_T trigger object is required to have $p_T > 24\,\text{GeV}$. Both leptons are required to be associated to the PV within $d_{xy} < 0.045\,\text{cm}$ in the transverse plane and $d_z < 0.2\,\text{cm}$ along the beam axis. Furthermore, the electron and muon are required to be separated by $\Delta R > 0.3$, have opposite charge and need to be isolated according to $I_{\text{rel}}^{e(\mu)} < 0.15\,(0.2)$. The working points of the electron MVA identification discriminator used in this analysis have an identification efficiency between $80\,\%$ and $90\,\%$. For muons the *medium* identification criteria [122] is used.

6.3.2 Event Selection in the $e\tau_h$ and $\mu\tau_h$ Final State

In the semi-leptonic channels events are selected by using a single-lepton trigger which requires the presence of one electron (muon) in the $e\tau_h$ ($\mu\tau_h$) final state.

The τ_h candidate is required to have $p_T > 30\,\text{GeV}$ and $|\eta| < 2.3$. It needs to be reconstructed by the HPS algorithm and pass the `Tight` working point of the τ_h MVA identification discriminant. Furthermore, it is required to pass the `Tight` (`VeryLoose`) working point of the anti-electron discriminator and the `Loose` (`Tight`) working point of the anti-muon discriminator in the $e\tau_h$ ($\mu\tau_h$) final state. Finally, the τ_h candidate needs to be associated to the PV along the beam axis within $d_z < 0.2\,\text{cm}$.

In the $\mu\tau_h$ channel the muon and the τ_h need to be separated by $\Delta R > 0.5$ and

have opposite charge. The muon is required to have $p_T > 23\,\mathrm{GeV}$ and $|\eta| < 2.1$. Furthermore, it should pass the medium identification requirements and needs to be associated to the PV within $d_{xy} < 0.045\,\mathrm{cm}$ in the transverse plane and $d_z < 0.2\,\mathrm{cm}$ along the beam axis. It is also required to be isolated according to $I_{\mathrm{rel}}^{\mu} < 0.15$ as defined in Section 4.2.3 using a cone size of $\Delta R = 0.4$.

Like the muon in the $\mu\tau_h$ channel the electron in the $e\tau_h$ channel is required to be separated from the τ_h by $\Delta R > 0.5$ and needs to have opposite charge. The p_T threshold is set to $p_T > 26\,\mathrm{GeV}$ and the electron is required to be within $|\eta| < 2.1$. The electron needs to pass the $80\,\%$ working point of the electron MVA identification with an isolation of $I_{\mathrm{rel}}^{e} < 0.1$ within a cone size of $\Delta R = 0.3$. Finally, it needs to be associated to the PV within $d_{xy} < 0.045\,\mathrm{cm}$ and $d_z < 0.2\,\mathrm{cm}$.

6.3.3 Event Selection in the $\tau_h\tau_h$ Final State

The trigger items used to select the events in the $\tau_h\tau_h$ final state require the presence of two τ_h candidates. Both candidates are required to have $p_T > 40\,\mathrm{GeV}$, $|\eta| < 2.1$ and should pass the `Medium` working point of the τ_h MVA identification discriminant. The τ_h candidates need to be separated by $\Delta R > 0.5$, be of opposite charge and need to be associated to the PV within $d_z < 0.2\,\mathrm{cm}$. To suppress the misidentification of electrons and muons as hadronic τ decays the anti-electron and anti-muon discriminants are applied with the `VeryLoose` and `Loose` working points, respectively.

6.4 Categorization

All events that satisfy the selection described in Section 6.3 are categorized to increase the sensitivity of the analysis. Events with at least one b-tagged jet with $p_T > 20\,\mathrm{GeV}$ and $|\eta| < 2.4$ are combined into a global *b-tag* category, targeting the Higgs boson production in association with b quarks. All other events are combined into a global *no b-tag* category, targeting the production via gluon fusion. Events in the $e\mu$ final state are split in three subcategories, as shown in Table 6.1, based on the D_ζ variable [123]. The D_ζ distribution in the $e\mu$ chan-

Low-D_ζ	Medium-D_ζ	High-D_ζ
$-50\,\mathrm{GeV} < D_\zeta < -10\,\mathrm{GeV}$	$-10\,\mathrm{GeV} < D_\zeta < 30\,\mathrm{GeV}$	$D_\zeta > 30\,\mathrm{GeV}$

**Table 6.1** – Categories based on the D_ζ variable used in the $e\mu$ channel.

nel with the definition of the subcategories is shown in Fig. 6.2a. The region in which most signal events are located for all masses tested is the Medium-D_ζ region.

The eτ_h and $\mu\tau_h$ final states are split in two categories based on the transverse mass m_T, as shown in Table 6.2. The m_T^μ distribution in the $\mu\tau_h$ channel is

Tight-m_T	Loose-m_T
$m_T^\ell < 40\,\text{GeV}$	$40\,\text{GeV} < m_T^\ell < 70\,\text{GeV}$

Table 6.2 – Categories based on the transverse mass m_T^ℓ, with $\ell = \text{e}, \mu$, used in the eτ_h and $\mu\tau_h$ channel.

shown in Fig. 6.2b. The bulk of the signal events, especially for the low-mass hypothesis, falls in the Tight-m_T region, while the Loose-m_T region increases the signal acceptance for the high-mass hypothesis. Events in the $\tau_h\tau_h$ channel are only split according to the number of b-tagged jets.
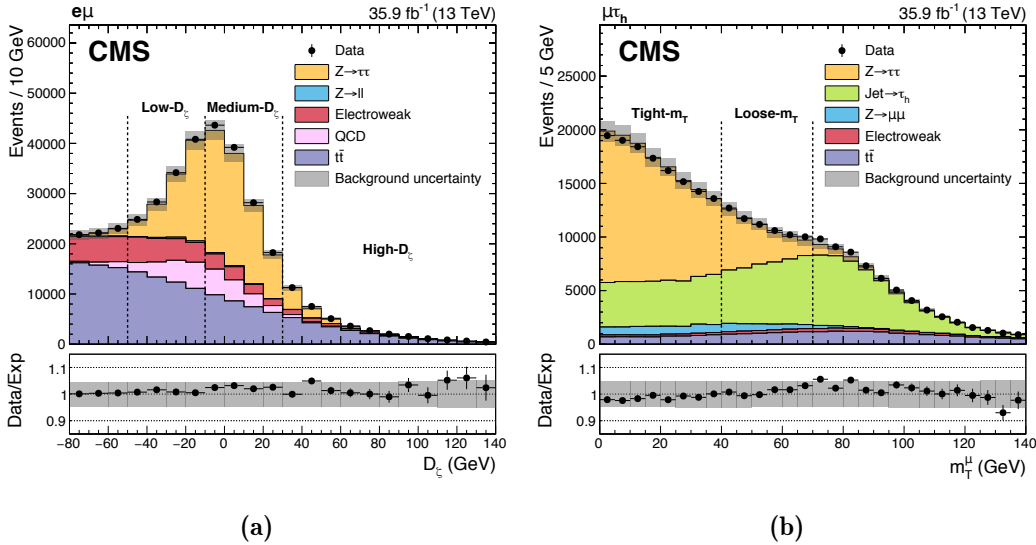
Figure 6.2 – Observed and expected distribution of D_ζ in the eμ (a) and m_T^μ in the $\mu\tau_h$ channel (b). The regions indicated with the dashed vertical lines correspond to the subcategorization as defined in Tables 6.1 and 6.2. The distributions are shown without displaying the shape altering uncertainties and before the fit for the signal extraction [13].

Fig. 6.3 shows the categorization scheme for all four final states as utilized for the statistical inference of the signal. In addition to the signal categories, two control regions, one enriched in t$\bar{\text{t}}$ events and one enriched with $\text{Z} \to \mu\mu$ events, are used to constrain the normalization of the t$\bar{\text{t}}$ and $\text{Z} \to \tau\tau$ backgrounds.

	No b-tag			b-tag		
$H \to \tau\tau \to e\mu$	Low-D_ζ	Medium-D_ζ	High-D_ζ	Low-D_ζ	Medium-D_ζ	High-D_ζ
$H \to \tau\tau \to e\tau_h$	Loose-m_T	Tight-m_T		Loose-m_T	Tight-m_T	
$H \to \tau\tau \to \mu\tau_h$	Loose-m_T	Tight-m_T		Loose-m_T	Tight-m_T	
$H \to \tau\tau \to \tau_h\tau_h$						
$Z \to \mu\mu$						
$t\bar{t}(e\mu)$						

☐ Signal region (SR)

◩ Control region

Figure 6.3 – Schematic overview of all subcategories that are used for the statistical inference of the signal. In total there are 16 signal categories complemented by three background categories [13].

6.5 Background Estimation

Table 6.3 shows a list of all background processes contributing to the event selection with the corresponding estimation technique given for each final state. In all four channels the $Z \to \tau\tau$ process is a dominant background since it decays to two genuine τ leptons and is, without sensitivity to CP-eigenvalues or spin in the analysis, only distinguishable from the signal via the difference in mass of the associated bosons. In the $e\tau_h$, $\mu\tau_h$ and $\tau_h\tau_h$ channels several backgrounds have in common that either a electron or muon is misidentified as τ_h candidate which is labelled as "$\ell \mapsto \tau_h$" with $\ell = e, \mu$. This includes events from $Z \to \ell\ell$, $t\bar{t}$, single t and diboson (WW, WZ and ZZ) production. Additionally, a fraction of events from $t\bar{t}$, single t and diboson production contains genuine hadronic τ decays and is labelled as "$\tau \mapsto \tau_h$". The backgrounds listed above are estimated using simulated events complemented with dedicated control regions to constrain the normalization of $Z \to \tau\tau$ and $t\bar{t}$ events. Another dominant source of backgrounds in the final states including a τ_h candidate is due to jets that are misidentified as τ_h candidates. The corresponding processes are labelled as jet $\mapsto \tau_h$ and are estimated using the F_F method. These are in particular events from W+jets and QCD multijet production but also events from $t\bar{t}$ and diboson production. The implementation of the F_F method that is used in this analysis is described in Section 6.5.1. In the $e\mu$ channel the shape and normalization from QCD multijet production is estimated from a CR in data, as described in Section 6.5.2.

Background process	Misidentification	$e\mu$	$e\tau_h$	$\mu\tau_h$	$\tau_h\tau_h$
H $\to \tau\tau$ (SM)		MC	MC	MC	MC
Z $\to \tau\tau$		MC†	MC†	MC†	MC†
W+jets		MC	F_F	F_F	F_F
QCD multijet production		CR	F_F	F_F	F_F
Z $\to \ell\ell$	Jet $\mapsto \tau_h$	MC	F_F	F_F	F_F
	$\ell \mapsto \tau_h$		MC	MC	MC
Diboson+single t †	Jet $\mapsto \tau_h$	MC	F_F	F_F	F_F
	$\tau/\ell \mapsto \tau_h$		MC	MC	MC
$t\bar{t}$ †	Jet $\mapsto \tau_h$	MC†	F_F	F_F	F_F
	$\tau/\ell \mapsto \tau_h$		MC†	MC†	MC†

† Normalization from a control region in data.

Table 6.3 – Background processes contributing to the event selection labelled with the corresponding estimation method for each final state.. Processes labelled with "MC" are estimated from simulation. The label "F_F" implies that the process is determined from data using the fake factor method. The label "CR" implies that both the shape and normalization of QCD multijet events are estimated from control regions in data. Adapted from Ref. [13].

6.5.1 Fake Factor Method

The idea and application of the F_F method used to estimate the backgrounds arising from jets that are misidentified as hadronic τ decay in the $e\tau_h$, $\mu\tau_h$ and $\tau_h\tau_h$ channels has been discussed thoroughly in Section 5.6.2. However, there are small differences in this analysis compared to the implementation used in Chapter 5. This concerns the binning of the fractions f_i and the binning of the raw F_F^i. Furthermore, there are minor differences in the definition of the DR_{QCD} as well as the corrections applied to the $F_F^{t\bar{t}}$.

Difference of the Fake Factor Measurement

The first difference is related to the binning of the measured F_F^i. As introduced in Section 5.6.2, the raw F_F^i are measured as a smooth function of the τ_h transverse momentum and as a function of jet multiplicity. In this analysis, the F_F^i are also measured as a function of the τ_h decay mode (1-prong and 3-prong).

A second difference is the definition of the $\mathrm{DR}_{\mathrm{QCD}}$ in the $\mathrm{e}\tau_{\mathrm{h}}$ and $\mu\tau_{\mathrm{h}}$ final states for the F_{F}^{i} measurement . In this region, an additional requirement on the transverse mass m_{T}^{ℓ} is applied. For the sake of completeness, all $\mathrm{DR}_{\mathrm{QCD}}$ requirements that differ from the selection in the SR, as defined in Section 6.3.2, are listed:

- The electric charges of the τ_{h} candidate and the light lepton are required to be the same.

- The relative lepton isolation has to be in the range $0.05 < I_{\mathrm{rel}}^{\ell} < 0.15$.

- Events are required to have a transverse mass $m_{\mathrm{T}}^{\ell} < 40\,\mathrm{GeV}$.

Difference of the Fake Factor Corrections

All corrections are derived using the same procedure as described in Section 5.6.2. A minor difference is that the data-to-simulation scale factors for the $F_{\mathrm{F}}^{\mathrm{t\bar{t}}}$ are not measured inclusively but as a function of the τ_{h} decay mode. The corresponding numbers are given in Table 6.4.

2016	1-prong	3-prong
$\mathrm{e}\tau_{\mathrm{h}}$	0.68	0.82
$\mu\tau_{\mathrm{h}}$	0.81	0.74

Table 6.4 − Data-to-simulation ratio for the $\mathrm{t\bar{t}}$ fake factors in $\mathrm{e}\tau_{\mathrm{h}}$ and $\mu\tau_{\mathrm{h}}$.

Difference of the Measurement of the Background Composition

Another change concerns the event selection applied in the AR in the $\tau_{\mathrm{h}}\tau_{\mathrm{h}}$ channel. Since the `Medium` working point of the τ_{h} ID discriminant is applied in the SR events are required to pass the `VeryLoose` but fail the `Medium` working point to be selected in the AR.

The procedure to measure the background composition in the AR and the application are identical to the instructions in Section 5.6.2. The difference is that the obtained fractions are measured as a function of the τ_{h} transverse momentum and the τ_{h} decay mode.

6.5.2 Estimation of QCD Multijet Production in $e\mu$

The QCD multijet background in the $e\mu$ final state is estimated from a control region (CR) with the same selection as in the signal region (SR) but requiring the electron and muon to have the same electric charge. The yield in the SR is estimated by applying same-to-opposite charge transfer factors on data in the CR, from which any non-QCD contribution is subtracted using simulation.

The transfer factors are measured in a sideband region (SB) without event categorization, where both leptons satisfy a relaxed isolation requirement of $I_{\text{rel}}^{\ell} < 0.4$ and at least one lepton does not satisfy the isolation requirement in the SR but $0.15 < I_{\text{rel}}^{e} < 0.4$ or $0.2 < I_{\text{rel}}^{\mu} < 0.4$ instead. They are measured as a function of the p_{T} of the two leptons and their spatial separation $\Delta R(e, \mu)$. Corrections are applied to account for differences between the transfer factors in the exclusive event categories and the inclusively measured factors in the SB. Finally, a correction is applied to account for differences in the SR and SB.

6.5.3 Estimation of Backgrounds from Simulation

All backgrounds that are not estimated with the methods described in Sections 6.5.1 and 6.5.2 are estimated via simulation. To improve the data-to-simulation agreement several corrections are applied on the individual processes. Furthermore, $Z \rightarrow \mu\mu$ and $t\bar{t}$ control regions are used to improve the modelling of $Z \rightarrow \tau\tau$ and $t\bar{t}$, respectively.

The uncertainty on the $Z \rightarrow \tau\tau$ normalization is constrained by including a $Z \rightarrow \mu\mu$ control region (CR) in the maximum likelihood fit. To correctly account for normalization effects of systematic uncertainties the $Z \rightarrow \mu\mu$ yield is correlated with the yield of $Z \rightarrow \tau\tau$. For this purpose, a similar event selection as discussed in Section 6.3.2 is used, however, the most isolated $\mu\mu$ pair is selected. The muon with the highest transverse momentum is required to satisfy $p_{\text{T}}^{\mu} > 24\,\text{GeV}$ and the other $p_{\text{T}}^{\mu} > 10\,\text{GeV}$. Furthermore, the invariant mass of the di-muon system is required to be within the range $70\,\text{GeV} < m_{\text{vis}} < 110\,\text{GeV}$. To align the CR with the categorization in the SR the selected events are further divided in a b-tag and no b-tag category. This selection leads to a purity of $Z \rightarrow \mu\mu$ events of about $88\,\%$ in the b-tag and $99\,\%$ in the no b-tag categories.

The $t\bar{t}$ normalization in all categories of the $e\mu$ channel as well as the normalization of the $t\bar{t}$ decays to genuine τ leptons in the remaining channels are correlated

with the $t\bar{t}$ normalization in this CR. This may be exploited to constrain the normalization of $t\bar{t}$ in all final states. This CR contains $e\mu$ events that pass the event selection outlined in Section 6.3.1 and two additional requirements of $D_\zeta < -50\,\text{GeV}$ and $p_\text{T}^\text{miss} > 80\,\text{GeV}$. This selection leads to a high purity region in which 85 % of the selected events originate from $t\bar{t}$.

6.6 Data-to-Simulation Corrections

The corrections to improve the agreement between data and simulation are to a large extent identical to the corrections applied on the 2016 dataset listed in Section 5.7. Not applied in this analysis are the corrections due to prefiring and the reweighting on the SM Higgs boson produced via gluon fusion.

The correction to the τ_h energy scale for events in which an electron is misidentified as τ_h is derived specifically for this analysis and differs from the one listed in Table 5.6. Furthermore, no correction has been derived for $\mu \mapsto \tau_\text{h}$. All corrections to the τ_h energy scale applied in this analysis are shown in Table 6.5.

Corrections for 2016 in %	1-prong	1-prong $+\,1\,\pi^0$	3-prong
$\tau \to \tau_\text{h}$	-0.5 ± 1.2	1.1 ± 1.2	0.6 ± 1.2
$e \mapsto \tau_\text{h}$	2.4 ± 0.5	7.6 ± 1.0	—
$\mu \mapsto \tau_\text{h}$	—	—	—

Table 6.5 – Corrections to the τ_h energy scale. The correction is derived for genuine τ_h decays, and events where an electron or muon is misidentified as τ_h candidate.

6.7 Uncertainty Model

The uncertainty model used for the statistical inference of the signal comprises theoretical, experimental and statistical uncertainties. Systematic uncertainties are implemented in the form of nuisance parameters in the likelihood [141], which can be further constrained by the fit to the data. A summary table of the dominant systematic uncertainties is given in Table 6.6.

6.7.1 Normalization Uncertainties

Luminosity
The uncertainties in the integrated luminosity amounts to 2.5 % [142].

Background cross section

An uncertainty in the background cross section of the W+jets, single t and diboson production contribution estimated from simulation is included. No theoretical cross section uncertainty is included for $Z \to \mu\mu$ and $t\bar{t}$ due to the inclusion of the respective control regions.

Electron and Muon identification, isolation and trigger efficiency

The uncertainties in the measurement of the identification, isolation and trigger efficiencies for both electrons and muons amount to $2\,\%$, adding all individual contributions in quadrature.

τ_h identification and trigger efficiency

The uncertainties for the τ_h identification efficiency is factorized into a correlated and uncorrelated part across the final states. It amounts to $4(8)\,\%$ in the correlated and $3(6)\,\%$ in the uncorrelated part for the $e\tau_\mathrm{h}$ and $\mu\tau_\mathrm{h}$ ($\tau_\mathrm{h}\tau_\mathrm{h}$) channel. A $7\,\%$ uncertainty is added to the uncorrelated part in the $\tau_\mathrm{h}\tau_\mathrm{h}$ final state to account for the uncertainty in the trigger efficiency measurement.

$e \mapsto \tau_\mathrm{h}$ and $\mu \mapsto \tau_\mathrm{h}$ misidentification rate

Uncertainties are included in the $e\tau_\mathrm{h}$, $\mu\tau_\mathrm{h}$ and $\tau_\mathrm{h}\tau_\mathrm{h}$ final states for the corresponding simulated $Z \to \ell\ell$ events in which one light lepton is misidentified as τ_h candidate. The uncertainty amounts to $11(3)\,\%$ in the $e\tau_\mathrm{h}$ ($\tau_\mathrm{h}\tau_\mathrm{h}$) channel for the $e \mapsto \tau_\mathrm{h}$ misidentification rate and $12(5)\,\%$ in the $\mu\tau_\mathrm{h}$ ($\tau_\mathrm{h}\tau_\mathrm{h}$) channel for the $\mu \mapsto \tau_\mathrm{h}$ misidentification rate.

Response and resolution of the $p_\mathrm{T}^\mathrm{miss}$

Uncertainties are included for $Z \to \ell\ell$ and W+jets production as well as all signal processes, and they are derived as part of the determination of the recoil corrections. They range from $1\,\%$ to $5\,\%$. For backgrounds from simulation for which no recoil corrections are included, jet energy scale and unclustered energy scale variations are propagated to the $p_\mathrm{T}^\mathrm{miss}$.

Jet energy scale and b tagging efficiency

An uncertainty in the jet energy scale is included for all processes estimated from simulation and affects the number of events entering each category. It ranges from $1\,\%$ to $6\,\%$ depending on the final state and subcategory. Similarly, uncertainties ranging from $1\,\%$ to $5\,\%$ are included on the efficiency of the b tagging discriminator.

QCD sideband extrapolation

An overall uncertainty of 30% due to the extrapolation from the control region to the signal regions is included. Additionally, category-dependent uncertainties in the measurement of the transfer factors are included, ranging from 4% to 29%

Signal acceptance of b-associated production

Uncertainties due to the renormalization and factorization scales, and of the internal generator matching scale Q_{sh} related to parton showering to account for the acceptance of signal events produced in association with b quarks [11].

PDF uncertainty

Uncertainties due to different choices of the factorization and renormalization scales in the signal predictions are included in the parameter scan of the model interpretations. They are calculated based on the choice of the parton distribution functions following the recommendation in Refs. [168, 169].

SM Higgs production cross section

Uncertainties on the calculation of the production cross section of the SM Higgs boson due to different choices of the factorization and renormalization scales for the calculation.

6.7.2 Shape Uncertainties

τ_{h} energy scale

Independent uncertainties, each amounting to 1.2%, are included for the 1-prong, 1-prong$+\pi^0$ and 3-prong decay modes of the τ_{h} candidate. A similar uncertainty of 0.5% to 1%, depending on the τ_{h} decay mode, is included for the $\mathrm{Z} \rightarrow \mathrm{ee}$ process in the $\mathrm{e}\tau_{\mathrm{h}}$ channel where an electron is misidentified as τ_{h}.

Electron energy scale

An uncertainty on the electron energy scale is included in the $\mathrm{e}\mu$ final state. The uncertainty is 1% in the barrel region and 2.5% in the endcap.

Top quark p_{T} reweighting

The maximum shape variations on the reweighting procedure, as described in Section 5.7, are obtained by not applying the weighting or by applying it twice.

Drell-Yan shape reweighting

In total five uncertainties are included to address the uncertainty in the Drell-Yan shape reweighting, which is discussed in Section 5.7. The uncertainties are comprised of three non-negligible statistical uncertainties when deriving the correction, one uncertainty due to the muon energy scale and one uncertainty in the $t\bar{t}$ cross section.

Extrapolation of the τ_h identification efficiency to the high-p_T regime

The τ_h identification efficiency is derived close to the Z boson mass peak in a low-p_T τ_h regime. To account for the extrapolation to the high-p_T regime a p_T dependent uncertainty with $\pm 1\sigma$ variations of $5\% \times p_T$ [TeV] and $-35\% \times p_T$ [TeV] is included.

jet $\mapsto$ e and jet $\mapsto$ μ misidentification rate

In the eμ channel uncertainties depending on the jet p_T with a minimum of $13(10)\%$ are included to account for the uncertainty on jet $\mapsto$ e (jet $\mapsto$ μ) misidentification rates.

6.7.3 Fake-Factor-specific Uncertainties

Statistical uncertainties on the fake factor

Statistical uncertainties in the measurement of the F_F^i in the DR_i obtained from the uncertainty of the fit. They are uncorrelated except for the uncertainty on $F_F^{t\bar{t}}$, which is correlated between the eτ_h and $\mu\tau_h$ final state. They amount to 4% in the $\mu\tau_h$ channel and range between 4%–7% in the eτ_h and between 2%–3% in the $\tau_h\tau_h$ channel.

Corrections to the QCD multijet fake factor

The following uncertainties are added quadratically, leading to one additional nuisance parameter per final state:

- Uncertainty of the non-closure correction, as a function of m_{vis}.

- Uncertainty on the correction function for I_{rel}^ℓ in the eτ_h and $\mu\tau_h$ channel or on the p_T of the other τ_h candidate in the $\tau_h\tau_h$ channel.

- Uncertainty on the correction function for the extrapolation from the same-to-opposite charge region.

Corrections to the W+jets fake factor

The following uncertainties are added quadratically, leading to one additional nuisance parameter in the eτ_h and $\mu\tau_h$ channel:

- Uncertainty of the non-closure correction, as a function of m_{vis}.

- Uncertainty on the correction depending on the transverse mass.

Corrections to the $t\bar{t}$ fake factor

The following uncertainties are added quadratically, leading to one additional nuisance parameter in the $e\tau_{\text{h}}$ and $\mu\tau_{\text{h}}$ channel:

- Uncertainty of the non-closure correction, as a function of m_{vis}.

- Uncertainty on the data-to-simulation correction factors

Background composition in the AR

Systematic uncertainty on the fraction of W/Z+jets and $t\bar{t}$ events with one misidentified τ_{h} candidate in the AR, adding two nuisance parameters for the $\tau_{\text{h}}\tau_{\text{h}}$ channel.

Subtraction of genuine τ background in the AR

Systematic uncertainty due to the subtraction of genuine τ backgrounds in the AR, which is correlated across categories.

6.8 Statistical inference of the Signal

The statistical inference of the signal is based on the total transverse mass distributions in the categories described in Section 6.4. The total transverse mass, $m_{\text{T}}^{\text{tot}}$ [158], is defined as

$$m_{\text{T}}^{\text{tot}} = \sqrt{m_{\text{T}}^2\left(\tau_1, \tau_2\right) + m_{\text{T}}^2\left(\tau_1, p_{\text{T}}^{\text{miss}}\right) + m_{\text{T}}^2\left(\tau_2, p_{\text{T}}^{\text{miss}}\right)}, \tag{6.1}$$

where the pair (τ_1, τ_2) can be (e, μ), (e, τ_{h}), (μ, τ_{h}) or $(\tau_{\text{h}}^1, \tau_{\text{h}}^2)$. The transverse mass m_{T} between two objects (i, j) is given by:

$$m_{\text{T}}(i, j) = \sqrt{2\, p_{\text{T}}^{\text{i}}\, p_{\text{T}}^{\text{j}}\left[1 - \cos\Delta\phi(i, j)\right]} \tag{6.2}$$

In Figs. 6.4 and 6.5 the distributions of $m_{\text{T}}^{\text{tot}}$ in the most sensitive subcategories for all four final states are shown. The distribution of $m_{\text{T}}^{\text{tot}}$ in the remaining subcategories can be found in Appendix B.1. The stack of filled histograms in the top panel of each subfigure represents the expectation of the background estimation model as described in Section 6.5. Each filled histogram corresponds to the estimated template distribution for a given process. The template labelled as "Jet $\to \tau_{\text{h}}$" comprises events from W+jets, $t\bar{t}$ and QCD multijet production, which are estimated using the F_{F} method. The remaining contributions from

Uncertainty	$e\mu$	$e\tau_h$	$\mu\tau_h$	$\tau_h\tau_h$	Process	Shape	Variation
Integrated luminosity	✓	✓	✓	✓	MC	—	2.5%
signal acceptance	✓	✓	✓	✓	Signal	—	3.2–16.5%
PDF/scale	✓	✓	✓	✓	Signal	—	15–25%
	✓	✓	✓	✓	SM Higgs	—	0.5–3.2%
Jet energy scale	✓	✓	✓	✓	MC	—	1–6%
b tagging	✓	✓	✓	✓	MC	—	1–5%
p_T^{miss} resp./res.	✓	✓	✓	✓	MC	—	1–5%
Bkgr. in signal categories	✓	✓	✓	✓	Diboson	—	5%
	✓	✓	✓	✓	single t	—	5%
	✓	—	—	—	W+jets	—	4%
Bkgr. in $DR_{QCD/W+jets}$	—	✓	✓	—	MC	—	3%
	—	—	—	✓	MC	—	4%
Sideband extrapolation	✓	✓	✓	✓	$Z \to \tau\tau$	—	1–7%
	✓	✓	✓	✓	$Z \to \ell\ell$	—	4%
	✓	✓	✓	✓	$t\bar{t}$	—	1%
	✓	—	—	—	QCD	—	4–29 (30)%
e/μ-trigger, ID, isolation	✓	✓	—	—	MC	—	2%
	✓	—	✓	—	MC	—	2%
$e \to \tau_h$ mis-ID	—	✓	—	—	$Z \to ee$	—	11%
	—	—	—	✓	$Z \to ee$	—	3%
$\mu \to \tau_h$ mis-ID	—	—	✓	—	$Z \to \mu\mu$	—	12%
	—	—	—	✓	$Z \to \mu\mu$	—	5%
τ_h-trigger	—	—	—	✓	MC	—	7%
τ_h-ID	—	✓	✓	—	MC	—	3 (4)%
	—	—	—	✓	MC	—	6 (8)%
τ_h-ID (high p_T)	—	✓	✓	✓	MC	✓	p_T dep.
τ_h energy scale	—	✓	✓	✓	MC	✓	1.2%
$e \to \tau_h$ energy scale	—	✓	—	—	$Z \to ee$	✓	0.5–1.0%
e energy scale	✓	—	—	—	MC	✓	1.0–2.5%
Jet $\to$ e mis-ID	✓	—	—	—	MC	✓	13%
Jet $\to \mu$ mis-ID	✓	—	—	—	MC	✓	10%
Top quark p_T reweighting	✓	✓	✓	✓	$t\bar{t}$	✓	100%
Z reweighting of LO MC	✓	✓	✓	✓	$Z \to \tau\tau, \ell\ell$	✓	See text
F_F^i stat. uncert.	—	✓	—	—	F_F	✓	4–7%
	—	—	✓	—	F_F	✓	4%
	—	—	—	✓	F_F	✓	2–3%
F_F^i corrections	—	✓	—	—	F_F	✓	7–10%
	—	—	✓	—	F_F	✓	5–7%
	—	—	—	✓	F_F	✓	10%

Table 6.6 – Overview of the systematic uncertainties used in the likelihood model for the statistical inference of the signal. The label "MC" refers to all processes that are obtained from simulation, the label "F_F" refers to all backgrounds that are obtained from the fake factor method. Values in parentheses correspond to additional uncertainties correlated across final states or event categories. Detailed descriptions are given in Section 6.7. Adapted from Ref. [13].
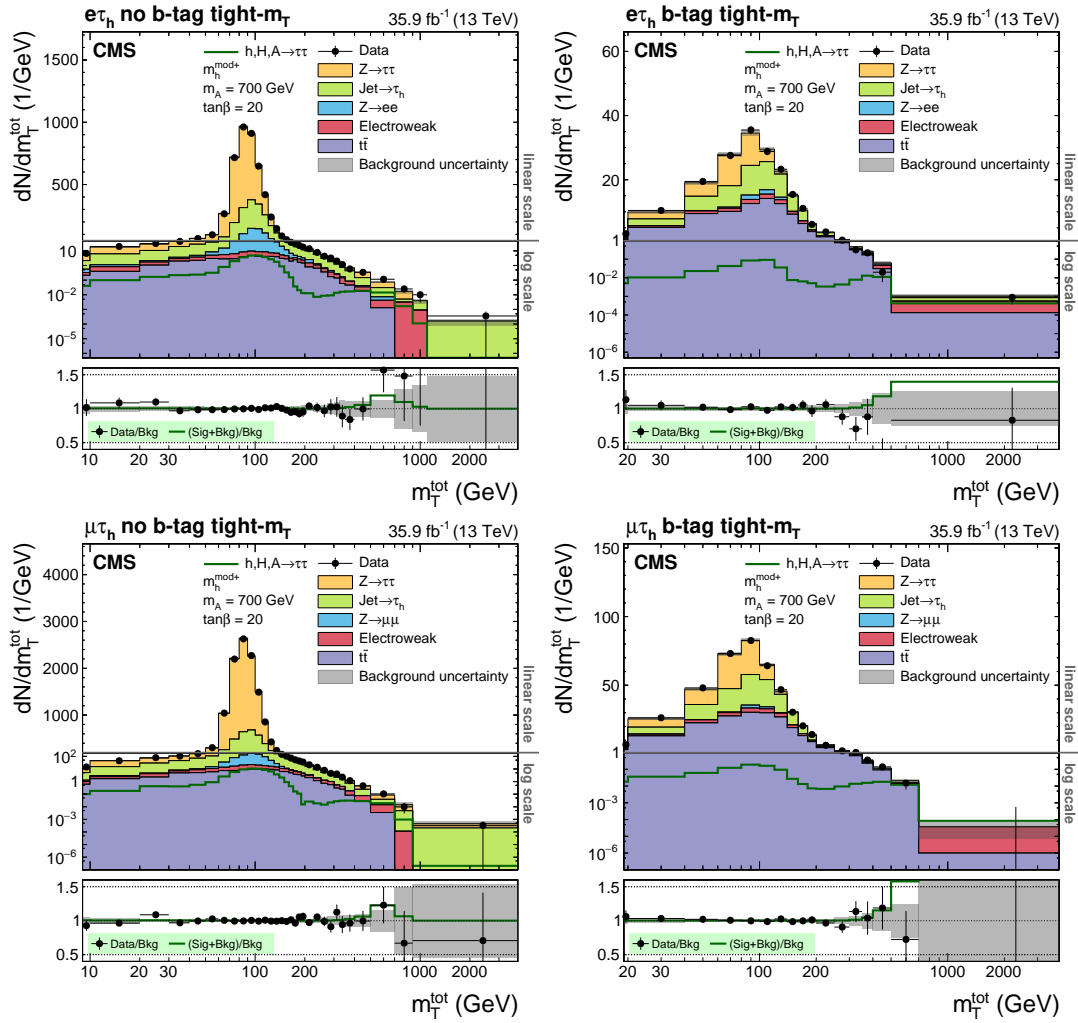
Figure 6.4 – Distribution of m_T^{tot} in the $e\tau_h$ (top row) and $\mu\tau_h$ (bottom row) final states. The distributions are shown after the fit to data. In all cases the most sensitive tight-m_T event subcategory is shown. The gray horizontal line in the top panel of each subfigure indicates the change from logarithmic to linear scale on the vertical axis [13].
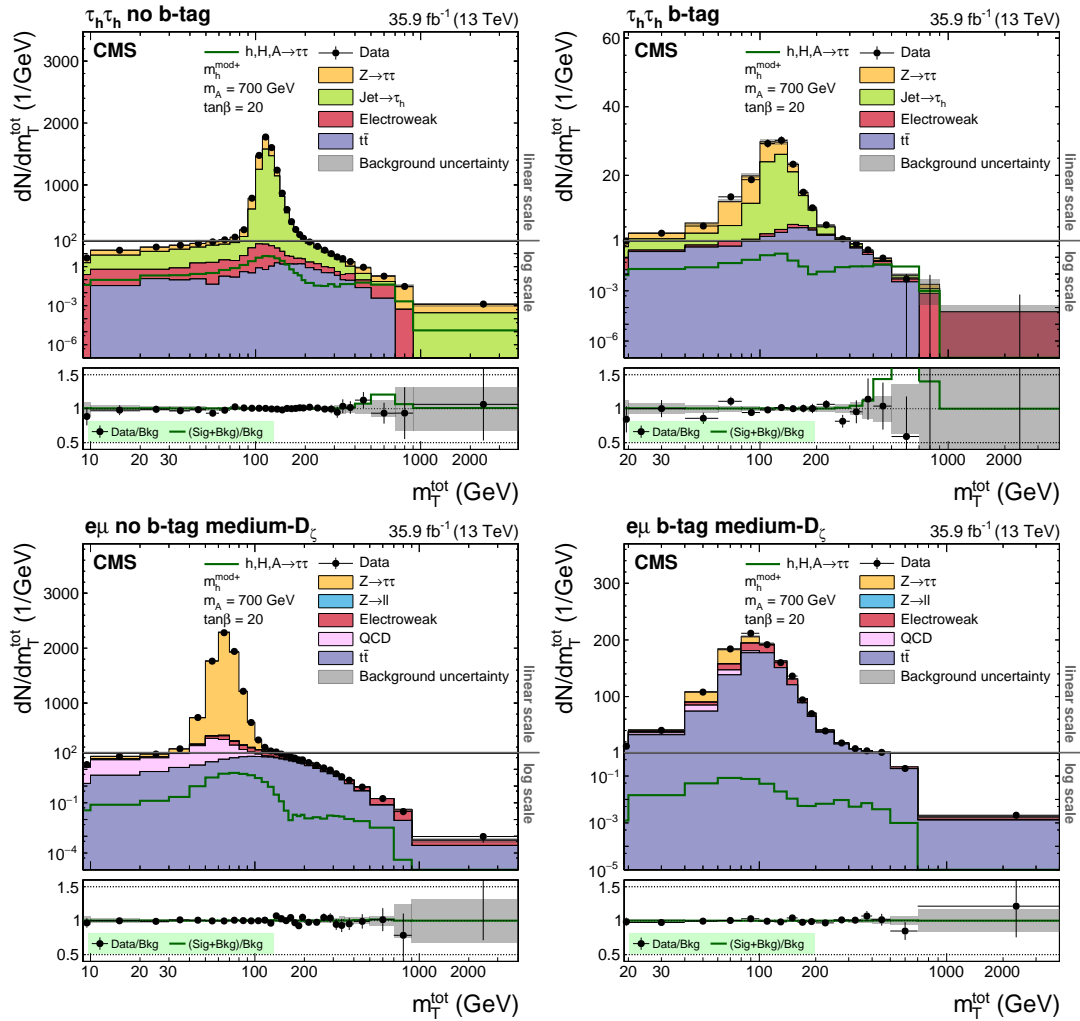
Figure 6.5 – Distribution of m_T^{tot} in the $\tau_h\tau_h$ (top row) and $e\mu$ (bottom row) final states. The distributions are shown after the fit to data. For the $e\mu$ final state the most sensitive medium-D_ζ event subcategory is shown. The gray horizontal line in the top panel of each subfigure indicates the change from logarithmic to linear scale on the vertical axis [13].

W+jets, single t and diboson production are subsumed in the template labelled as "Electroweak". The bottom panel shows the ratio of the observed number of events to the expectation from the background estimation. The shaded band in the bottom panel and associated to the stacked histograms in the top panel corresponds to the combination of all shape-altering and normalization uncertainties in all background processes. The vertical bars on the data points represent the statistical uncertainties associated to the observed number of events.

For the statistical inference of the signal the analysis makes use of the likelihood function that has been introduced in Eq. (5.12). A simultaneous maximum likelihood fit to the $m_{\mathrm{T}}^{\mathrm{tot}}$ distributions in all final states and subcategories is performed and the control regions designed to constrain the Drell-Yan and $t\bar{t}$ backgrounds are included in the likelihood model. The signal extraction is performed under the background-only and several signal-plus-background hypotheses to search for excesses due to potential additional Higgs bosons, with the 125 GeV Higgs boson included as background.

The data are interpreted in a model-independent way and also in terms of representative MSSM benchmark scenarios. The model-independent interpretation searches for a single resonance ϕ produced by gluon fusion (ggϕ) or in association with a b quark (bbϕ). For each production mode 28 single resonances with mass m_ϕ between 90 GeV and 3.2 TeV are simulated. In the second interpretation, exclusion contours in the m_{A}–$\tan\beta$ plane for the $m_{\mathrm{h}}^{\mathrm{mod+}}$ and hMSSM benchmark scenarios, as introduced in Section 2.2, are determined.

6.9 Results

The distributions in Figs. 6.4 and 6.5 are shown after an $m_{\mathrm{h}}^{\mathrm{mod+}}$ signal-plus-background hypothesis has been fitted to the data. The hypothesis corresponds to $m_{\mathrm{A}} = 700\,\mathrm{GeV}$ and $\tan\beta = 20$, with the expectation for a signal of three neutral Higgs bosons shown in the $m_{\mathrm{T}}^{\mathrm{tot}}$ distributions. No excess over the expected SM background has been found in the investigated mass range between 90 GeV and 3.2 TeV. Upper limits are set on the presence of a signal for the two discussed production modes following the modified frequentist approach [170, 171]. In this approach, a test statistic q_μ, which is based on the profile likelihood ratio defined

in Refs. [33, 172], is defined as:

$$q_\mu = -2\ln \frac{\mathcal{L}\Big(\{k_i\}\big|\mu s(\hat{\theta}_\mu) + b(\hat{\theta}_\mu)\Big)}{\mathcal{L}\Big(\{k_i\}\big|\hat{\mu}s(\hat{\theta}_{\hat{\mu}}) + b(\hat{\theta}_{\hat{\mu}})\Big)}, \quad \text{with } 0 \le \hat{\mu} \le \mu, \tag{6.3}$$

where the index of q_μ indicates that the fit to data has been performed for a fixed value of μ. The hat in $\hat{\mu}$, $\hat{\theta}_\mu$ and $\hat{\theta}_{\hat{\mu}}$ indicates the respective maximum likelihood best fit estimate. This test statistic is used to obtain 95 % CL upper limits [173] on the product of the Higgs boson production cross section and the branching fraction to τ leptons.

For the model-independent interpretation of the data, 95 % CL upper limits are set on the product of Higgs boson production cross section times branching fraction for the decay into a τ lepton pair. Separate limits, which are shown in Fig. 6.6, are set on the production of a single resonance ϕ via gluon fusion or in association with a b quark. The SM Higgs boson at 125 GeV is treated as background. Furthermore, as described in Section 6.2, the p_T spectrum of the ϕ is reweighted to NLO precision, which affects the sensitivity of the analysis at low masses where the Higgs boson p_T contributes significantly to the p_T of its decay products. The limit for each production mode is determined by treating the normalization of the corresponding other process as freely varying parameter in the fit that is performed before the limit calculation. The expected limits range between 18 pb at $m_\phi = 90$ GeV and 3.5 fb at $m_\phi = 3.2$ TeV for gluon fusion and between 15 pb at $m_\phi = 90$ GeV and 2.5 fb at $m_\phi = 3.2$ TeV for the Higgs boson production in association with b quarks. In addition to the presented 95 % CL upper limits for each production mode, a 2D likelihood scan as a function of the Higgs boson production via gluon fusion, $\sigma(\text{gg}\phi)B(\phi \to \tau\tau)$, and the production in association with b quarks, $\sigma(\text{bb}\phi)B(\phi \to \tau\tau)$, for the tested mass points is performed. Figure 6.7 shows the likelihood scan at a subset of six mass points.

The second interpretation of the data is in terms of exclusion contours for two representative benchmark scenarios of the MSSM, the $m_h^{\text{mod}+}$ and hMSSM, introduced in Section 2.2. For this purpose, the simulated neutral Higgs boson signals are combined into a multiresonance signal structure, comprised of the three neutral Higgs bosons (h,H,A), expected from each of the benchmark scenarios for the given values of m_A and $\tan\beta$. The data are fit for each value of m_A and $\tan\beta$ under the signal-plus-background and the background-only hypotheses. The utilized test statistic differs slightly from the one introduced in Eq. (6.3). A fixed value of $\mu = 1$ is used in the numerator and a fixed value of $\mu = 0$ in the denominator.
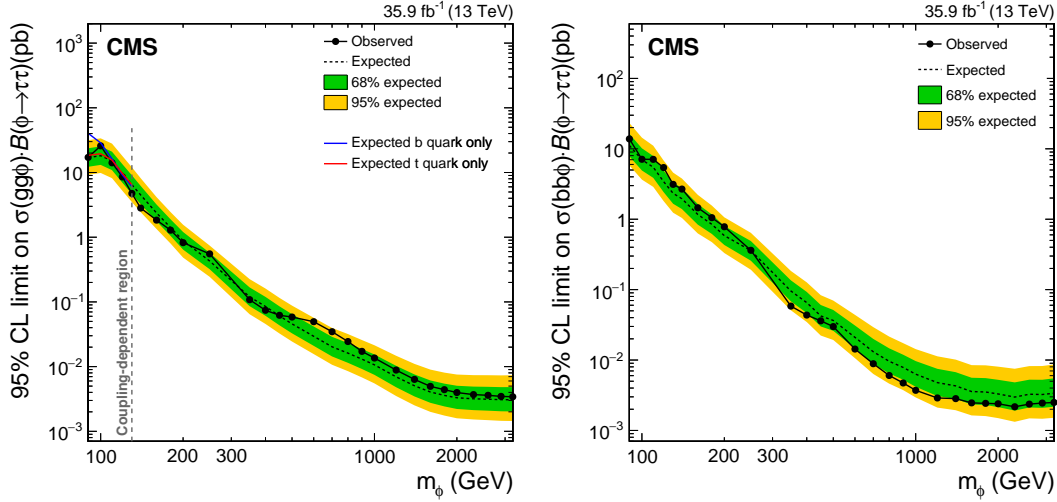
Figure 6.6 – Expected and observed 95 % CL upper limits for the production of a single narrow resonance ϕ with a mass between 90 GeV and 3.2 TeV in the $\tau\tau$ final state for the production via gluon fusion (ggϕ, left) and in association with b quarks (bbϕ, right). In the left panel the expected exclusion limits for the cases where (blue continuous line) only the b quark and (red continuous line) only the t quark are taken into account in the fermion loop are shown [13].

For a given value of m_A and $\tan\beta$, $\mu = 1$ corresponds to the signal prediction and $\mu = 0$ corresponds to the background-only prediction, including the SM Higgs boson as background process. The observed and expected exclusion contours for the $m_h^{\text{mod}+}$ and hMSSM scenarios are shown in Fig. 6.8. The excluded mass region is extended by almost a factor of two compared to the previous publication from the CMS collaboration [157] and reaches up to 1.6 TeV. For both the $m_h^{\text{mod}+}$ and hMSSM scenarios, the exclusion contours extend down to $\tan\beta \approx 6$ for the mass region $m_A \lesssim 250$ GeV, which is similar to the sensitivity of the previous publication. The reason why this analysis does not improve the sensitivity in the low mass region is due to three main factors:

- The previous analysis used the fully reconstructed mass of the ditau system as discriminating variable instead of m_T^{tot}. While m_T^{tot} provides more sensitivity in the high mass region it gives less sensitivity for lower masses.

- The reweighting to NLO precision gives generally a softer p_T spectrum for the Higgs boson production via gluon fusion, which dominates for low values of $\tan\beta$ and leads to a smaller signal acceptance.

- Higher p_T thresholds are needed at the increased instantaneous luminosity due to the choice of single lepton triggers in the semi-leptonic channels. This leads to a reduced signal acceptance.
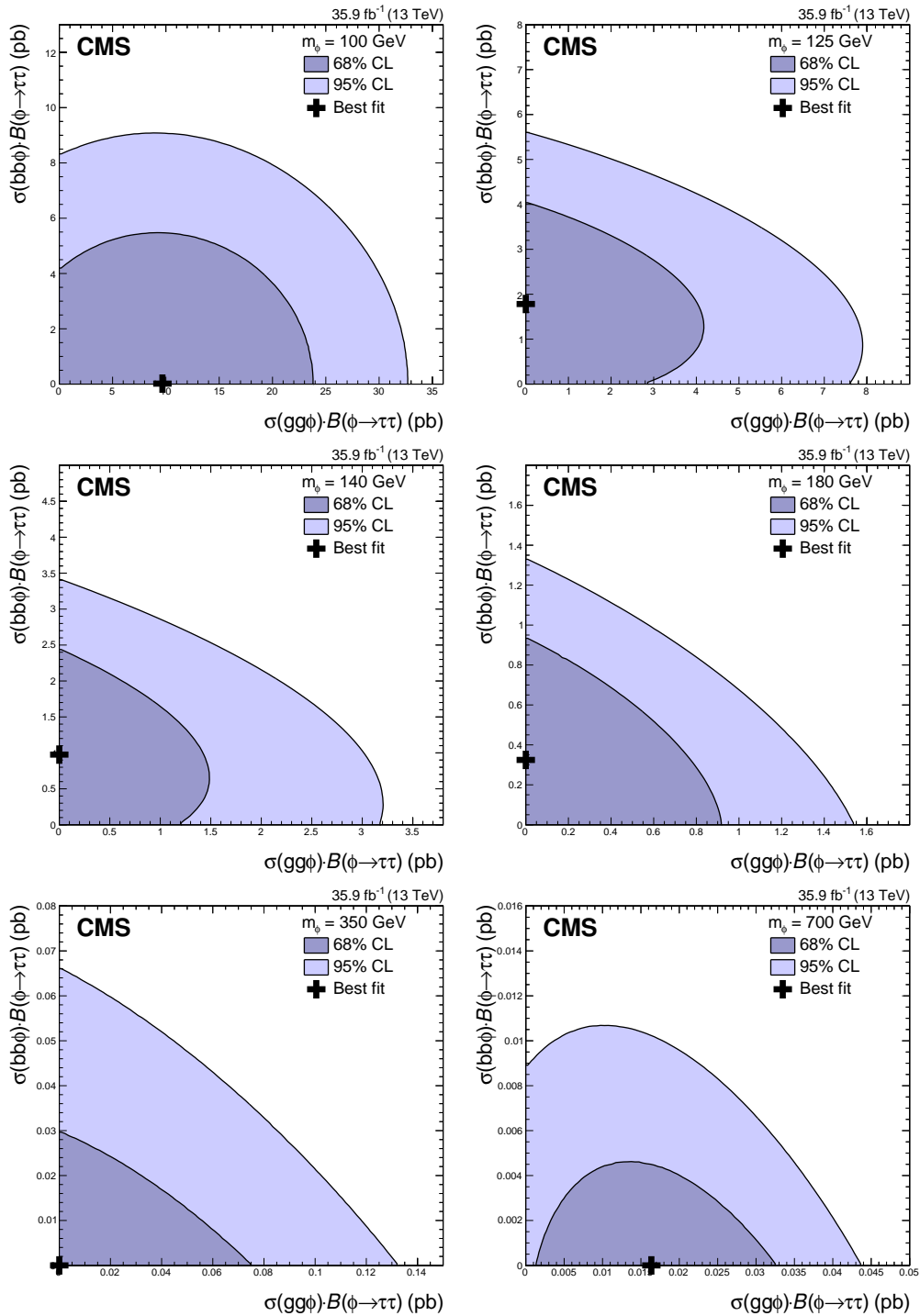
Figure 6.7 – Scan of the likelihood function for the search in the $\tau\tau$ final state for a single narrow resonance ϕ produced via gluon fusion (ggϕ) or in association with b quarks (bbϕ). A subset of the tested mass points is shown, with the corresponding mass m_ϕ indicated at the top right of each panel [13].
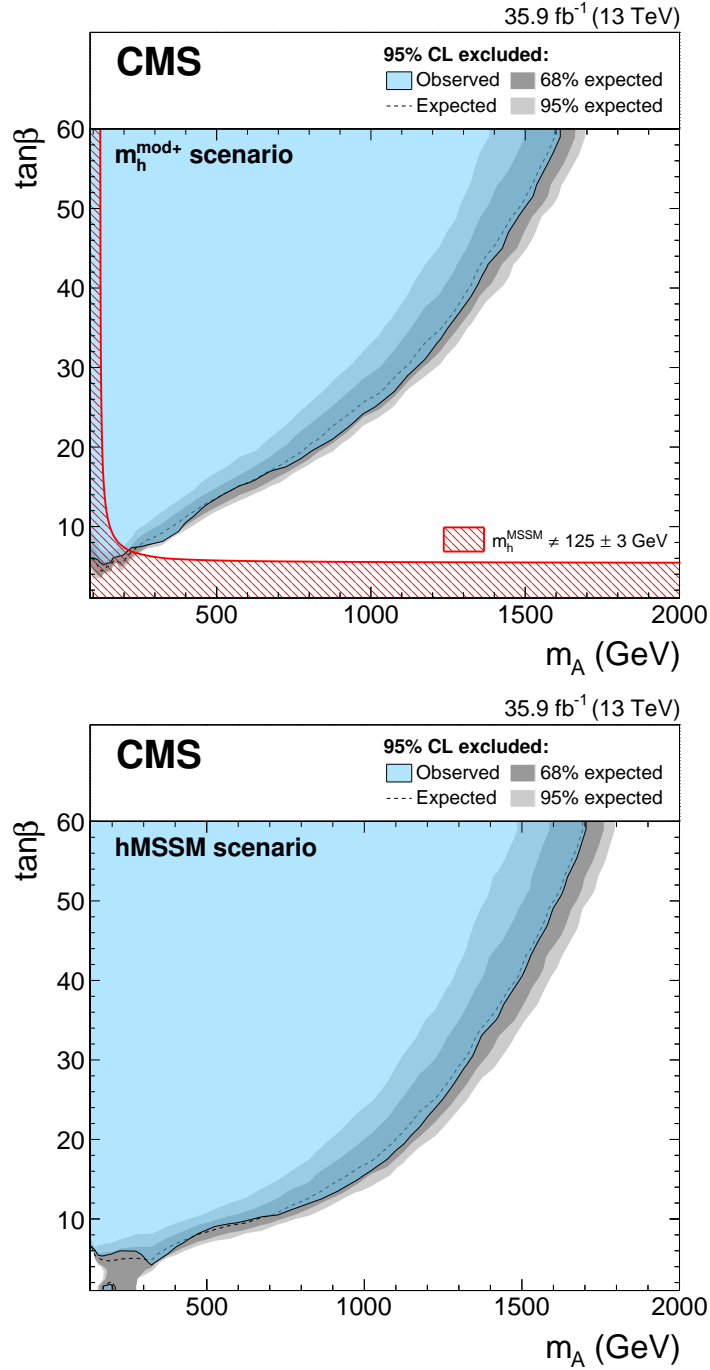
Figure 6.8 – Expected and observed 95 % CL exclusion contour in the MSSM $m_\mathrm{h}^\mathrm{mod+}$ (top) and in the hMSSM (bottom) scenarios. For the $m_\mathrm{h}^\mathrm{mod+}$ scenario, those parts of the parameter space where m_h deviates by more then $\pm 3\,\mathrm{GeV}$ from the mass of the observed Higgs boson at $125\,\mathrm{GeV}$ are indicated by a red hatched area [13].

CONCLUSIONS AND OUTLOOK

After the discovery of a particle compatible with the SM Higgs boson predictions during the LHC Run-1 the experimental effort in the Higgs sector during LHC Run-2 has been shifted towards precision measurements and searches for signatures hinting at physics beyond the standard model, as discussed in Section 2.3. The two analyses presented in this thesis are at the front of these efforts by targeting the $e\mu$, $e\tau_h$, $\mu\tau_h$ and $\tau_h\tau_h$ final states of the Higgs boson decay into a pair of τ leptons.

In Chapter 5 the measurement of the $pp \to H \to \tau\tau$ cross section has been presented in various levels of differentiation. The measurement is based on proton-proton collisions recorded in 2016 and 2017 by CMS at a centre-of-mass energy of 13 TeV, corresponding to an integrated luminosity of 77.4 fb^{-1}. This analysis utilizes the τ-embedding technique and the F_F method, which allows to estimate about 90 % of the background processes from data. A new method to differentiate between signal and background has been implemented. It is based on an NN classification algorithm that splits the recorded data into several signal and background categories, which are finally used to extract the signal and to constrain the background contributions. The inclusive measurement of the Higgs production cross section times the branching fraction into τ leptons results in $\sigma_{incl}\mathcal{B}(H \to \tau\tau) = 2.56$ pb ± 0.48 (stat) pb ± 0.34 (syst) pb. The inclusive signal strength is measured as $\mu_{incl} = 0.75^{+0.18}_{-0.17}$, which is 1.4σ below 1, the value predicted by the SM. The measurements split by the $gg \to H$, bbH and VBF+V(qq)H production modes give $\sigma(gg \to H, bbH)\mathcal{B}(H \to \tau\tau) = 1.11$ pb ± 0.81 (stat) pb ± 0.78 (syst) pb and $\sigma(VBF+V(qq)H)\mathcal{B}(H \to \tau\tau) = 0.34$ pb ± 0.08 (stat) pb ± 0.09 (syst) pb. Finally, measurements in terms of simplified template cross sections have been presented, which probe the kinematic properties of the Higgs boson production.

This analysis provides an intermediate result towards a measurement on the full LHC Run-2 dataset. An additional purpose is to evaluate the performance and establish the newly introduced NN classification approach in combination with the data-driven background estimation methods. The NN approach is a powerful

analysis tool, however, there is room for further improvement towards a full LHC Run-2 analysis. One item to be addressed is a revision of the observables used as input to the NN as given in Table 5.2. This includes the pruning of variables that have negligible impact on the classification task by exploiting the approach described in Section 5.4.2 and aligning the input parameters between the individual data-taking periods. Furthermore, apparent mismodeling of important discriminating parameters, e.g. m_{vis} in 2017, needs to be investigated and corrected for to avoid suboptimal performance. Another item is to use events from τ-embedding and the F_{F} method for training in order to decrease the generalization error of the classification task. Furthermore, the baseline selection, which is based on a cut-based approach described in Ref. [10], may be optimized with focus on the performance of the NN approach.

In Chapter 6 a search for additional neutral Higgs bosons in the $\tau\tau$ final state has been presented. The measurement is based on proton-proton collisions recorded in 2016 by CMS at a centre-of-mass energy of $13\,\text{TeV}$, corresponding to an integrated luminosity of $35.9\,\text{fb}^{-1}$. No significant excess hinting at the presence of additional neutral Higgs bosons was found in the investigated mass range. Upper limits have been set on the cross section for the Higgs boson production via ggF and in association with b quarks times the branching fraction to a pair of τ leptons. Furthermore, the results have been interpreted in terms of the $m_{\text{h}}^{\text{mod}+}$ and the hMSSM benchmark scenarios. The model-dependent interpretations exclude additional neutral Higgs bosons below $m_{\text{A}} = 250\,\text{GeV}$ for all $\tan\beta$ values and the exclusion contours reach $1.6\,\text{TeV}$ for $\tan\beta = 60$.

A future search based on the full LHC Run-2 dataset can benefit from the excellent performance of the CMS detector during this data-taking period, which in total recorded about 4 times the amount of data compared to 2016. Further improvements are possible by exploiting additional τ_{h} triggers in the $e\tau_{\text{h}}$, $\mu\tau_{\text{h}}$ and $\tau_{\text{h}}\tau_{\text{h}}$ final states. In the semi-leptonic channels the use of a lepton + tau trigger gives access to kinematic regions with lower transverse momentum extending the sensitivity in the low mass region. The use of a single tau trigger targeting τ leptons with large transverse momentum increases the signal acceptance in the high mass region. Further improvement is possible with advanced analysis techniques such as the NN approach described in Chapter 5.

# APPENDIX: HIGGS BOSON PRODUCTION AND DECAY TO THE $\tau\tau$ FINAL STATE

A.1 Definition of Input Parameters

Parameter	Definition
$p_{\mathrm{T}}^{\tau_1}$	Transverse momentum of the leading τ candidate.
$p_{\mathrm{T}}^{\tau_2}$	Transverse momentum of the sub-leading τ candidate.
$p_{\mathrm{T}}^{\mathrm{vis}}$	Vector sum of the transverse momenta of the visible $\tau\tau$ decay products.
m_{vis}	Visible mass of the $\tau\tau$ system.
$p_{\mathrm{T}}(\mathrm{jet}_1)$	Transverse momentum of the leading jet.
$\eta\ (\mathrm{jet}_1)$	Pseudorapidity of the leading jet.
$p_{\mathrm{T}}(\mathrm{jet}_2)$	Transverse momentum of the sub-leading jet.
$\eta\ (\mathrm{jet}_2)$	Pseudorapidity of the leading jet.
N_{jet}	Number of jets with $p_{\mathrm{T}} > 30\,\mathrm{GeV}$.
$p_{\mathrm{T}}(\mathrm{b\ jet}_1)$	Transverse momentum of the leading b-tagged jet within $\lvert\eta\rvert < 2.4$.
$p_{\mathrm{T}}(\mathrm{b\ jet}_2)$	Transverse momentum of the sub-leading b-tagged jet within $\lvert\eta\rvert < 2.4$.
$N_{\mathrm{b\ jet}}$	Number of b-tagged jets with $p_{\mathrm{T}} > 20\,\mathrm{GeV}$ and within $\lvert\eta\rvert < 2.4$.
$p_{\mathrm{T}}^{\mathrm{miss}}$	Missing transverse momentum as defined in Eq. (4.4).
$p_{\mathrm{T}}^{\mathrm{jj}}$	Vector sum of the transverse momenta of the two leading jets with $p_{\mathrm{T}} > 30\,\mathrm{GeV}$.
m_{jj}	Mass of the di-jet system of the leading and sub-leading jet.
$m_{\tau\tau}^{\mathrm{SV}}$	Mass estimate of the $\tau\tau$ system using the SVFIT algorithm [174]. This algorithm combines the missing transverse momentum $p_{\mathrm{T}}^{\mathrm{miss}}$ with the four-vectors of the visible τ decay products for a likelihood based mass estimate.
$p_{\mathrm{T}\tau\tau}^{\mathrm{SV}}$	Transverse momentum estimate of the $\tau\tau$ system using the SVFIT algorithm [174].
$m_{\mathrm{T}\tau\tau}^{\mathrm{SV}}$	Transverse mass estimate of the $\tau\tau$ system using the SVFIT algorithm [174].

$p_{\mathrm{T}}^{\tau\tau+\mathrm{miss}}$	Vector sum of the reconstructed τ momenta and the missing transverse momentum.
$p_{\mathrm{T}}^{\tau\tau\mathrm{jj}+\mathrm{miss}}$	Vector sum of the reconstructed τ momenta, the di-jet system and the missing transverse momentum.
$\Delta\eta_{\mathrm{jj}}$	Relative distance in η between the leading and sub-leading jet.
m_{T}^{μ}	Transverse mass calculated with Eq. (5.1) using the reconstructed transverse momentum of the muon and the missing transverse momentum.
$m_{\mathrm{T}}^{\mathrm{e}}$	Transverse mass calculated with Eq. (5.1) using the reconstructed transverse momentum of the electron and the missing transverse momentum.
$m_{\mathrm{T}}^{\mathrm{e}+\mu}$	Transverse mass calculated with Eq. (5.1) using the reconstructed transverse momentum sum of the electron and muon in the eμ channel and the missing transverse momentum.
$\max(m_{\mathrm{T}}^{\mu}, m_{\mathrm{T}}^{\mathrm{e}})$	Maximum of the transverse masses m_{T}^{μ} and $m_{\mathrm{T}}^{\mathrm{e}}$ in the eμ channel.
$m_{\mathrm{T}}^{\tau_{\mathrm{h}}}$	Transverse mass calculated with Eq. (5.1) using the reconstructed transverse momentum of the τ_{h} candidate and the missing transverse momentum.
$\Delta R^{\mathrm{e}\mu}$	Distance between the electron and the muon in the eμ channel using Eq. (3.2).
D_{ζ}	D_{ζ} variable defined as in Eq. (5.2).

Table A.1 – Definition of the 28 parameters used as input to the neural network.

A.2 Fake Factors Measured for the 2016 Dataset
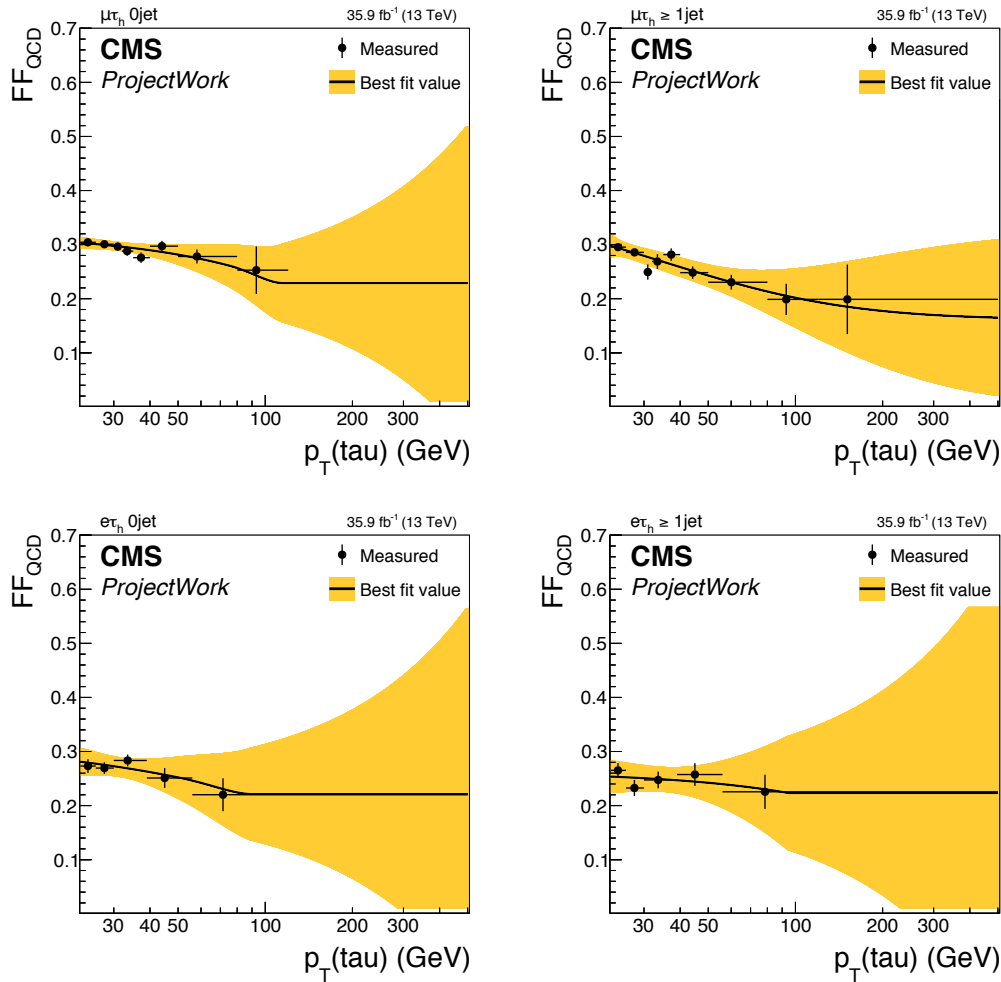
Figure A.1 – The uncorrected fake factors as measured for the 2016 dataset in the $\mathrm{DR_{QCD}}$ for different jet multiplicities as indicated in the plots. Shown are the measurements for $\mu\tau_\mathrm{h}$ (top) and $e\tau_\mathrm{h}$ (bottom). The fake faktor dependency versus p_T is fit with the sum of a Landau function and a polynomial of order one and set constant at high p_T.
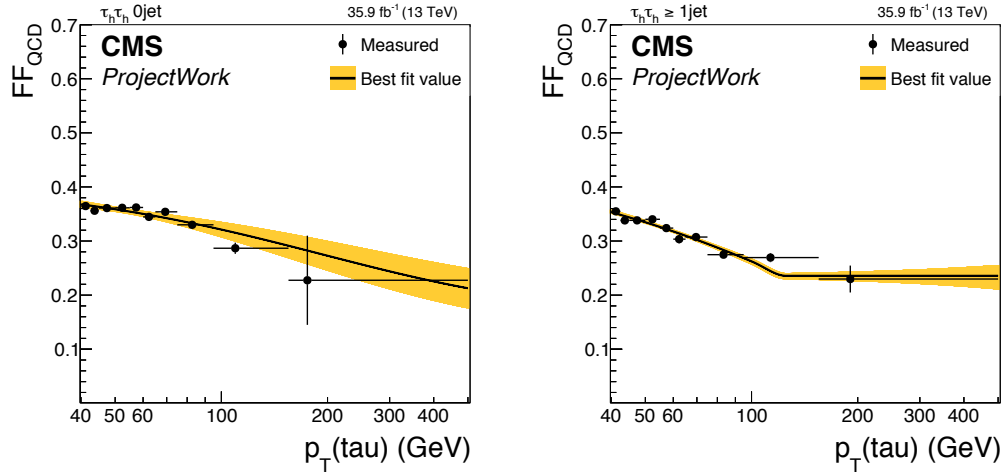
Figure A.2 – The uncorrected fake factors as measured for the 2016 dataset in the $\mathrm{DR_{QCD}}$ for different jet multiplicities as indicated in the plots. Shown are the measurements for $\tau_\mathrm{h}\tau_\mathrm{h}$. The fake faktor dependency versus p_T is fit with the sum of a Landau function and a polynomial of order one and set constant at high p_T.
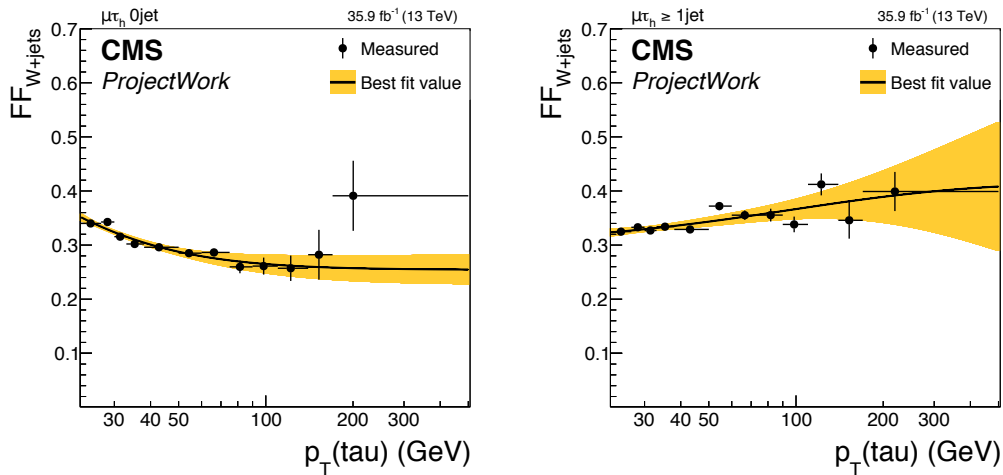
Figure A.3 – The uncorrected fake factors as measured for the 2016 dataset in the $\mathrm{DR_{W+jets}}$ for different jet multiplicities as indicated in the plots. Shown are the measurements for $\mu\tau_\mathrm{h}$. The fake faktor dependency versus p_T is fit with the sum of a Landau function and a polynomial of order one and set constant at high p_T.
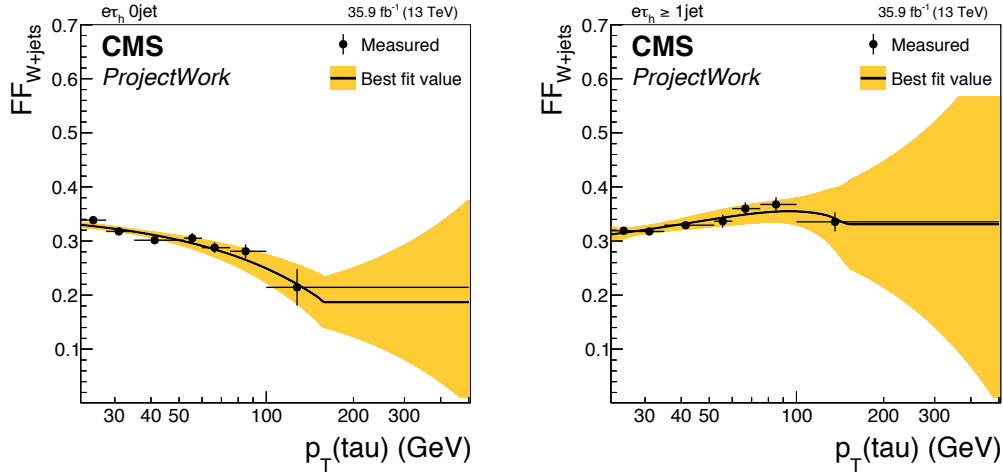
Figure A.4 – The uncorrected fake factors as measured for the 2016 dataset in the $\mathrm{DR_{W+jets}}$ for different jet multiplicities as indicated in the plots. Shown are the measurements for $\mathrm{e}\tau_\mathrm{h}$. The fake faktor dependency versus p_T is fit with the sum of a Landau function and a polynomial of order one and set constant at high p_T.
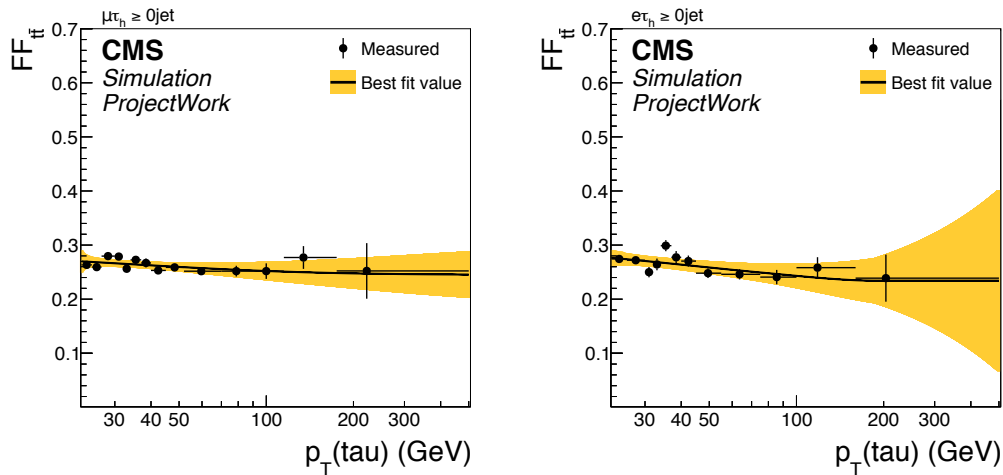
Figure A.5 – The uncorrected fake factors as estimated from $\mathrm{t\bar{t}}$ simulation for the 2016 dataset. Shown are the measurements for $\mu\tau_\mathrm{h}$ (left) and $\mathrm{e}\tau_\mathrm{h}$ (right). The fake faktor dependency versus p_T is fit with the sum of a Landau function and a polynomial of order one and set constant at high p_T.

A.3 Fake Factor Corrections Measured for the 2016 Dataset
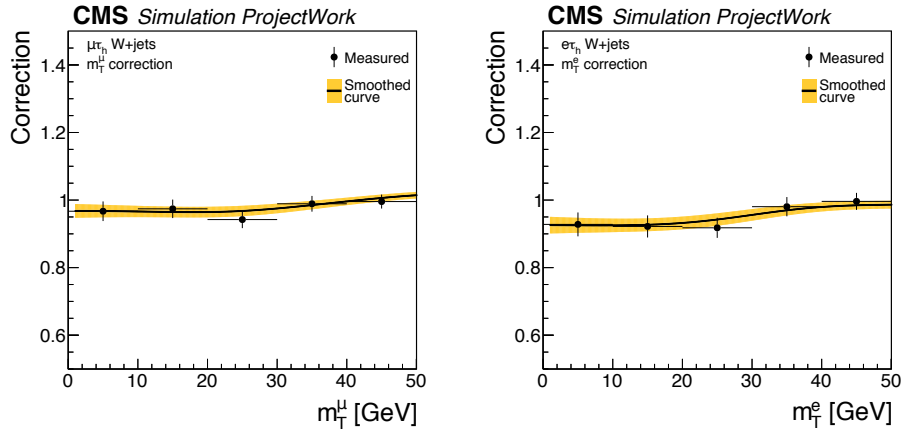
Figure A.6 – W+jets m_T correction: Observed over estimated W+jets background ratio as a function of the transverse mass, obtained from simulation for the 2016 dataset. The smoothed ratio and its statistical uncertainty are shown for $\mu\tau_h$ (left) and $e\tau_h$ (right).
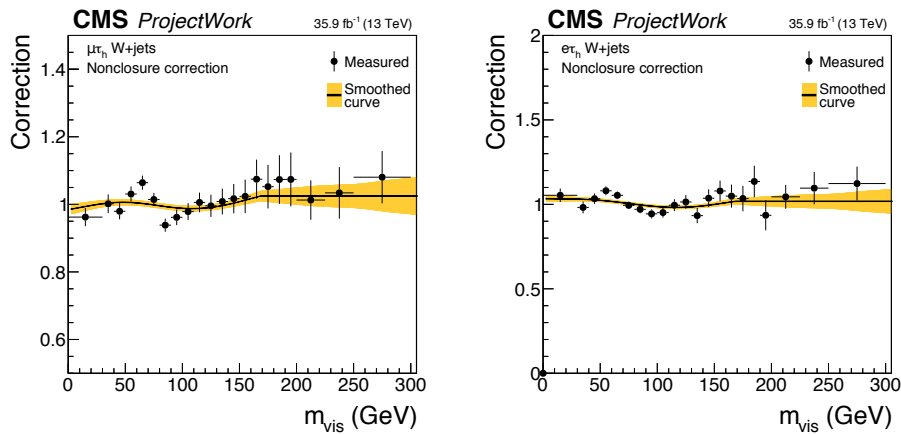
Figure A.7 – W+jets non-closure correction: Observed over estimated W+jets background ratio in the $m_T > 70\,\text{GeV}$, with b jet veto, as a function of m_{vis} for the 2016 dataset. The smoothed ratio and its statistical uncertainty are shown for $\mu\tau_h$ (left) and $e\tau_h$ (right).
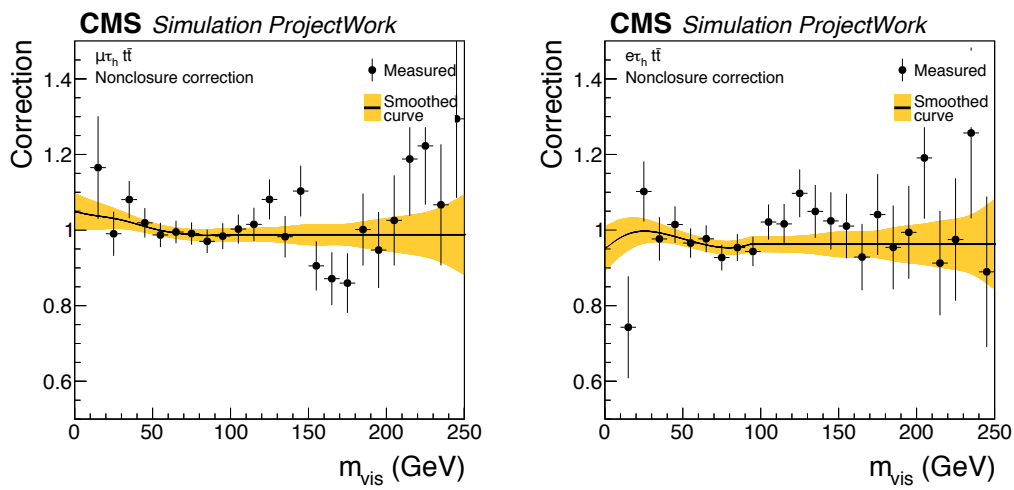
Figure A.8 $-$ $t\bar{t}$ non-closure correction: Actual over estimated background ratio from simulation, as a function of m_{vis} for the 2016 dataset. The smoothed ratio and its statistical uncertainty are shown for $\mu\tau_h$ (left) and $e\tau_h$ (right).
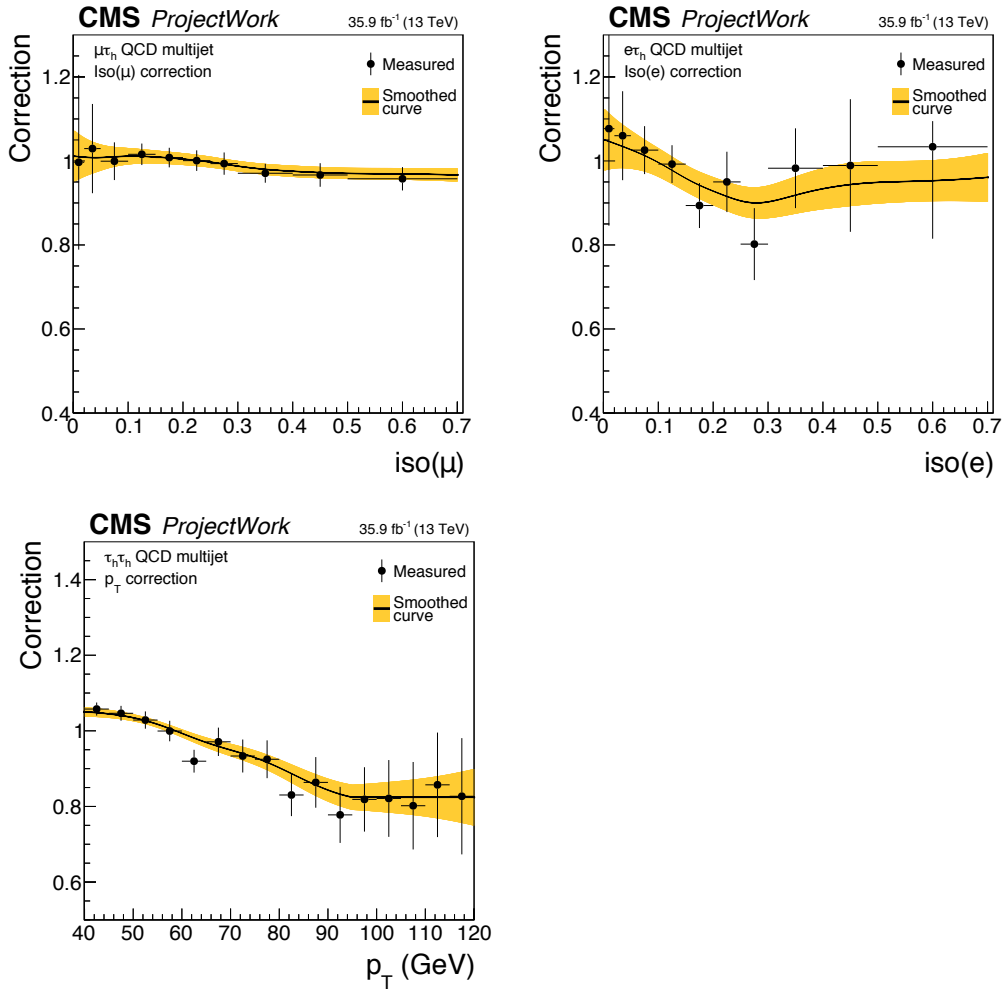
Figure A.9 – QCD multijet lepton correction: Observed over estimated QCD multijet background ratio in the same-sign region, as a function of the muon isolation ($\mu\tau_h$, top left) electron isolation ($e\tau_h$, top right) and the p_T of the second τ_h ($\tau_h\tau_h$, bottom left) for the 2016 dataset. The smoothed ratio and its statisitical uncertainty are shown.
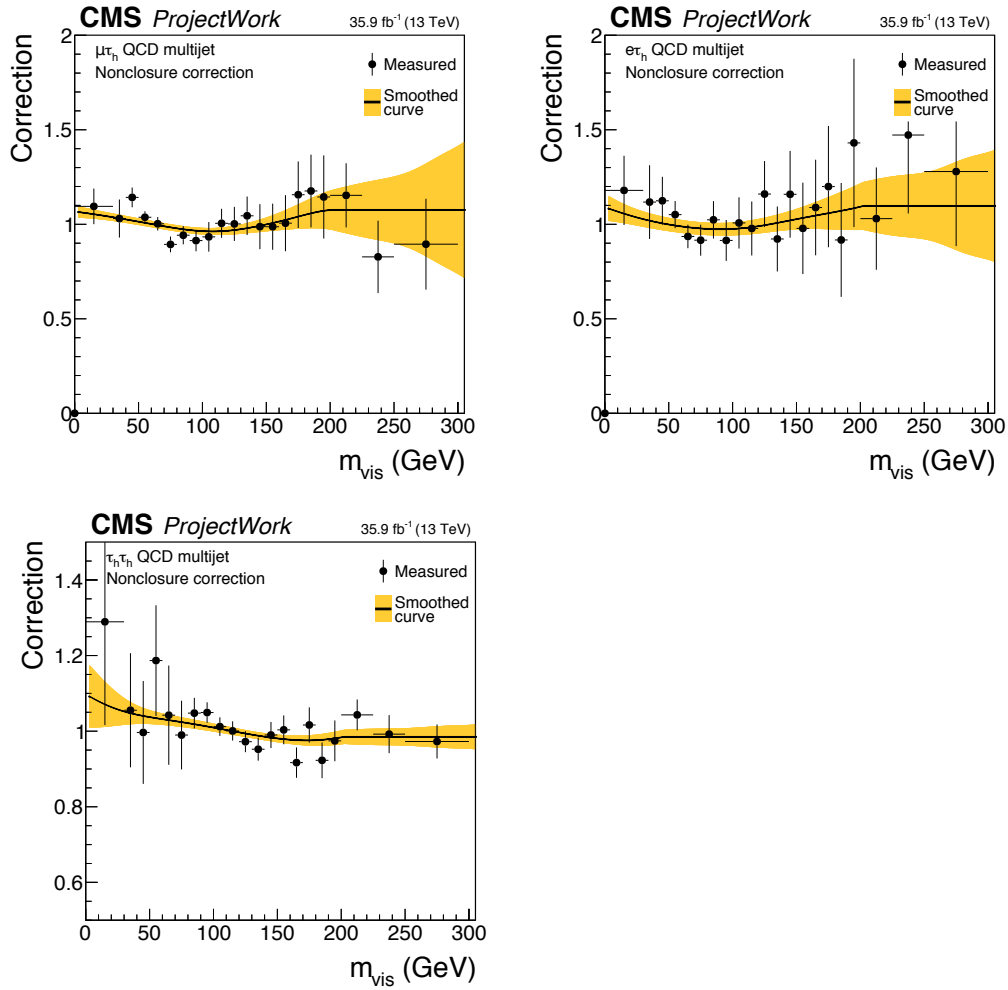
Figure A.10 – QCD multijet non-closure correction: Observed over estimated background ratio in the same-sign region with anti-isolated lepton, as a function of the visible mass for the 2016 dataset. The smoothed ratio and its statisitical uncertainty are shown for $\mu\tau_h$ (top left), $e\tau_h$ (top right) and $\tau_h\tau_h$ (bottom left).
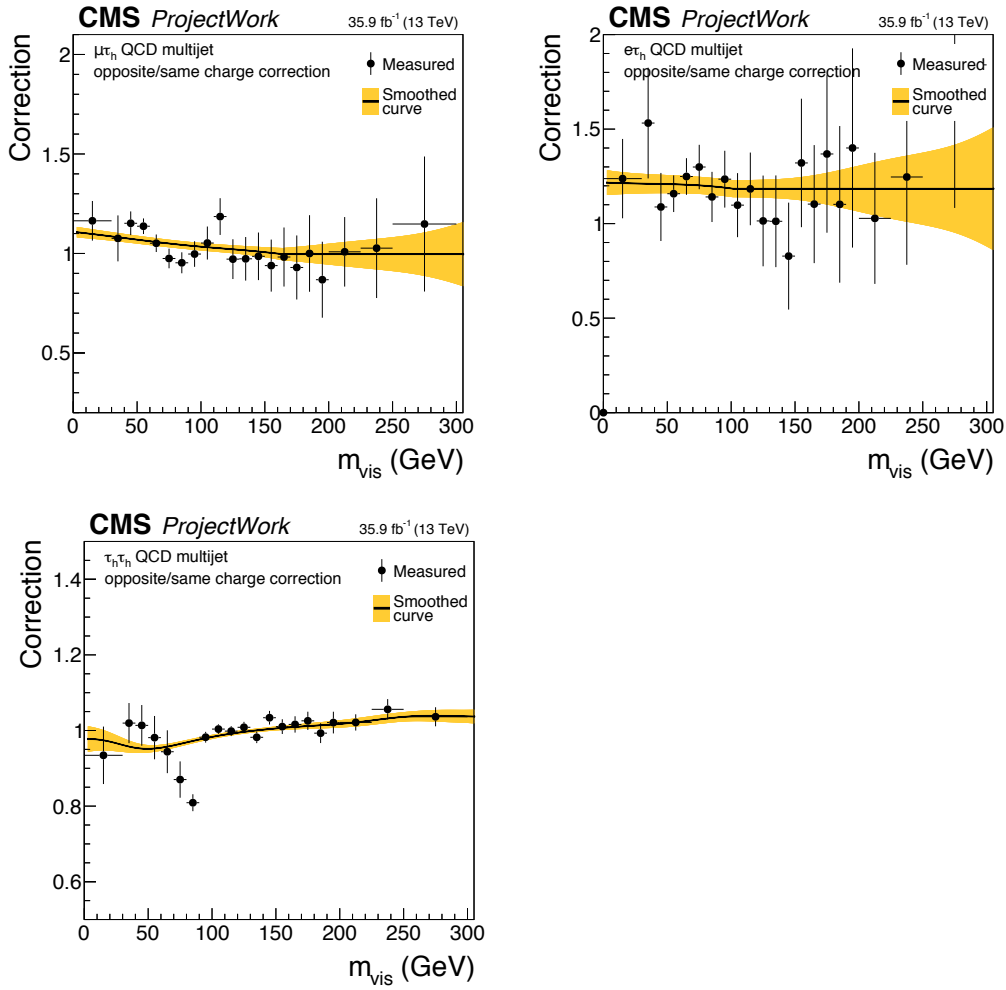
Figure A.11 – QCD FF multijet opposite/same-sign correction: Observed over estimated multijet contribution in the opposite-sign region with $I_{\mathrm{rel}}^{\ell} > 0.15$ ($\mu\tau_{\mathrm{h}}$, top left and $e\tau_{\mathrm{h}}$, top right) or when requiring the second τ_{h} to pass the `VeryLoose` but not the `Tight` τ_{h} ID discriminant working point (bottom left), as a function of m_{vis} after applying special fake fators derived in the same-sign region for the 2016 dataset. The smoothed ratio and its statisitical uncertainty are shown.

A.4 Fake Factors Measured for the 2017 Dataset
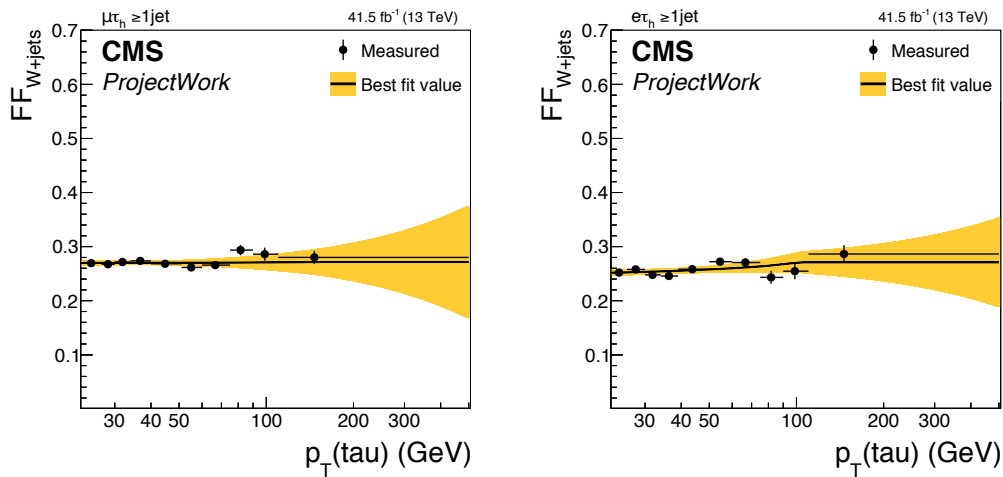
Figure A.12 – The uncorrected fake factors as measured for the 2017 dataset in the $\mathrm{DR_{W+jets}}$ with at least on jet in the event selection. Shown are the measurements for $e\tau_h$. The fake faktor dependency versus p_T is fit with the sum of a Landau function and a polynomial of order one and extrapolated with a constant at high p_T.

A.5 Fake Factor Corrections Measured in $e\tau_h$ for the 2017 Dataset
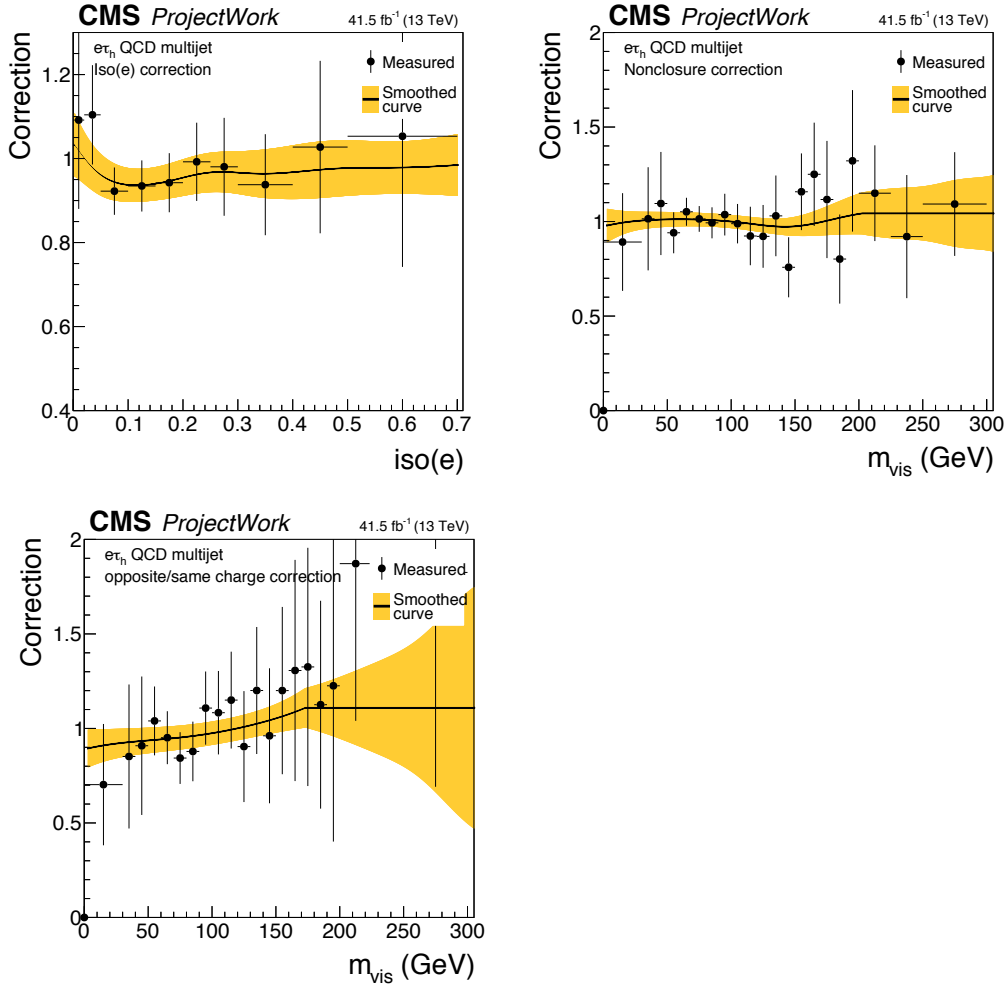
Figure A.13 – Shown are the QCD multijet lepton isolation (top left) non-closure correction (top right) and opposite/same-sign correction (bottom left) for the 2017 dataset in $e\tau_h$. The smoothed ratio and its statisitical uncertainty are shown.

A.6 Background Categories for the 2016 Dataset
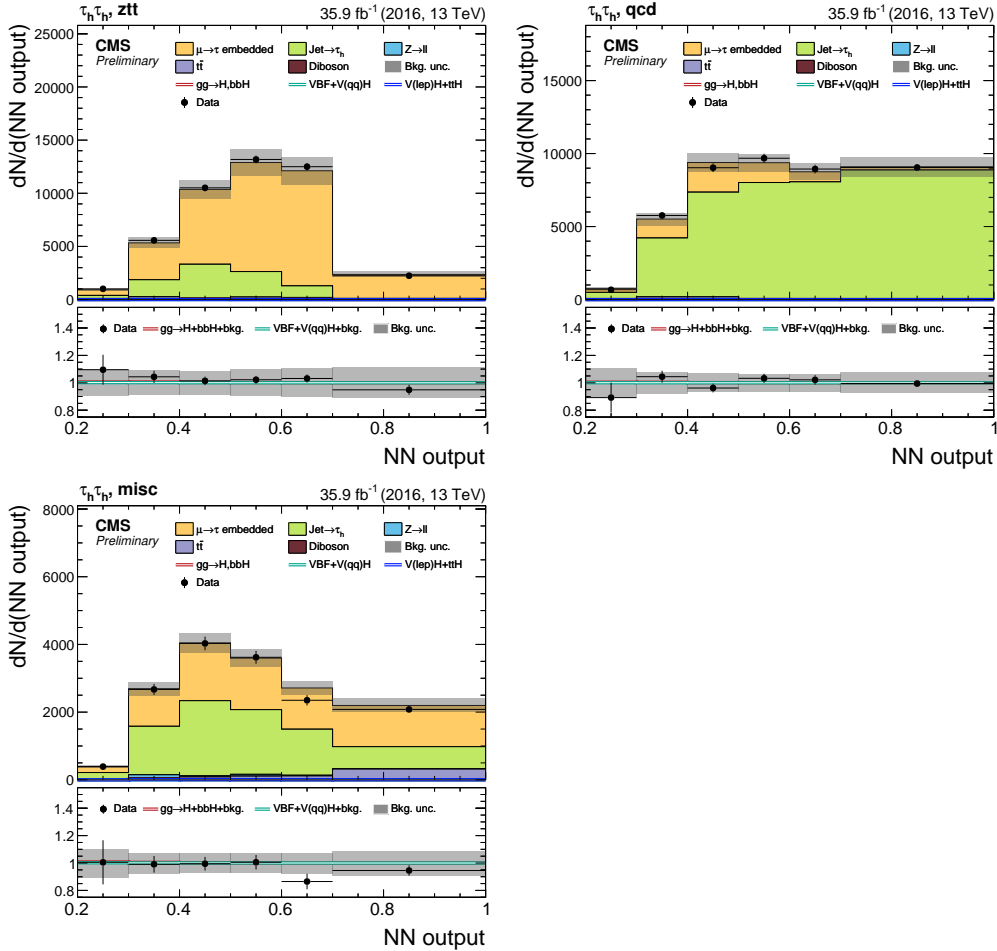
Figure A.14 − Distributions of the NN output in the `ztt` (top left), `qcd` (top right), `misc` (bottom centre) event categories, in the $\tau_h\tau_h$ final state. All distributions and uncertainties are shown after the fit that has been applied to the 2016 dataset [12].
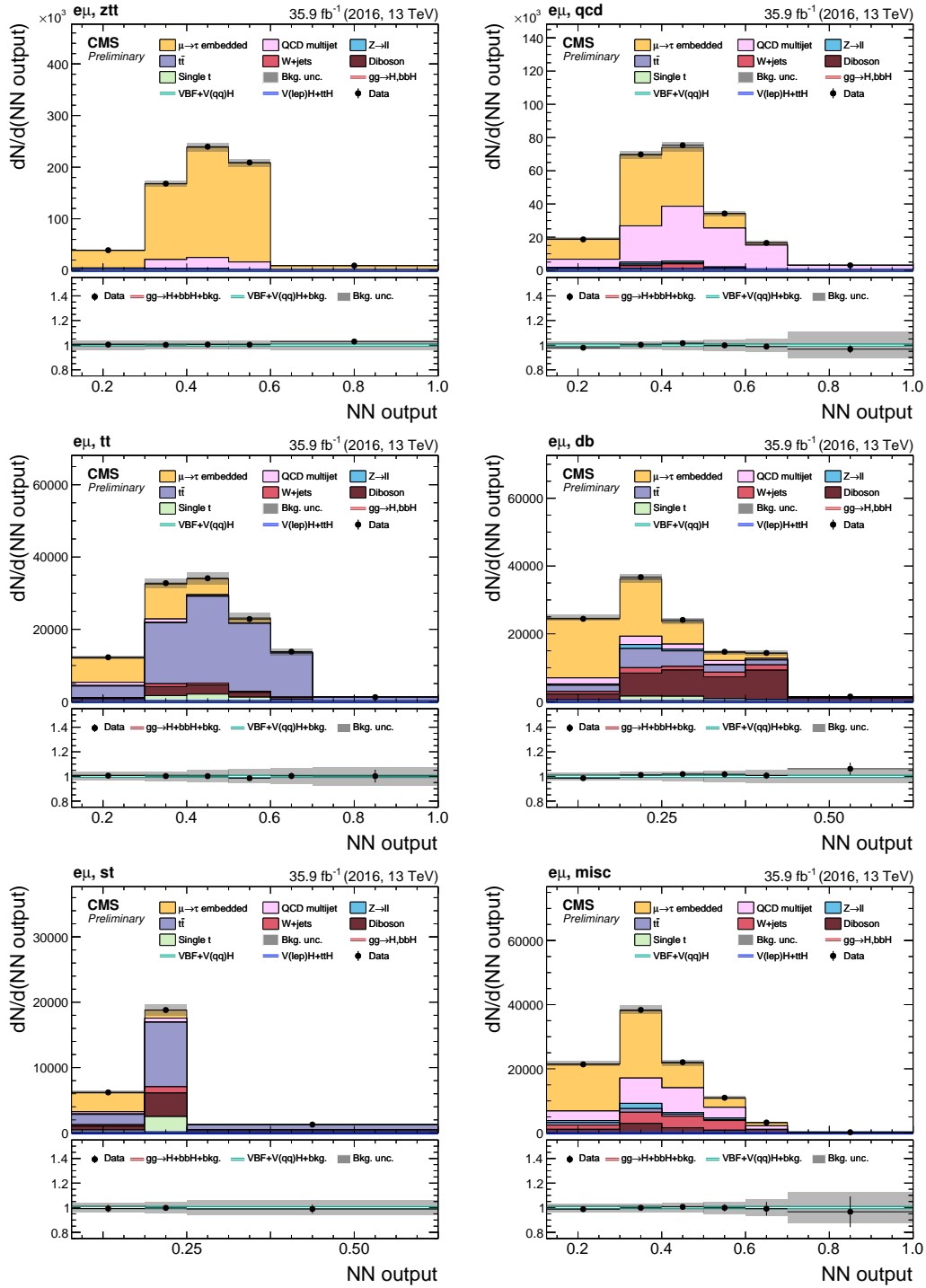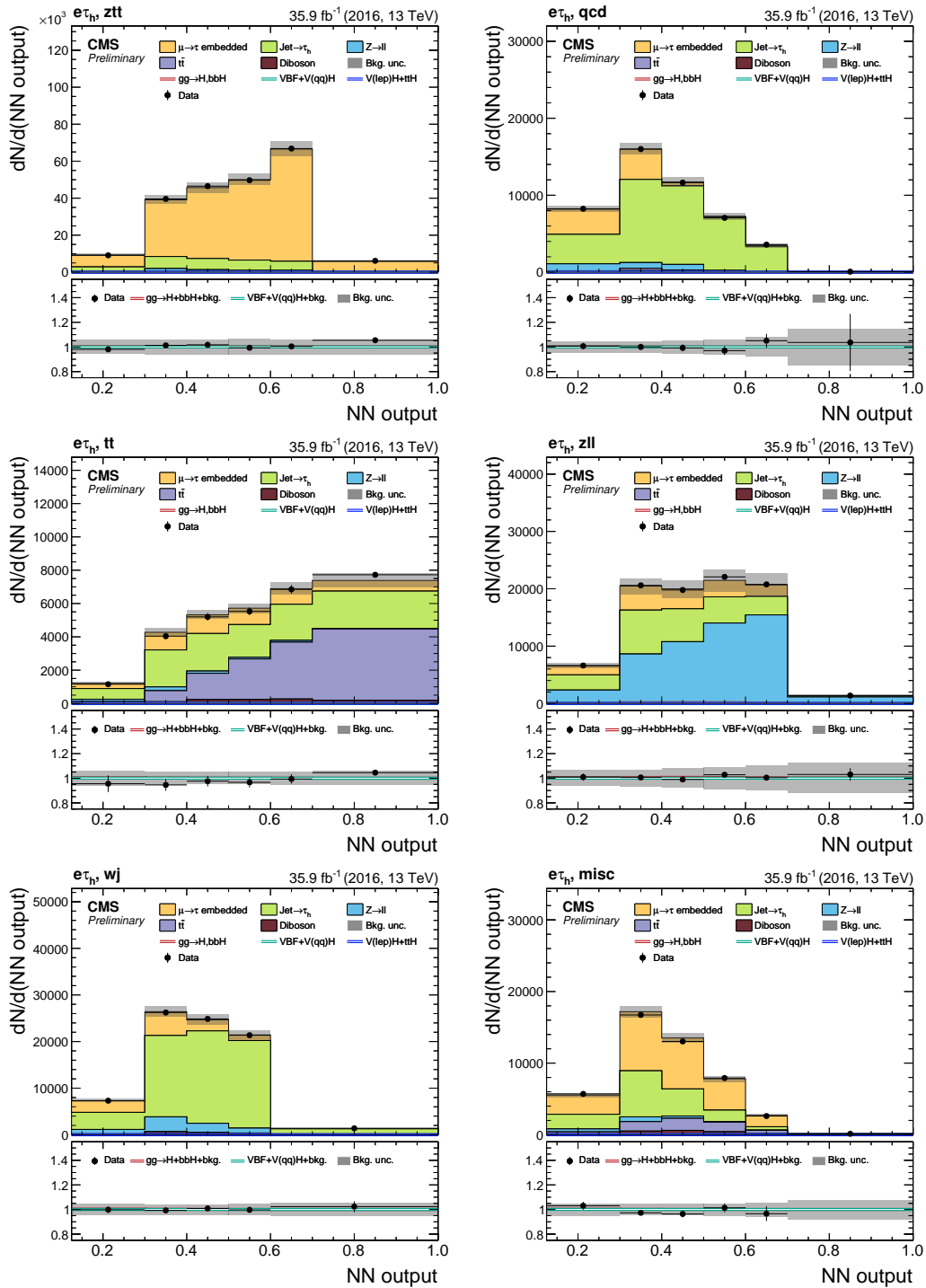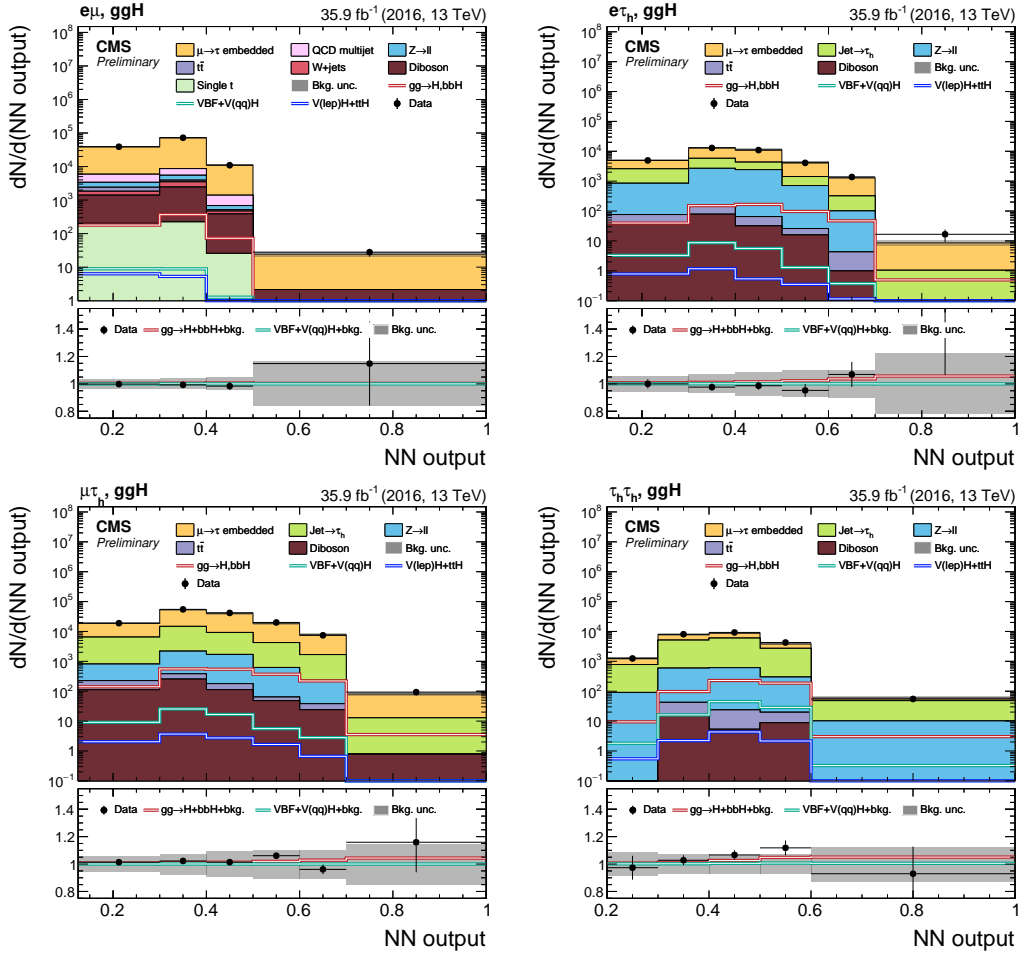
Figure A.15 – Distributions of the NN output in the `ztt` (top left), `qcd` (top right), `tt` (centre left), `db` (centre right), `st` (bottom left), and `misc` (bottom right) event categories, in the eμ final state. All distributions and uncertainties are shown after the fit that has been applied to the 2016 dataset [12].
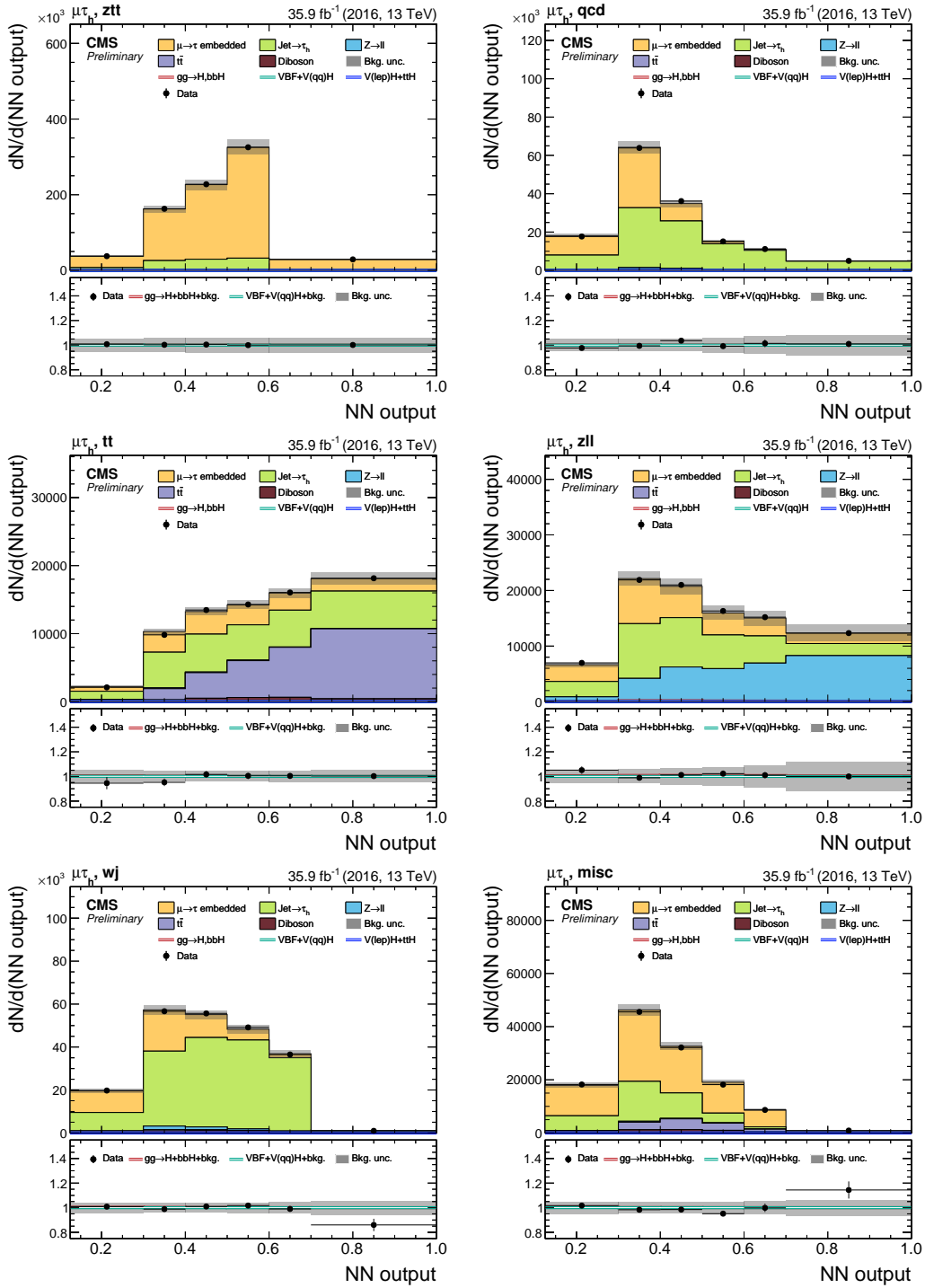
Figure A.16 – Distributions of the NN output in the `ztt` (top left), `qcd` (top right), `tt` (centre left), `zll` (centre right), `wj` (bottom left), and `misc` (bottom right) event categories, in the $\mu\tau_\mathrm{h}$ final state. All distributions and uncertainties are shown after the fit that has been applied to the 2016 dataset [12].

Figure A.17 – Distributions of the NN output in the `ztt` (top left), `qcd` (top right), `tt` (centre left), `zll` (centre right), `wj` (bottom left), and `misc` (bottom right) event categories, in the $e\tau_h$ final state. All distributions and uncertainties are shown after the fit that has been applied to the 2016 dataset [12].

A.7 Stage-0 Signal Categories on the 2016 Dataset

Figure A.18 − Distributions of the NN output in the ggH event category, in the eμ (top left), eτ_h (top right), $\mu\tau_\mathrm{h}$ (bottom left), and $\tau_\mathrm{h}\tau_\mathrm{h}$ (bottom right) final states, on the 2016 dataset. All distributions and uncertainties are shown after the fit that has been applied to the data [12].
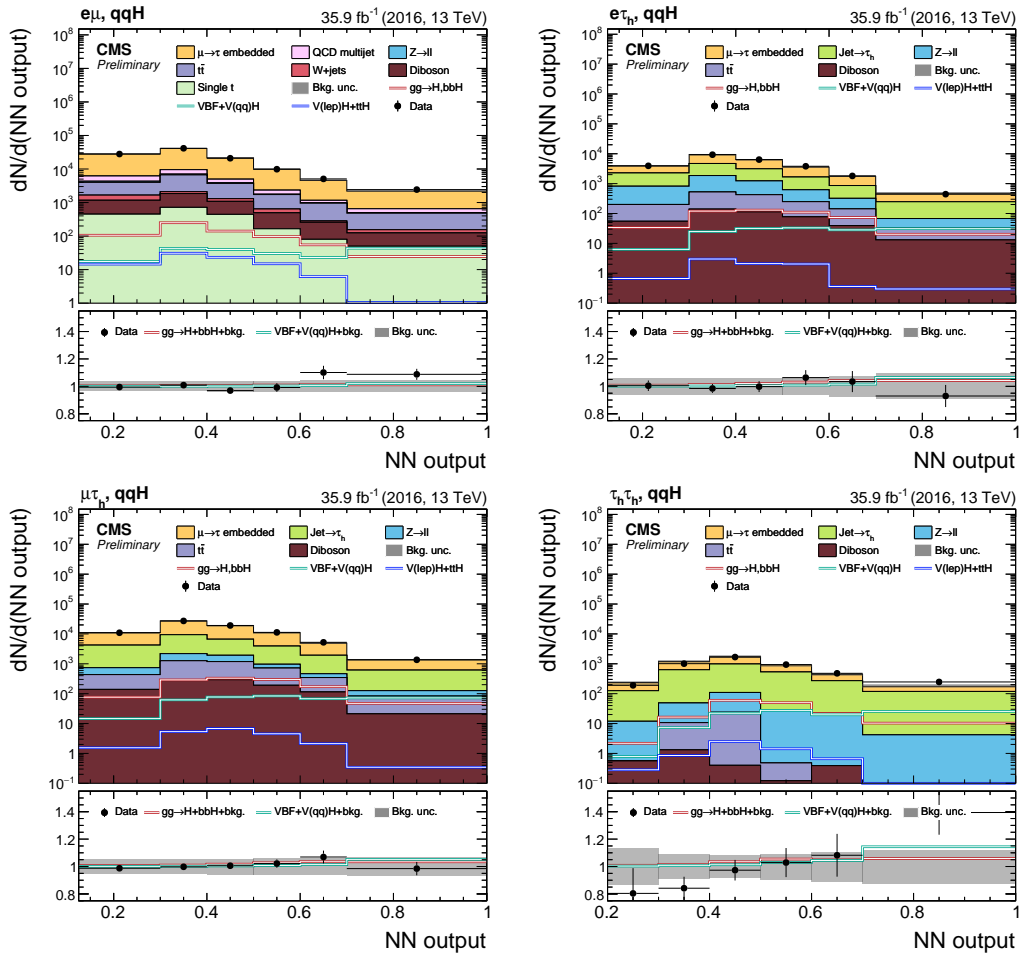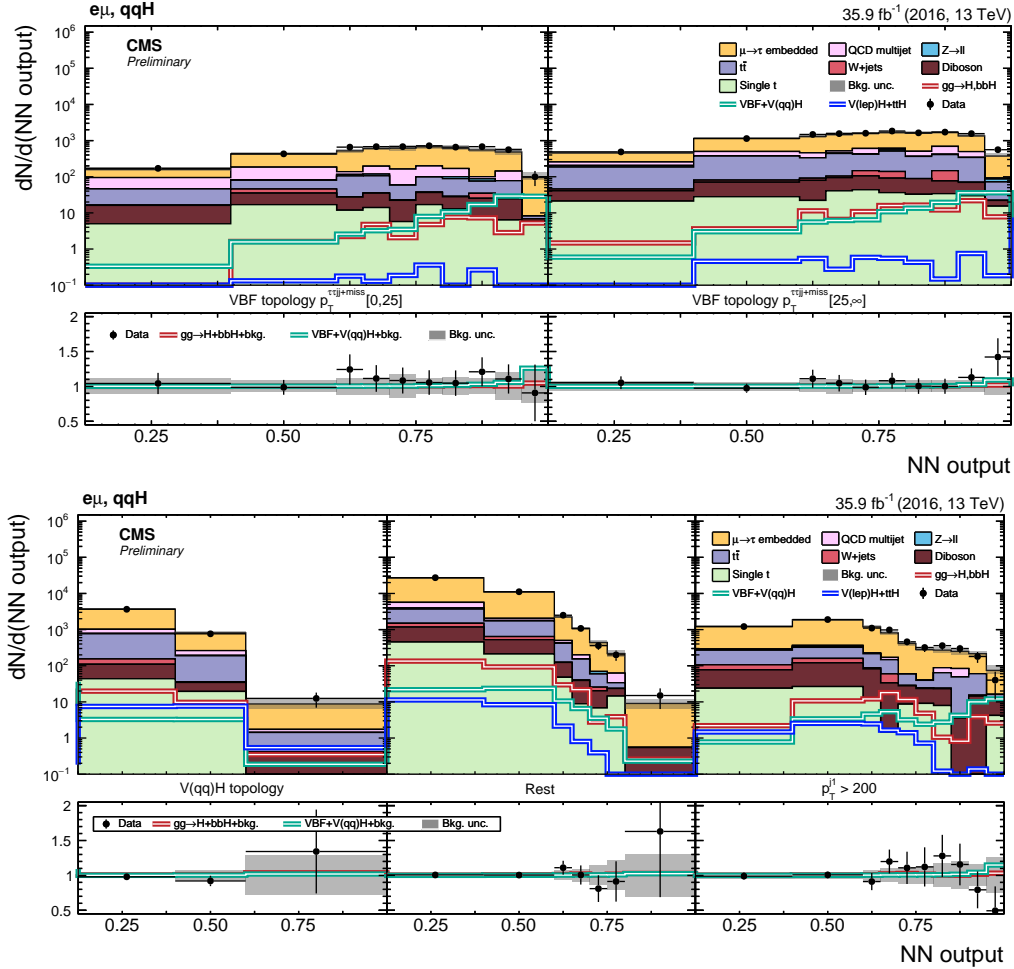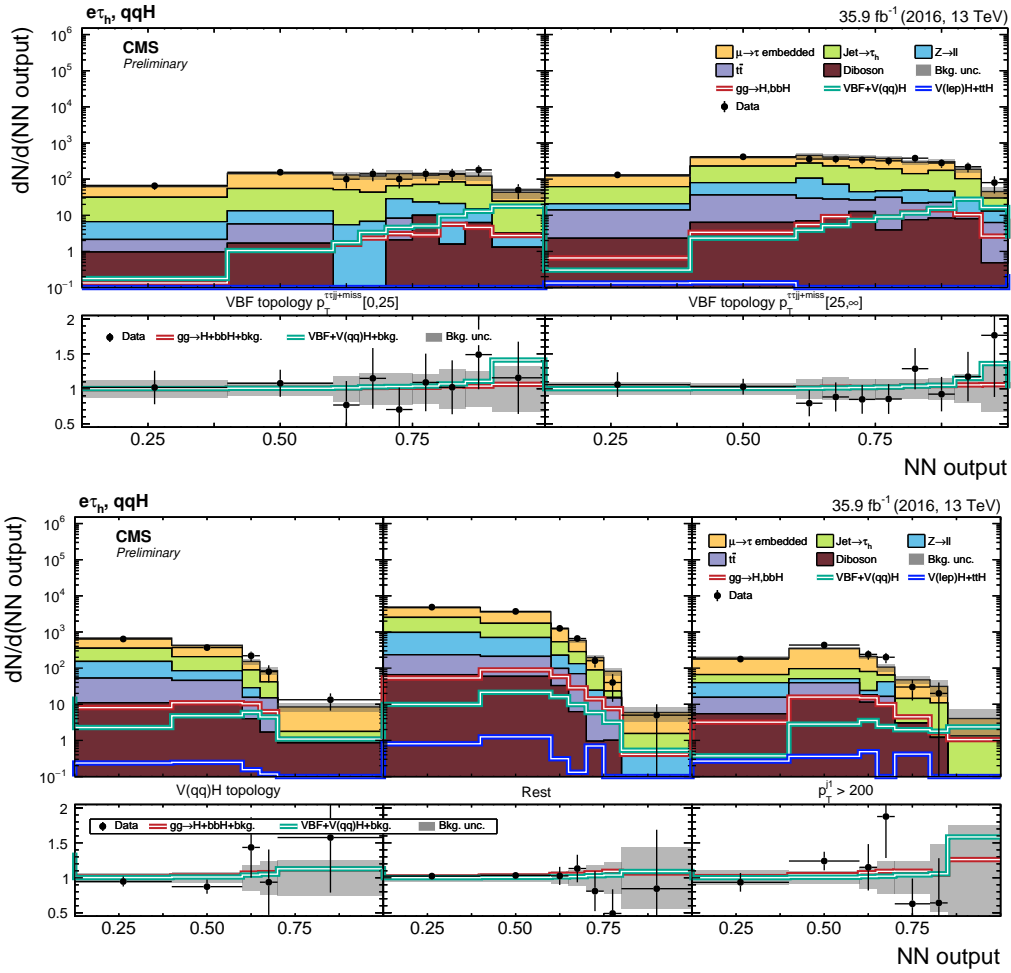
Figure A.19 − Distributions of the NN output in the qqH event category, in the eμ (top left), eτ_h (top right),) $\mu\tau_h$ (bottom left, and $\tau_h\tau_h$ (bottom right) final states, on the 2016 dataset. All distributions and uncertainties are shown after the fit that has been applied to the data [12].
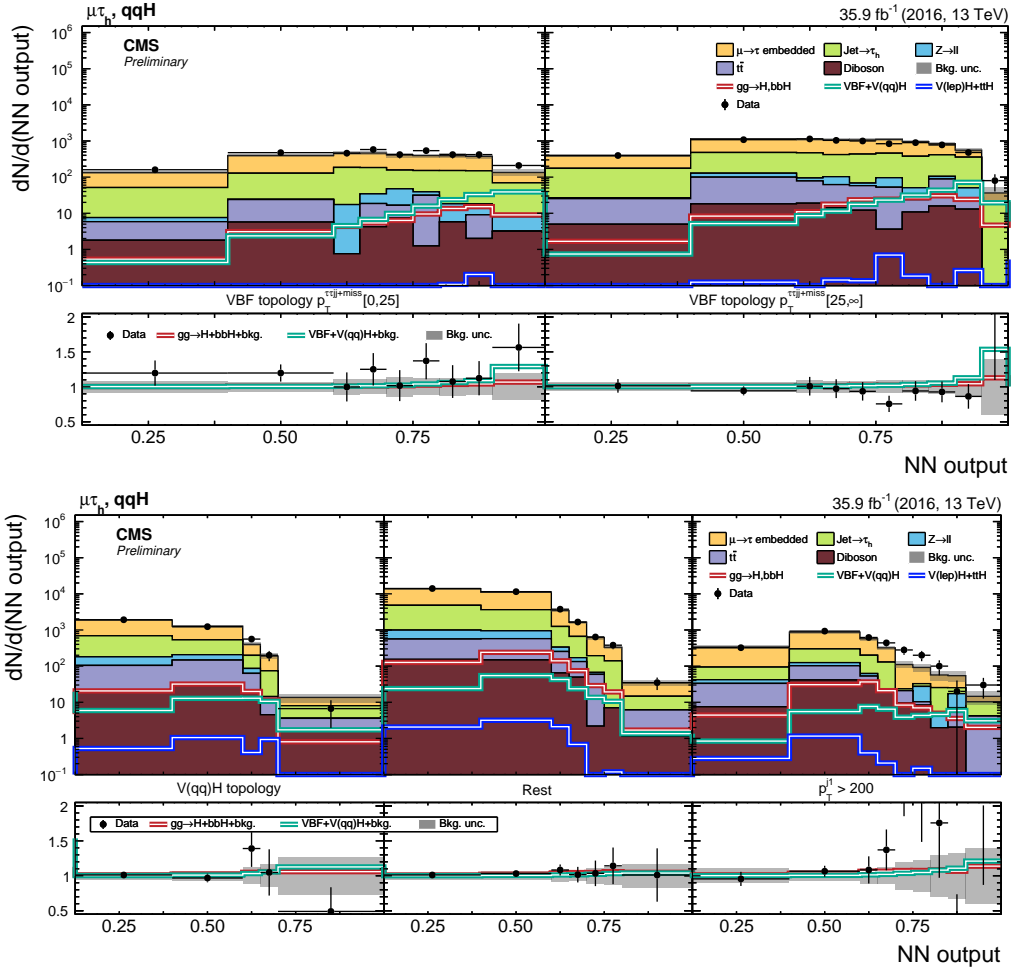
A.8 Stage-1 Signal Categories

Figure A.20 – Distributions of the NN output in the event categories used for the determination of the stage-1 simplified template cross sections for VBF+V(qq)H in the eμ final state. Shown are the categories with VBF topology in bins of $p_T^{\tau\tau jj+miss}$ (top row) and the event category with V(qq)H topology, the "Rest" category and the event category with $p_T^{j_1} > 200$ GeV (bottom row), on the 2016 dataset [12].
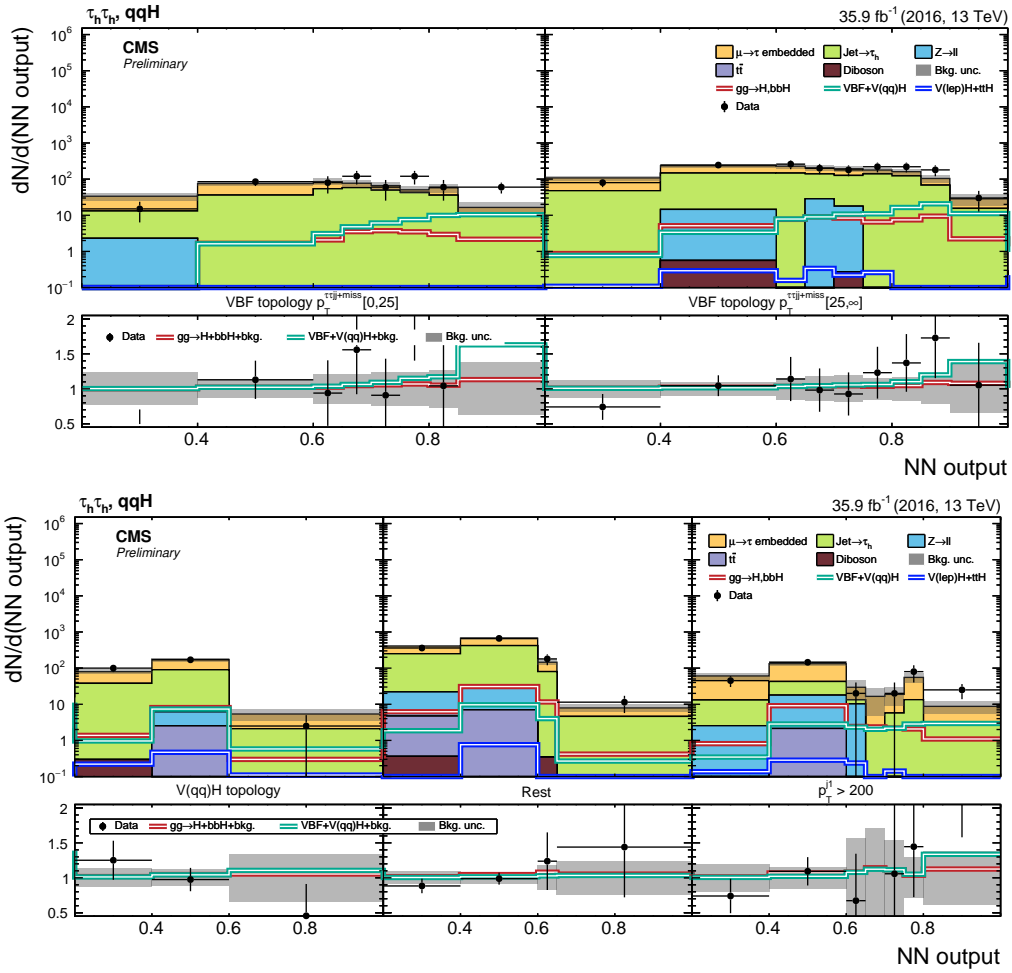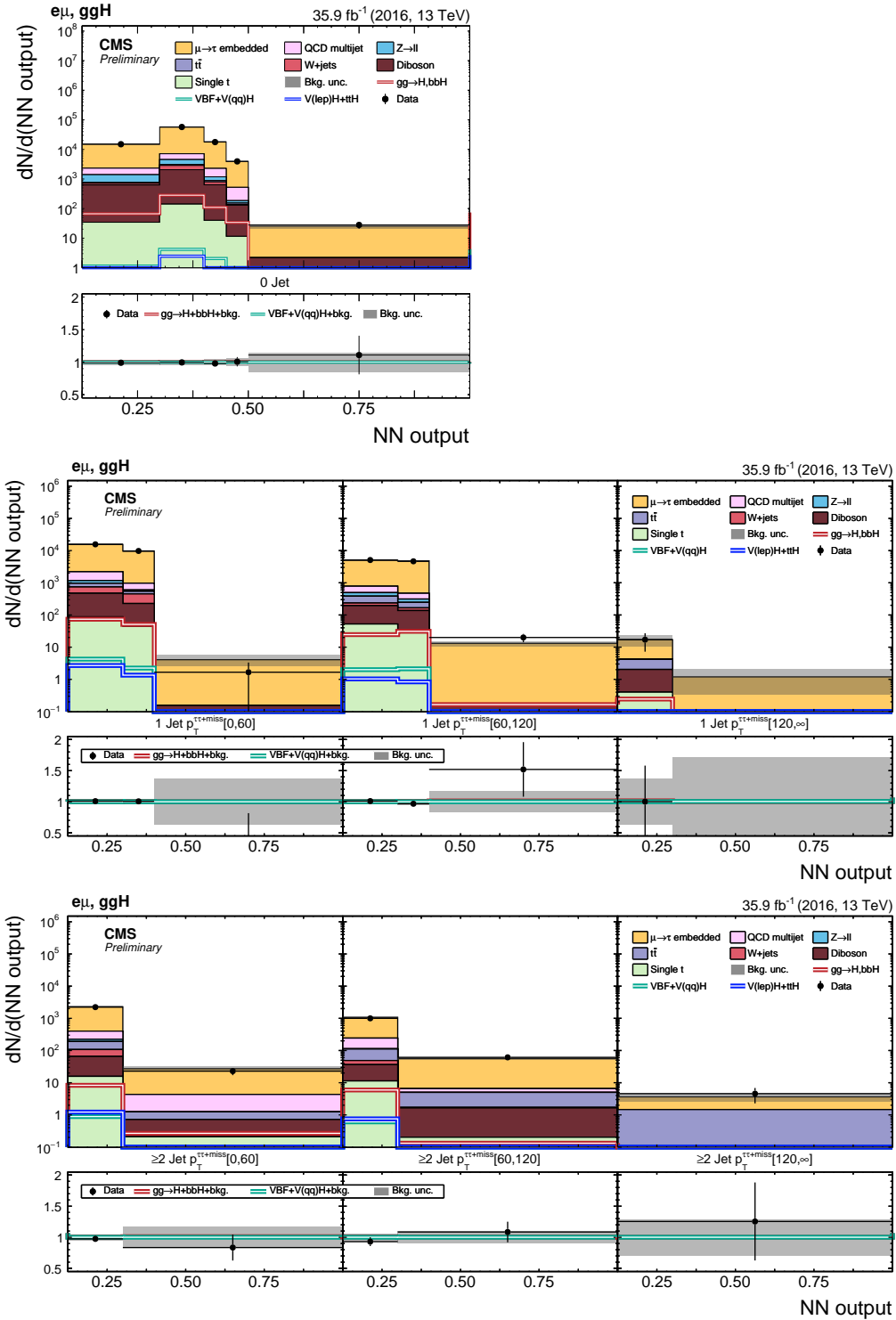
Figure A.21 – Distributions of the NN output in the event categories used for the determination of the stage-1 simplified template cross sections for VBF+V(qq)H in the eτ_h final state. Shown are the categories with VBF topology in bins of $p_\mathrm{T}^{\tau\tau\mathrm{jj+miss}}$ (top row) and the event category with V(qq)H topology, the "Rest" category and the event category with $p_\mathrm{T}^{\mathrm{j_1}} > 200\,\mathrm{GeV}$ (bottom row), on the 2016 dataset [12].
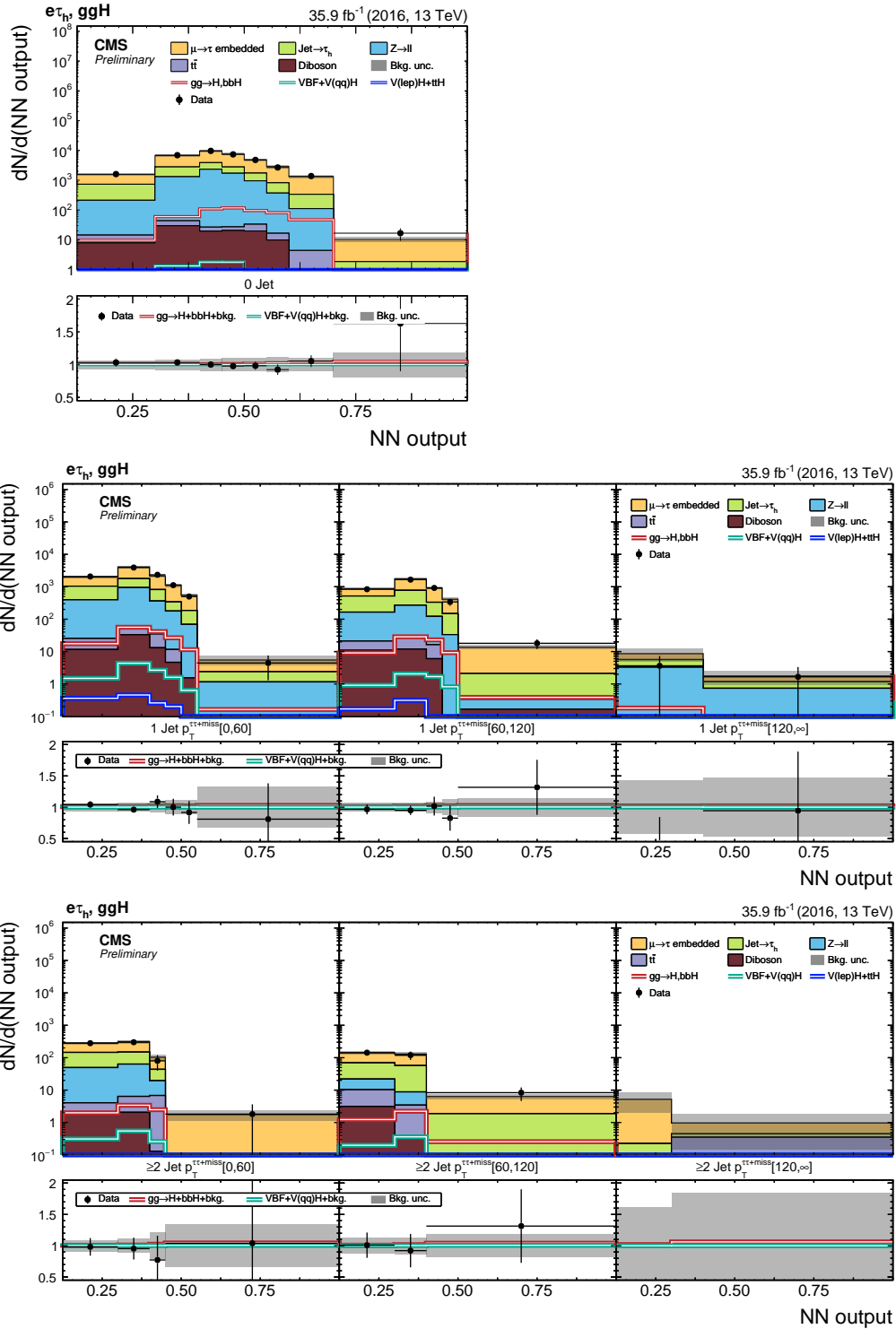
Figure A.22 – Distributions of the NN output in the event categories used for the determination of the stage-1 simplified template cross sections for VBF+V(qq)H in the $\mu\tau_h$ final state. Shown are the categories with VBF topology in bins of $p_T^{\tau\tau\text{jj}+\text{miss}}$ (top row) and the event category with V(qq)H topology, the "Rest" category and the event category with $p_T^{j_1} > 200$ GeV (bottom row), on the 2016 dataset [12].

Figure A.23 – Distributions of the NN output in the event categories used for the determination of the stage-1 simplified template cross sections for VBF+V(qq)H in the $\tau_h\tau_h$ final state. Shown are the categories with VBF topology in bins of $p_T^{\tau\tau jj+\text{miss}}$ (top row) and the event category with V(qq)H topology, the "Rest" category and the event category with $p_T^{j_1} > 200$ GeV (bottom row), on the 2016 dataset [12].
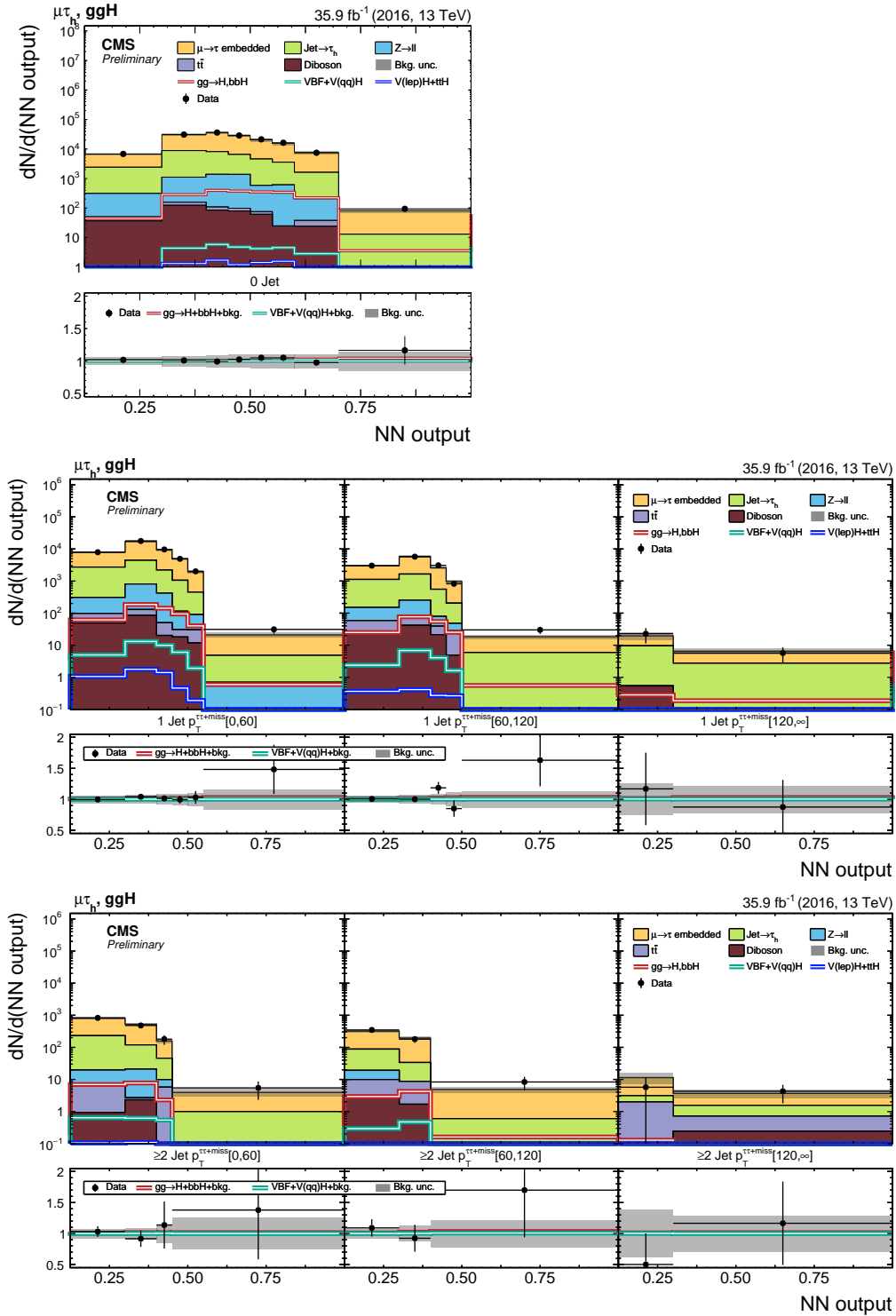
Figure A.24 – Distributions of the NN output in the event categories used for the determination of the stage-1 simplified template cross sections for gg → H, bbH in the eμ final state. Shown are the 0 Jet (top left), 1 Jet (centre row), and ≥ 2 Jet (bottom row) event categories in bins of $p_{\mathrm{T}}^{\tau\tau+\mathrm{miss}}$, on the 2016 dataset [12].
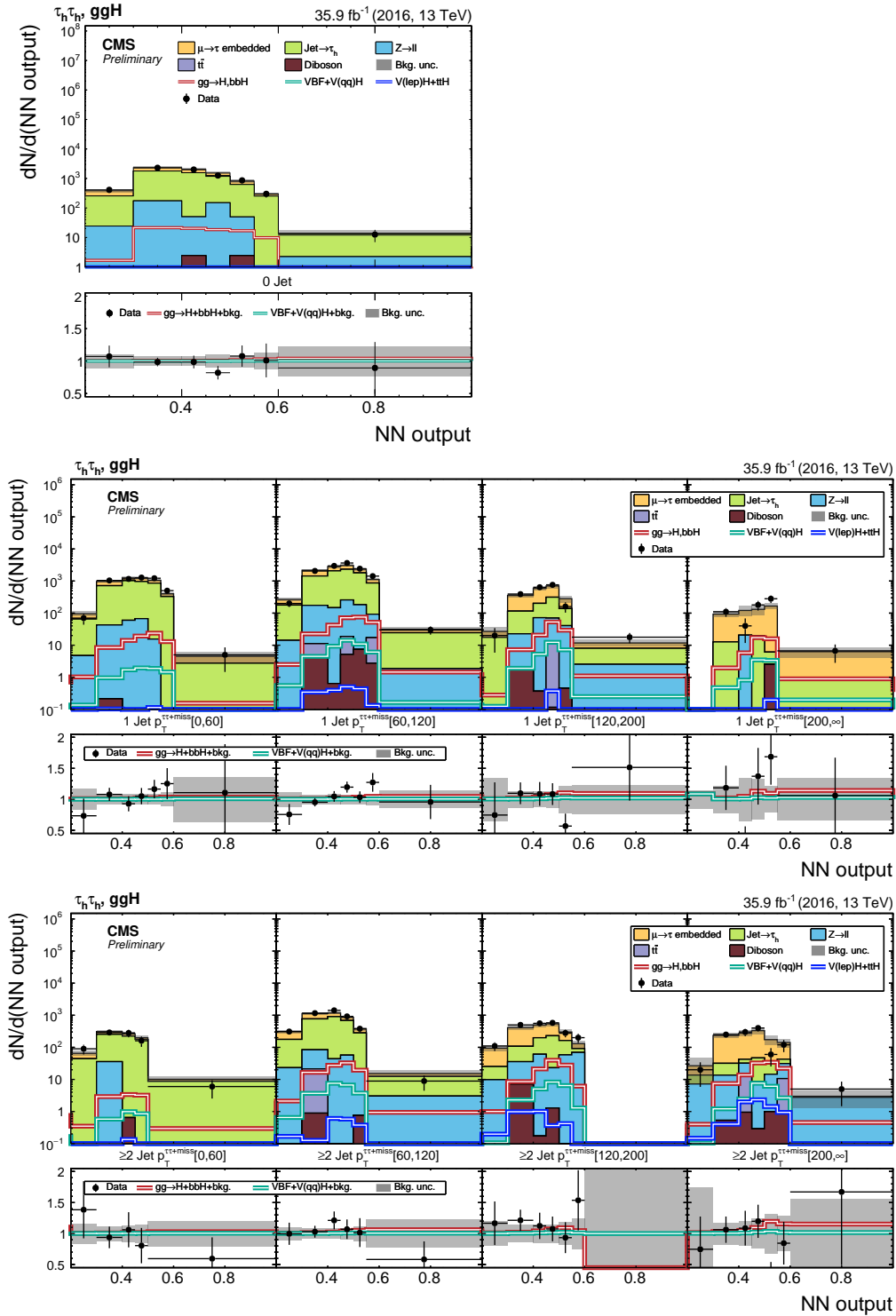
Figure A.25 − Distributions of the NN output in the event categories used for the determination of the stage-1 simplified template cross sections for $gg \rightarrow H, bbH$ in the $e\tau_h$ final state. Shown are the 0 Jet (top left), 1 Jet (centre row), and ≥ 2 Jet (bottom row) event categories in bins of $p_T^{\tau\tau+\text{miss}}$, on the 2016 dataset [12].

Figure A.26 – Distributions of the NN output in the event categories used for the determination of the stage-1 simplified template cross sections for $gg \to H, bbH$ in the $\mu\tau_h$ final state. Shown are the 0 Jet (top left), 1 Jet (centre row), and ≥ 2 Jet (bottom row) event categories in bins of $p_T^{\tau\tau+\text{miss}}$, on the 2016 dataset [12].
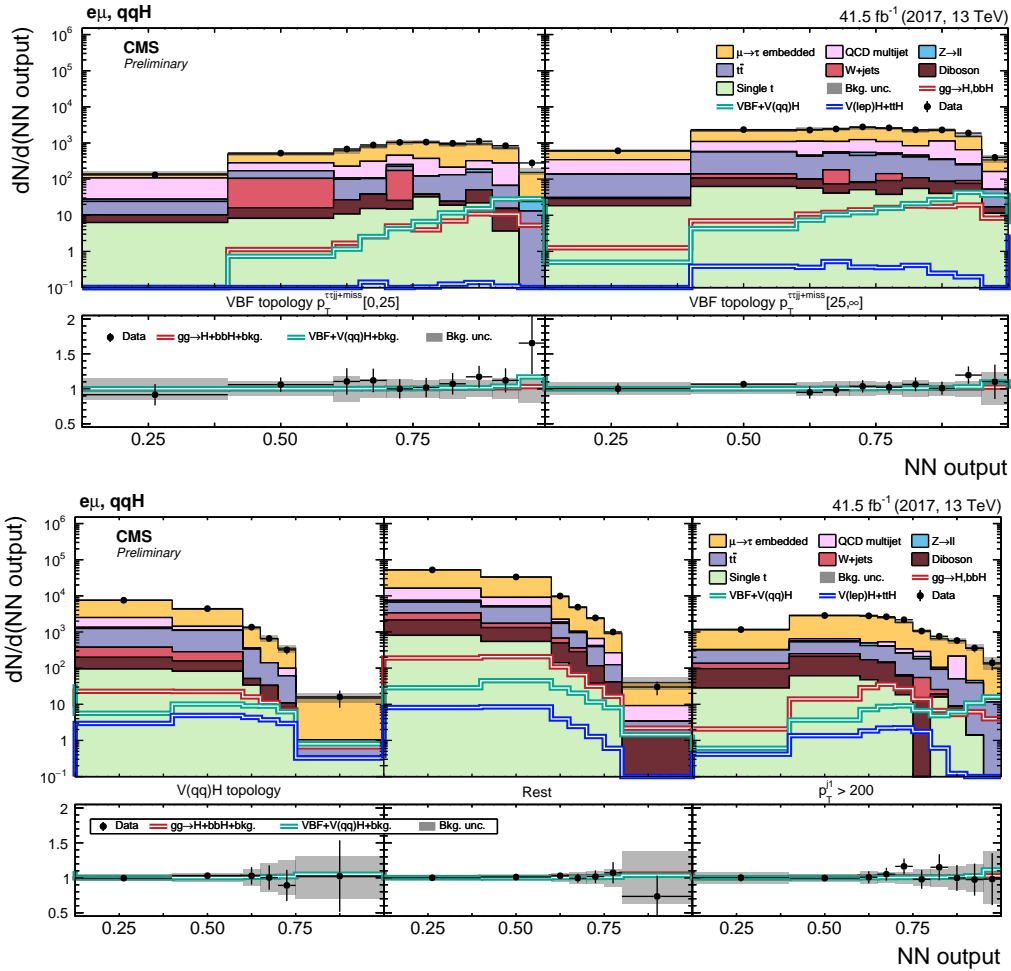
Figure A.27 – Distributions of the NN output in the event categories used for the determination of the stage-1 simplified template cross sections for $gg \to H, bbH$ in the $\tau_h \tau_h$ final state. Shown are the 0 Jet (top left), 1 Jet (centre row), and $\geq$ 2 Jet (bottom row) event categories in bins of $p_T^{\tau\tau+\mathrm{miss}}$, on the 2016 dataset [12].

Figure A.28 – Distributions of the NN output in the event categories used for the determination of the stage-1 simplified template cross sections for VBF+V(qq)H in the eμ final state. Shown are the categories with VBF topology in bins of $p_T^{\tau\tau jj+\text{miss}}$ (top row) and the event category with V(qq)H topology, the "Rest" category and the event category with $p_T^{j_1} > 200\,\text{GeV}$ (bottom row), on the 2017 dataset [12].
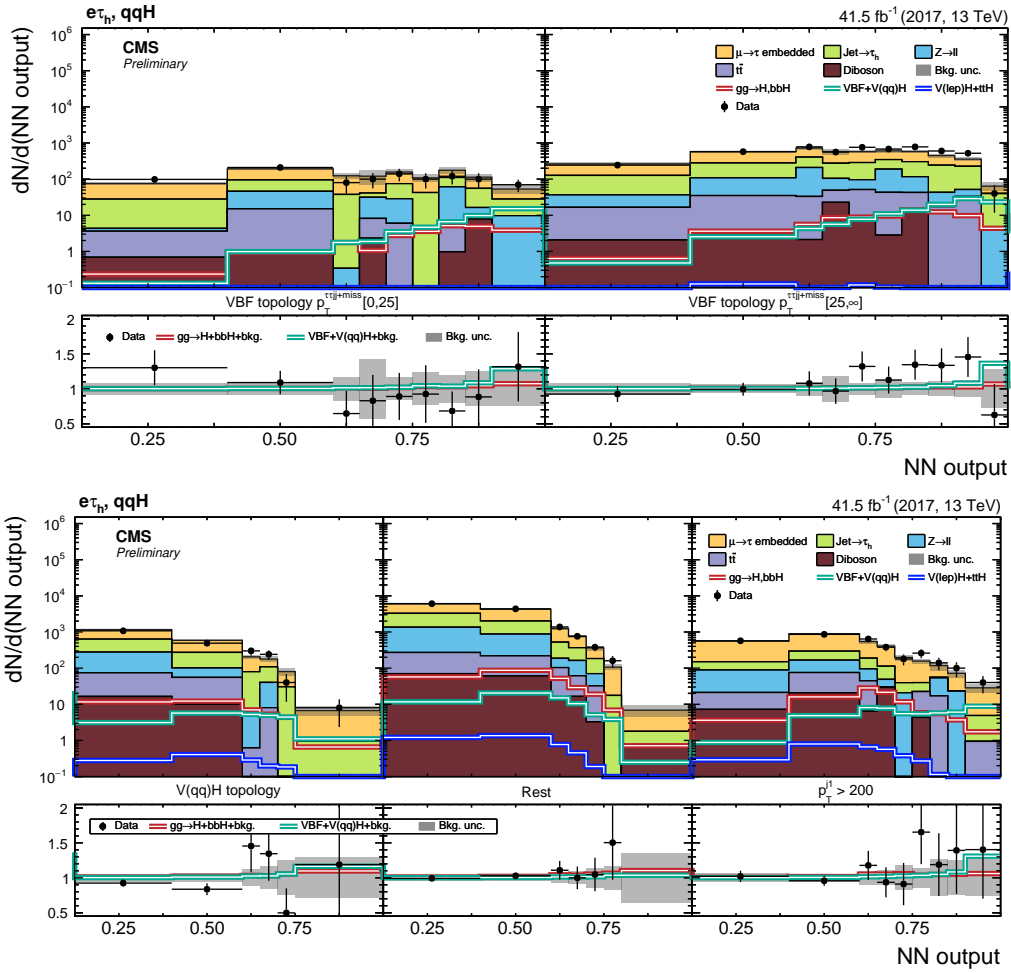
176

Figure A.29 – Distributions of the NN output in the event categories used for the determination of the stage-1 simplified template cross sections for VBF+V(qq)H in the eτ_h final state. Shown are the categories with VBF topology in bins of $p_T^{\tau\tau\mathrm{jj}+\mathrm{miss}}$ (top row) and the event category with V(qq)H topology, the "Rest" category and the event category with $p_T^{\mathrm{j}_1} > 200\,\mathrm{GeV}$ (bottom row), on the 2017 dataset [12].
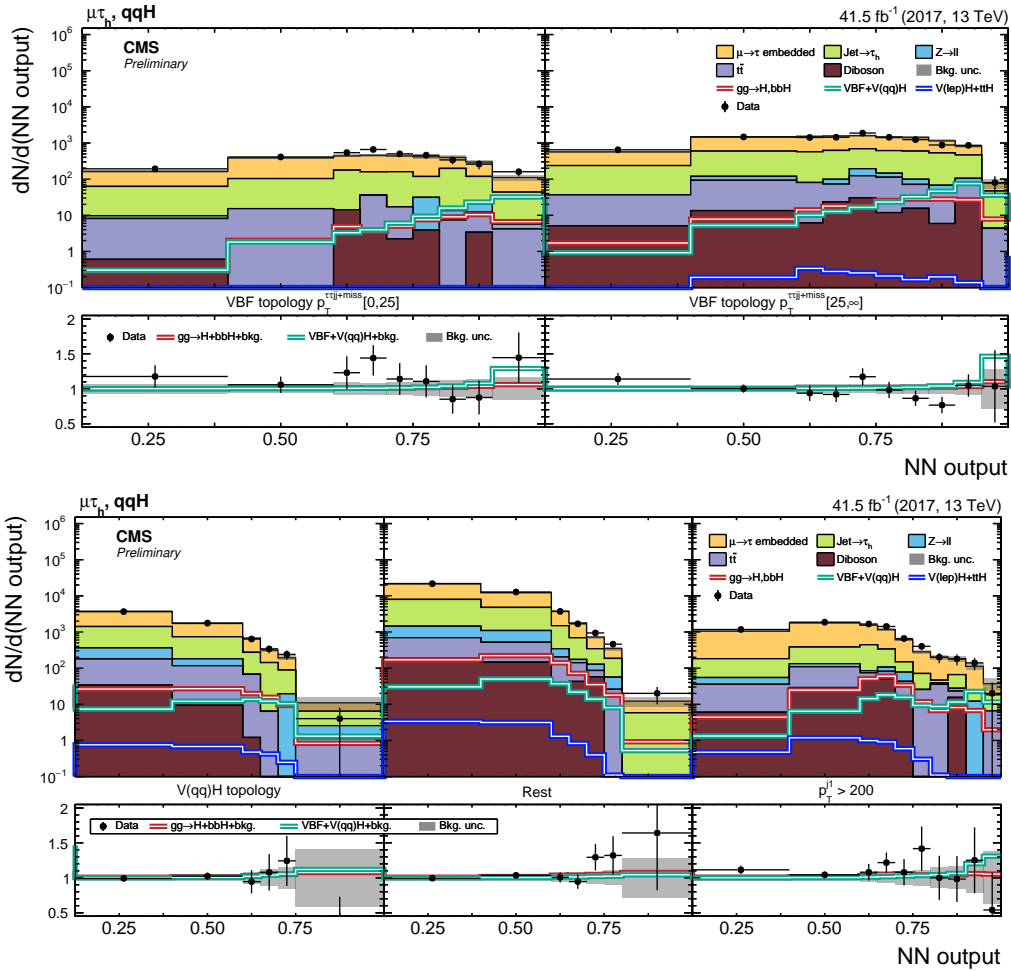
Figure A.30 – Distributions of the NN output in the event categories used for the determination of the stage-1 simplified template cross sections for VBF+V(qq)H in the $\mu\tau_{\rm h}$ final state. Shown are the categories with VBF topology in bins of $p_{\rm T}^{\tau\tau{\rm jj+miss}}$ (top row) and the event category with V(qq)H topology, the "Rest" category and the event category with $p_{\rm T}^{\rm j_1} > 200\,{\rm GeV}$ (bottom row), on the 2017 dataset [12].
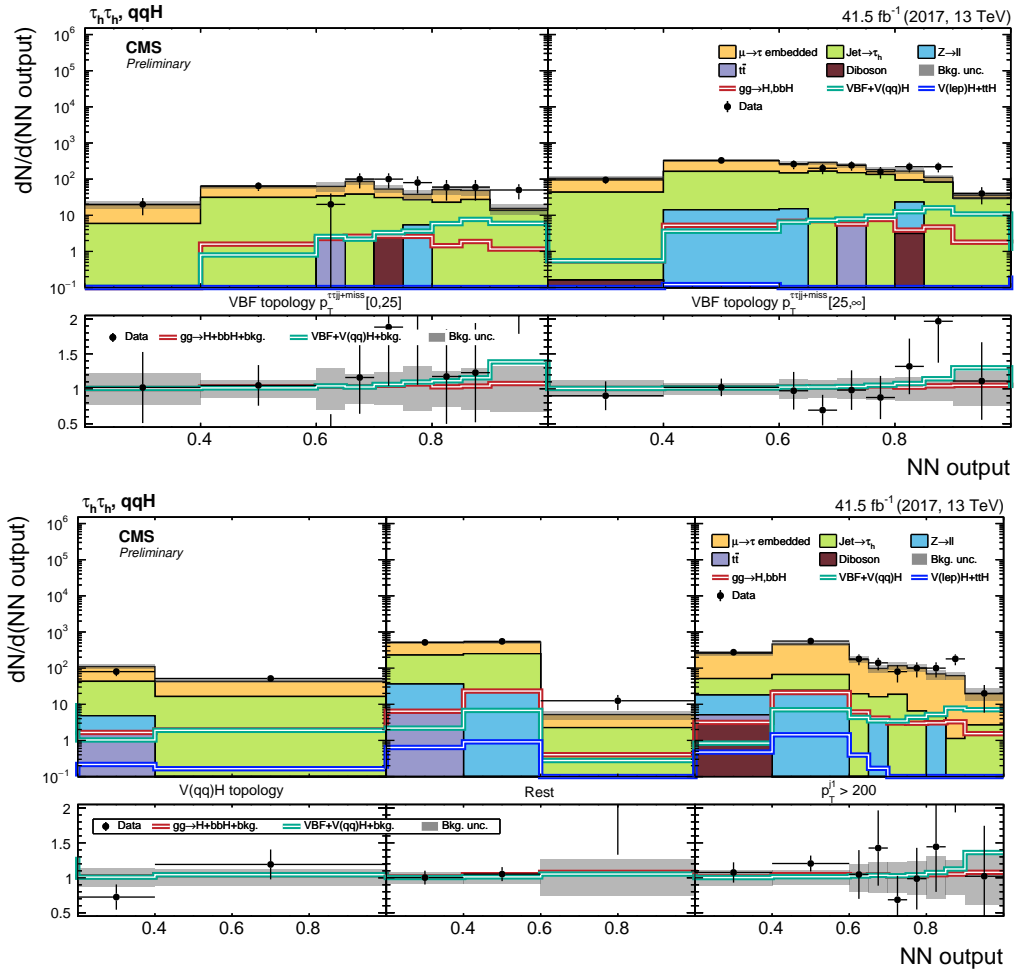
Figure A.31 – Distributions of the NN output in the event categories used for the determination of the stage-1 simplified template cross sections for VBF+V(qq)H in the $\tau_h\tau_h$ final state. Shown are the categories with VBF topology in bins of $p_T^{\tau\tau jj+\mathrm{miss}}$ (top row) and the event category with V(qq)H topology, the "Rest" category and the event category with $p_T^{j_1} > 200\,\mathrm{GeV}$ (bottom row), on the 2017 dataset [12].
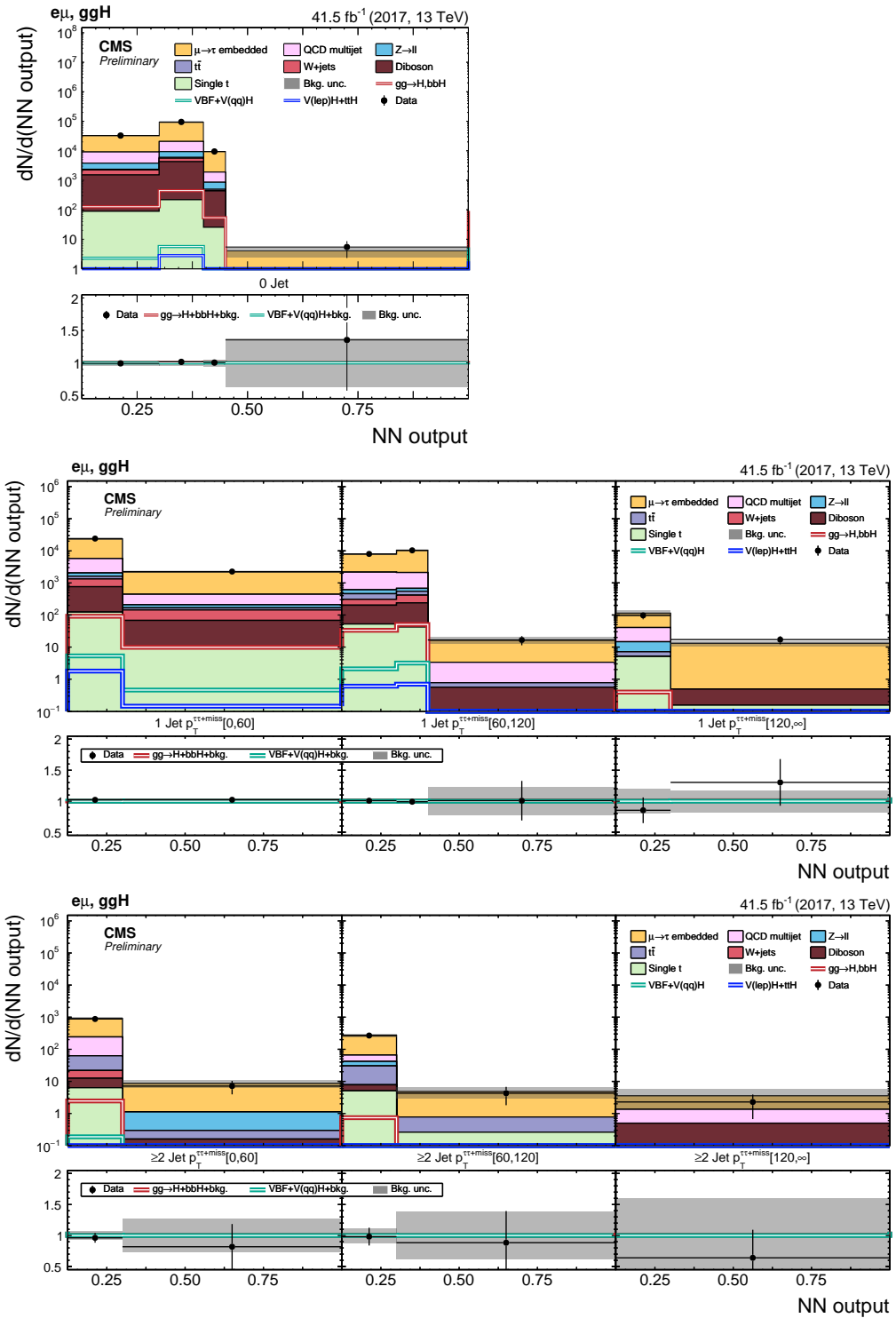
Figure A.32 – Distributions of the NN output in the event categories used for the determination of the stage-1 simplified template cross sections for gg → H, bbH in the eμ final state. Shown are the 0 Jet (top left), 1 Jet (centre row) , and ≥ 2 Jet (bottom row) event categories in bins of $p_T^{\tau\tau+\text{miss}}$, on the 2017 dataset [12].
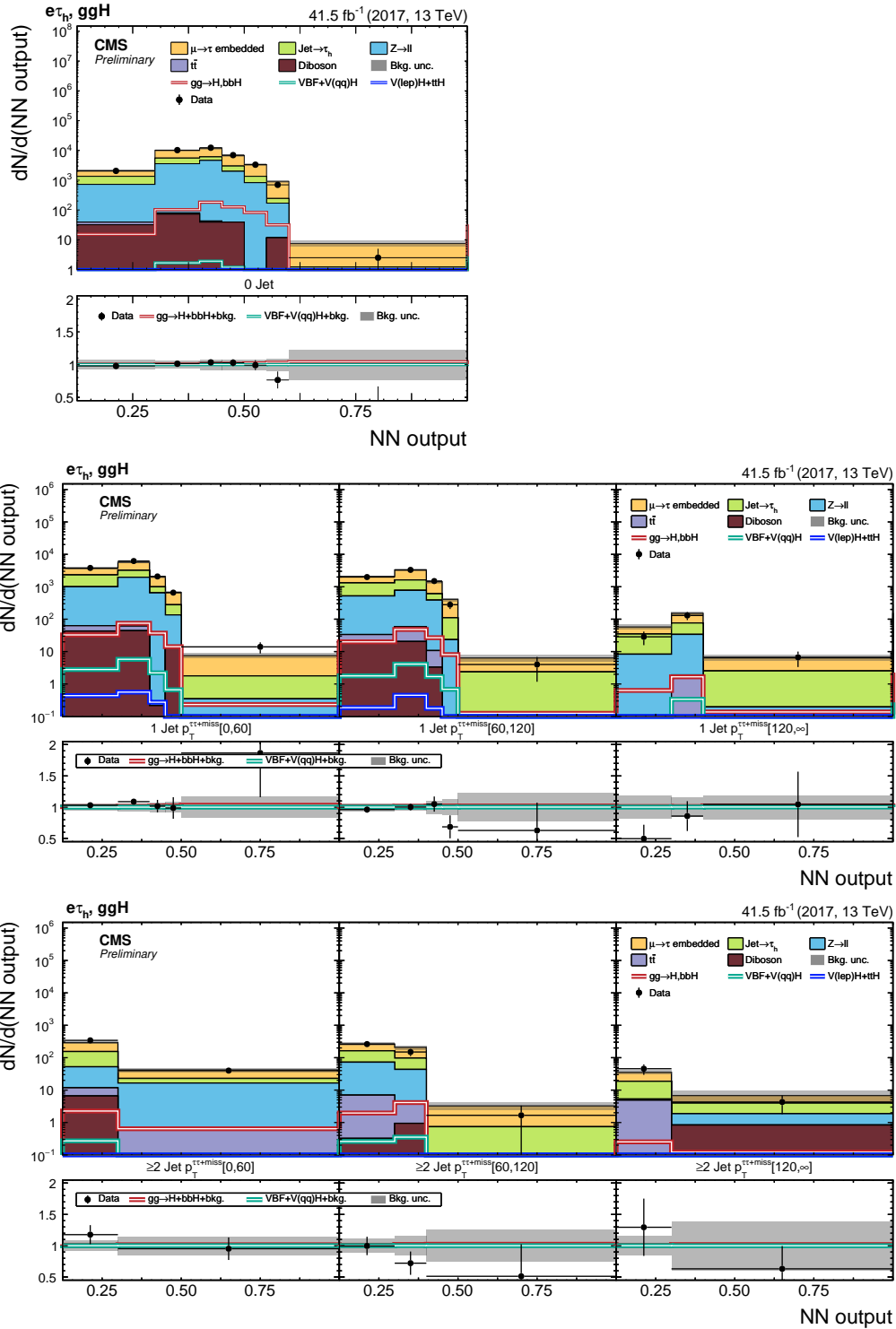
Figure A.33 – Distributions of the NN output in the event categories used for the determination of the stage-1 simplified template cross sections for $gg \rightarrow H, bbH$ in the $e\tau_h$ final state. Shown are the 0 Jet (top left), 1 Jet (centre row), and ≥ 2 Jet (bottom row) event categories in bins of $p_T^{\tau\tau+\text{miss}}$, on the 2017 dataset [12].

Figure A.34 – Distributions of the NN output in the event categories used for the determination of the stage-1 simplified template cross sections for $gg \to H, bbH$ in the $\mu\tau_h$ final state. Shown are the 0 Jet (top left), 1 Jet (centre row), and ≥ 2 Jet (bottom row) event categories in bins of $p_T^{\tau\tau+\text{miss}}$, on the 2017 dataset [12].
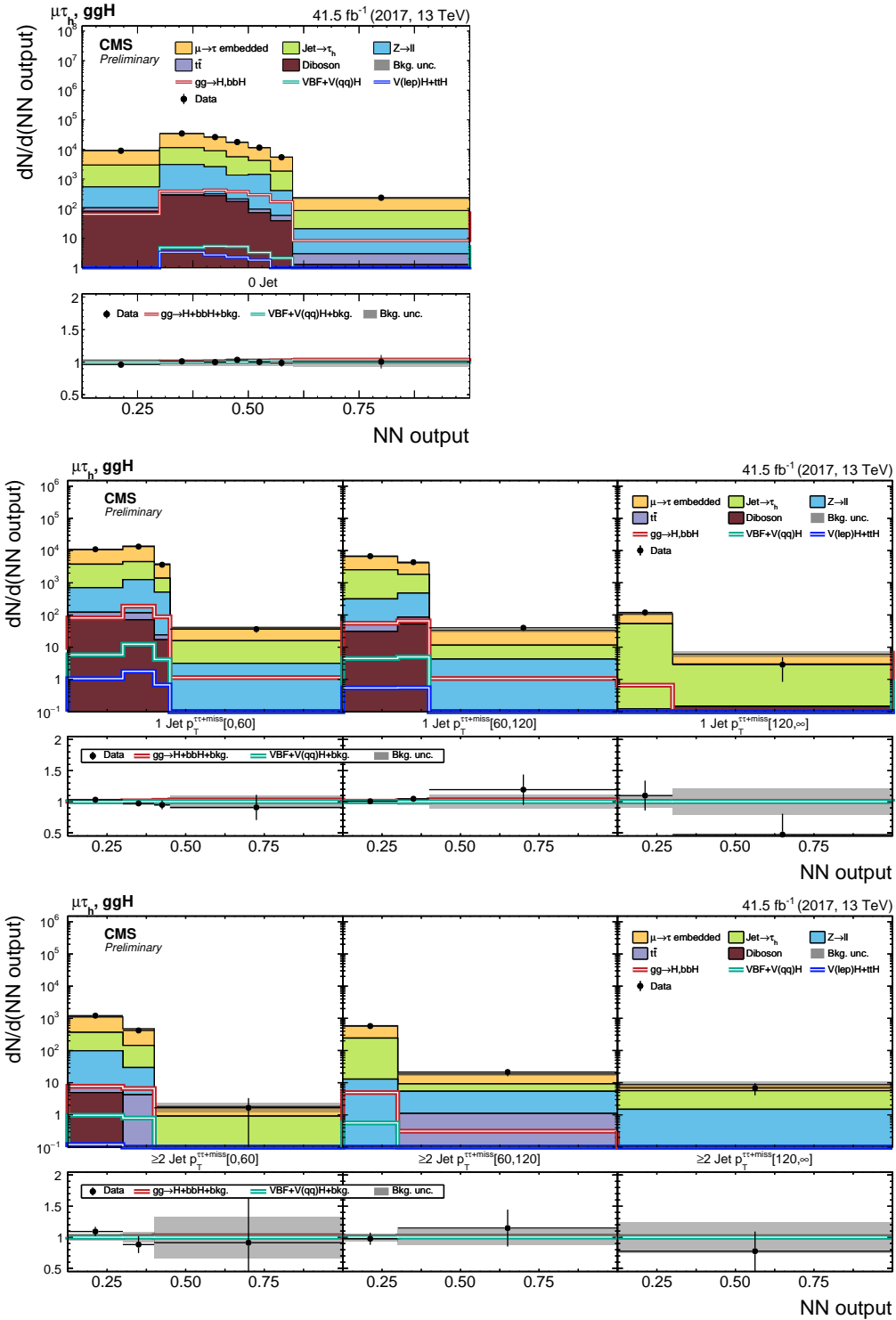
Figure A.35 − Distributions of the NN output in the event categories used for the determination of the stage-1 simplified template cross sections for $gg \to H, bbH$ in the $\tau_h \tau_h$ final state. Shown are the 0 Jet (top left), 1 Jet (centre row), and ≥ 2 Jet (bottom row) event categories in bins of $p_T^{\tau\tau+\text{miss}}$, on the 2017 dataset [12].

APPENDIX: ADDITIONAL NEUTRAL MSSM HIGGS BOSONS IN THE $\tau\tau$ FINAL STATE

B.1 Distribution of $m_{\mathrm{T}}^{\mathrm{tot}}$ in the Less Sensitive Subcategories.

Figure B.1 – Distribution of $m_{\mathrm{T}}^{\mathrm{tot}}$ in the eμ final state. The distributions are shown after the fit to data. Shown are the low-D_ζ (top) and high-D_ζ (bottom) event subcategory. The gray horizontal line in the upper panel of each subfigure indicates the change from logarithmic to linear scale on the vertical axis [13].

Figure B.2 – Distribution of $m_{\mathrm{T}}^{\mathrm{tot}}$ in the $\mathrm{e}\tau_{\mathrm{h}}$ (upper row) and $\mu\tau_{\mathrm{h}}$ (lower row) final states. The distributions are shown after the fit to data. In all cases the loose-m_{T} event subcategory is shown. The gray horizontal line in the upper panel of each subfigure indicates the change from logarithmic to linear scale on the vertical axis [13].

List of Abbreviations

AR	application region
BDT	Boosted Decision Tree
BEH	Brout-Englert-Higgs
BSM	beyond the standard model
CERN	European Organisation for Nuclear Research
CL	confidence level
CMS	Compact Muon Solenoid
CP	charge-parity
CR	control region
CSC	Cathode Strip Chambers
CTF	Combinatorial Track Finder
DA	Deterministic Annealing
$\mathbf{DR}_i$	determination region
DT	Drift Tubes
EB	ECAL Barrel
ECAL	Electromagnetic Calorimeter
EE	ECAL Endcap
EFT	effective field theories
$\boldsymbol{F}_\mathbf{F}$	Fake Factor
ggF	gluon fusion
GSF	Gaussian Sum Filter
HB	Hadron Barrel
HCAL	Hadron Calorimeter
HE	Hadron Endcap
HF	Hadron Forward
HLT	High-Level Trigger
HO	Hadron Outer
HPS	Hadron Plus Strips
ID	identification
KF	Kalman Filter

L1	Level-1
LEP	Large Electron-Positron Collider
LHC	Large Hadron Collider
LO	leading order
LS1	long shutdown 1
MC	Monte Carlo
ML	machine learning
MSSM	minimal supersymmetric extension of the standard model
MVA	Multivariate Analysis
NLO	next-to-LO
NN	neural network
NNLO	next-to-next-to-LO
PDF	parton density function
PF	Particle-Flow
POIs	parameters of interest
PPV	Positive Predictive Value
PS	Proton Synchrotron
PSB	Proton Synchrotron Booster
PU	pile-up
PV	primary-vertex
QCD	Quantum Chromodynamics
QED	Quantum Electrodynamics
RPC	Resistive Plate Chambers
SB	sideband region
SM	standard model
SPS	Super Proton Synchrotron
SR	signal region
STXS	simplified template cross sections
SUSY	supersymmetry
ttH	Higgs production in association with t quarks
VBF	vector boson fusion
VH	Higgs strahlung
VR	validation region

Bibliography

[1] ATLAS Collaboration. *"Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC"*. In: *Phys. Lett.* B716 (2012), pp. 1–29.
DOI: `10.1016/j.physletb.2012.08.020`. arXiv: `1207.7214`.

[2] CMS Collaboration. *"Observation of a New Boson at a Mass of 125 GeV with the CMS Experiment at the LHC"*. In: *Phys. Lett.* B716 (2012), pp. 30–61.
DOI: `10.1016/j.physletb.2012.08.021`. arXiv: `1207.7235`.

[3] CMS Collaboration. *"Observation of a New Boson with Mass Near 125 GeV in pp Collisions at $\sqrt{s} = 7$ and 8 TeV"*. In: *JHEP* 06 (2013), p. 081.
DOI: `10.1007/JHEP06(2013)081`. arXiv: `1303.4571`.

[4] P. W. Anderson. *"Plasmons, Gauge Invariance, and Mass"*. In: *Phys. Rev.* 130 (1 Apr. 1963), pp. 439–442.
DOI: `10.1103/PhysRev.130.439`.

[5] F. Englert and R. Brout. *"Broken Symmetry and the Mass of Gauge Vector Mesons"*. In: *Phys. Rev. Lett.* 13 (9 Aug. 1964), pp. 321–323.
DOI: `10.1103/PhysRevLett.13.321`.

[6] Peter W. Higgs. *"Broken Symmetries and the Masses of Gauge Bosons"*. In: *Phys. Rev. Lett.* 13 (1964), pp. 508–509.
DOI: `10.1103/PhysRevLett.13.508`.

[7] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble. *"Global Conservation Laws and Massless Particles"*. In: *Phys. Rev. Lett.* 13 (20 Nov. 1964), pp. 585–587.
DOI: `10.1103/PhysRevLett.13.585`.

[8] Stephen P. Martin. *"A Supersymmetry primer"*. In: (1997), pp. 1–98.
DOI: `10.1142/9789812839657_0001`. arXiv: `hep-ph/9709356`.

[9] ATLAS Collaboration. *"Cross-section measurements of the Higgs boson decaying into a pair of τ-leptons in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector"*. In: *Phys. Rev.* D99 (2019), p. 072001.
DOI: `10.1103/PhysRevD.99.072001`. arXiv: `1811.08856`.

[10] CMS Collaboration. *"Observation of the Higgs boson decay to a pair of τ leptons with the CMS detector"*. In: *Phys. Lett.* B779 (2018), pp. 283–316.
DOI: `10.1016/j.physletb.2018.02.004`. arXiv: `1708.00373`.

[11] LHC Higgs Cross Section Working Group Collaboration. *"Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector"*. In: (2016), pp. 437–449.

DOI: `10.23731/CYRM-2017-002`. arXiv: `1610.07922`.

[12] CMS Collaboration. Measurement of Higgs boson production and decay to the $\tau\tau$ final state. CMS-PAS-HIG-18-032. 2019. URL: `https://cds.cern.ch/record/2668685`.

[13] CMS Collaboration. *"Search for additional neutral MSSM Higgs bosons in the $\tau\tau$ final state in proton-proton collisions at $\sqrt{s} = 13$ TeV"*. In: *JHEP* 09 (2018), p. 007.
DOI: `10.1007/JHEP09(2018)007`. arXiv: `1803.06553`.

[14] W Noel Cottingham and Derek A. Greenwood. An introduction to the standard model of particle physics. 2nd. Cambridge University Press.

[15] C. Burgess and Guy Moore. *"The Standard Model: A Primer"*. In: (Jan. 2007).
DOI: `10.1017/CBO9780511819698`.

[16] S. L. Glashow. *"Partial Symmetries of Weak Interactions"*. In: *Nucl. Phys.* 22 (1961), pp. 579–588.
DOI: `10.1016/0029-5582(61)90469-2`.

[17] Steven Weinberg. *"A Model of Leptons"*. In: *Phys. Rev. Lett.* 19 (1967), pp. 1264–1266.
DOI: `10.1103/PhysRevLett.19.1264`.

[18] Abdus Salam. *"Weak and Electromagnetic Interactions"*. In: *Conf. Proc.* C680519 (1968), pp. 367–377.

[19] Particle Data Group Collaboration. *"Review of Particle Physics"*. In: *Phys. Rev. D* 98 (2018 and 2019 upgrade), p. 030001.
DOI: `10.1103/PhysRevD.98.030001`.

[20] KATRIN Collaboration. *"An improved upper limit on the neutrino mass from a direct kinematic method by KATRIN"*. In: (2019). arXiv: `1909.06048`.

[21] Paul A. M. Dirac. *"Quantum theory of emission and absorption of radiation"*. In: *Proc. Roy. Soc. Lond.* A114 (1927), p. 243.
DOI: `10.1098/rspa.1927.0039`.

[22] Steven Weinberg. *"New Approach to the Renormalization Group"*. In: *Phys. Rev. D* 8 (10 Nov. 1973), pp. 3497–3509.
DOI: `10.1103/PhysRevD.8.3497`.

[23] G. 't Hooft. *"Dimensional regularization and the renormalization group"*. In: *Nuclear Physics B* 61 (1973), pp. 455–468.
DOI: `https://doi.org/10.1016/0550-3213(73)90376-3`.

[24] Martin L. Perl et al. *"Evidence for Anomalous Lepton Production in e^+-e^- Annihilation"*. In: *Phys. Rev. Lett.* 35 (1975), pp. 1489–1492.
DOI: `10.1103/PhysRevLett.35.1489`.

[25]   Martin L. Perl et al. *"Properties of Anomalous e μ Events Produced in e^+-e^- Annihilation"*. In: *Phys. Lett.* 63B (1976), p. 466.
        DOI: `10.1016/0370-2693(76)90399-3`.

[26]   PLUTO Collaboration. *"Anomalous Muon Production in e^+-e^- Annihilation as Evidence for Heavy Leptons"*. In: *Phys. Lett.* 68B (1977), p. 297.
        DOI: `10.1016/0370-2693(77)90292-1`.

[27]   PLUTO Collaboration. *"Evidence for Heavy Leptons from Anomalous μ e Production in e^+-e^- Annihilation"*. In: *Phys. Lett.* 68B (1977), pp. 301–304.
        DOI: `10.1016/0370-2693(77)90293-3`.

[28]   DASP Collaboration. *"On the Origin of Inclusive electron Events in e^+-e^- Annihilation Between 3.6-GeV and 5.2-GeV"*. In: *Phys. Lett.* 70B (1977), p. 125.
        DOI: `10.1016/0370-2693(77)90360-4`.

[29]   Roger Wolf. The Higgs Boson Discovery at the Large Hadron Collider. Vol. 264. Springer, 2015.
        DOI: `10.1007/978-3-319-18512-5`.

[30]   Jeffrey Goldstone, Abdus Salam, and Steven Weinberg. *"Broken Symmetries"*. In: *Phys. Rev.* 127 (1962), pp. 965–970.
        DOI: `10.1103/PhysRev.127.965`.

[31]   LEP Working Group for Higgs boson searches, ALEPH, DELPHI, L3, OPAL Collaboration. *"Search for the standard model Higgs boson at LEP"*. In: *Phys. Lett.* B565 (2003), pp. 61–75.
        DOI: `10.1016/S0370-2693(03)00614-2`. arXiv: `hep-ex/0306033`.

[32]   CDF, D0 Collaboration. *"Updated Combination of CDF and D0 Searches for Standard Model Higgs Boson Production with up to 10.0 fb^{-1} of Data"*. In: 2012. arXiv: `1207.0449`. URL: `http://lss.fnal.gov/cgi-bin/find_paper.pl?conf-12-318`.

[33]   CMS Collaboration. *"Combined results of searches for the standard model Higgs boson in pp collisions at $\sqrt{s} = 7$ TeV"*. In: *Phys. Lett.* B710 (2012), pp. 26–48.
        DOI: `10.1016/j.physletb.2012.02.064`. arXiv: `1202.1488`.

[34]   ATLAS Collaboration. *"Combined search for the Standard Model Higgs boson using up to 4.9 fb^{-1} of pp collision data at $\sqrt{s} = 7$ TeV with the ATLAS detector at the LHC"*. In: *Phys. Lett.* B710 (2012), pp. 49–66.
        DOI: `10.1016/j.physletb.2012.02.044`. arXiv: `1202.1408`.

[35]   CMS Collaboration. *"Measurement of the properties of a Higgs boson in the four-lepton final state"*. In: *Phys. Rev.* D89.9 (2014), p. 092007.
        DOI: `10.1103/PhysRevD.89.092007`. arXiv: `1312.5353`.

[36] ATLAS Collaboration. *"Measurement of Higgs boson production in the diphoton decay channel in pp collisions at center-of-mass energies of 7 and 8 TeV with the ATLAS detector"*. In: *Phys. Rev.* D90.11 (2014), p. 112015. DOI: 10.1103/PhysRevD.90.112015. arXiv: 1408.7084.

[37] ATLAS Collaboration. *"Measurements of Higgs boson production and couplings in the four-lepton channel in pp collisions at center-of-mass energies of 7 and 8 TeV with the ATLAS detector"*. In: *Phys. Rev.* D91.1 (2015), p. 012006. DOI: 10.1103/PhysRevD.91.012006. arXiv: 1408.5191.

[38] CMS Collaboration. *"Observation of the diphoton decay of the Higgs boson and measurement of its properties"*. In: *Eur. Phys. J.* C74.10 (2014), p. 3076. DOI: 10.1140/epjc/s10052-014-3076-z. arXiv: 1407.0558.

[39] ATLAS Collaboration. *"Evidence for the spin-0 nature of the Higgs boson using ATLAS data"*. In: *Phys. Lett.* B726 (2013), pp. 120–144. DOI: 10.1016/j.physletb.2013.08.026. arXiv: 1307.1432.

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

[41] LHC Higgs Cross Section Working Group Collaboration. *"LHC HXSWG interim recommendations to explore the coupling structure of a Higgs-like particle"*. In: (2012). arXiv: 1209.0040.

[42] ATLAS, CMS Collaboration. *"Combined Measurement of the Higgs Boson Mass in pp Collisions at $\sqrt{s} = 7$ and 8 TeV with the ATLAS and CMS Experiments"*. In: *Phys. Rev. Lett.* 114 (2015), p. 191803. DOI: 10.1103/PhysRevLett.114.191803. arXiv: 1503.07589.

[43] CMS Collaboration. *"Measurements of the Higgs boson width and anomalous HVV couplings from on-shell and off-shell production in the four-lepton final state"*. In: *Phys. Rev.* D99.11 (2019), p. 112003. DOI: 10.1103/PhysRevD.99.112003. arXiv: 1901.00174.

[44] L. D. Landau. *"On the angular momentum of a system of two photons"*. In: *Dokl. Akad. Nauk Ser. Fiz.* 60.2 (1948), pp. 207–209. DOI: 10.1016/B978-0-08-010586-4.50070-5.

[45] C. N. Yang. *"Selection Rules for the Dematerialization of a Particle into Two Photons"*. In: *Phys. Rev.* 77 (2 Jan. 1950), pp. 242–245. DOI: 10.1103/PhysRev.77.242.

[46] CMS Collaboration. *"Study of the Mass and Spin-Parity of the Higgs boson candidate via its decays to Z boson pairs"*. In: *Phys. Rev. Lett.* 110.8 (2013), p. 081803.

DOI: `10.1103/PhysRevLett.110.081803`. arXiv: `1212.6639`.

[47] CMS Collaboration. *"Measurements of properties of the Higgs boson decaying to a W boson pair in pp collisions at $\sqrt{s} = 13$ TeV"*. In: *Phys. Lett.* B791 (2019), p. 96.
DOI: `10.1016/j.physletb.2018.12.073`. arXiv: `1806.05246`.

[48] ATLAS Collaboration. *"Measurements of Higgs boson production and couplings in diboson final states with the ATLAS detector at the LHC"*. In: *Phys. Lett.* B726 (2013), pp. 88–119.
DOI: `10.1016/j.physletb.2014.05.011`. arXiv: `1307.1427`.

[49] CMS Collaboration. *"Observation of Higgs boson decay to bottom quarks"*. In: *Phys. Rev. Lett.* 121.12 (2018), p. 121801.
DOI: `10.1103/PhysRevLett.121.121801`. arXiv: `1808.08242`.

[50] ATLAS Collaboration. *"Observation of $H \rightarrow b\bar{b}$ decays and VH production with the ATLAS detector"*. In: *Phys. Lett.* B786 (2018), pp. 59–86.
DOI: `10.1016/j.physletb.2018.09.013`. arXiv: `1808.08238`.

[51] CMS Collaboration. *"Observation of $t\bar{t}H$ production"*. In: *Phys. Rev. Lett.* 120.23 (2018), p. 231801.
DOI: `10.1103/PhysRevLett.120.231801`. arXiv: `1804.02610`.

[52] ATLAS Collaboration. *"Observation of Higgs boson production in association with a top quark pair at the LHC with the ATLAS detector"*. In: *Phys. Lett.* B784 (2018), pp. 173–191.
DOI: `10.1016/j.physletb.2018.07.035`. arXiv: `1806.00425`.

[53] ATLAS Collaboration. *"Search for the dimuon decay of the Higgs boson in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector"*. In: *Phys. Rev. Lett.* 119.5 (2017), p. 051802.
DOI: `10.1103/PhysRevLett.119.051802`. arXiv: `1705.04582`.

[54] ATLAS Collaboration. *"Search for the Decay of the Higgs Boson to Charm Quarks with the ATLAS Experiment"*. In: *Phys. Rev. Lett.* 120.21 (2018), p. 211802.
DOI: `10.1103/PhysRevLett.120.211802`. arXiv: `1802.04329`.

[55] CMS Collaboration. *"Search for the Higgs boson decaying to two muons in proton-proton collisions at $\sqrt{s} = 13$ TeV"*. In: *Phys. Rev. Lett.* 122.2 (2019), p. 021801.
DOI: `10.1103/PhysRevLett.122.021801`. arXiv: `1807.06325`.

[56] CMS Collaboration. Search for the standard model Higgs boson decaying to charm quarks. Tech. rep. CMS-PAS-HIG-18-031. Geneva: CERN, 2019. URL: `https://cds.cern.ch/record/2682638`.

[57] Gianfranco Bertone, Dan Hooper, and Joseph Silk. *"Particle dark matter: Evidence, candidates and constraints"*. In: *Phys. Rept.* 405 (2005), pp. 279–390.
DOI: `10.1016/j.physrep.2004.08.031`. arXiv: `hep-ph/0404175`.

[58] Super-Kamiokande Collaboration. *"Evidence for oscillation of atmospheric neutrinos"*. In: *Phys. Rev. Lett.* 81 (1998), pp. 1562–1567.
DOI: 10.1103/PhysRevLett.81.1562. arXiv: hep-ex/9807003.

[59] SNO Collaboration. *"Measurement of the rate of $\nu_e + d \to p + p + e^-$ interactions produced by ^8B solar neutrinos at the Sudbury Neutrino Observatory"*. In: *Phys. Rev. Lett.* 87 (2001), p. 071301.
DOI: 10.1103/PhysRevLett.87.071301. arXiv: nucl-ex/0106015.

[60] Marcela Carena and Howard E. Haber. *"Higgs Boson Theory and Phenomenology"*. In: *Prog. Part. Nucl. Phys.* 50 (2003), pp. 63–152.
DOI: 10.1016/S0146-6410(02)00177-1. arXiv: hep-ph/0208209.

[61] M. Carena et al. *"MSSM Higgs Boson Searches at the LHC: Benchmark Scenarios after the Discovery of a Higgs-like Particle"*. In: *Eur. Phys. J.* C73.9 (2013), p. 2552.
DOI: 10.1140/epjc/s10052-013-2552-1. arXiv: 1302.7033.

[62] A. Djouadi et al. *"The post-Higgs MSSM scenario: Habemus MSSM?"* In: *Eur. Phys. J.* C73 (2013), p. 2650.
DOI: 10.1140/epjc/s10052-013-2650-0. arXiv: 1307.5205.

[63] Emanuele Bagnaschi et al. *"MSSM Higgs Boson Searches at the LHC: Benchmark Scenarios for Run 2 and Beyond"*. In: *Eur. Phys. J.* C79.7 (2019), p. 617.
DOI: 10.1140/epjc/s10052-019-7114-8. arXiv: 1808.07542.

[64] Rick S. Gupta, Heidi Rzehak, and James D. Wells. *"How well do we need to measure Higgs boson couplings?"* In: *Phys. Rev.* D86 (2012), p. 095001.
DOI: 10.1103/PhysRevD.86.095001. arXiv: 1206.3560.

[65] Matthew Bowen, Yanou Cui, and James D. Wells. *"Narrow trans-TeV Higgs bosons and H —> hh decays: Two LHC search paths for a hidden sector Higgs boson"*. In: *JHEP* 03 (2007), p. 036.
DOI: 10.1088/1126-6708/2007/03/036. arXiv: hep-ph/0701035.

[66] Robert M. Schabinger and James D. Wells. *"A Minimal spontaneously broken hidden sector and its impact on Higgs boson physics at the large hadron collider"*. In: *Phys. Rev.* D72 (2005), p. 093007.
DOI: 10.1103/PhysRevD.72.093007. arXiv: hep-ph/0509209.

[67] N. Arkani-Hamed et al. *"The Littlest Higgs"*. In: *JHEP* 07 (2002), p. 034.
DOI: 10.1088/1126-6708/2002/07/034. arXiv: hep-ph/0206021.

[68] Roberto Contino, Yasunori Nomura, and Alex Pomarol. *"Higgs as a holographic pseudoGoldstone boson"*. In: *Nucl. Phys.* B671 (2003), pp. 148–174.
DOI: 10.1016/j.nuclphysb.2003.08.027. arXiv: hep-ph/0306259.

[69] Kaustubh Agashe, Roberto Contino, and Alex Pomarol. *"The Minimal composite Higgs model"*. In: *Nucl. Phys.* B719 (2005), pp. 165–187.

DOI: `10.1016/j.nuclphysb.2005.04.035`. arXiv: `hep-ph/0412089`.

[70]   ATLAS Collaboration. *"Combined measurements of Higgs boson production and decay using up to 80 fb^{-1} of proton-proton collision data at $\sqrt{s}$ = 13 TeV collected with the ATLAS experiment"*. In: (2019). arXiv: `1909.02845`.

[71]   Lyndon Evans and Philip Bryant. *"LHC Machine"*. In: *JINST* 3 (2008), S08001. DOI: `10.1088/1748-0221/3/08/S08001`.

[72]   Esma Mobs. *"The CERN accelerator complex."* In: (Aug. 2018). General Photo. URL: `https://cds.cern.ch/record/2636343`.

[73]   J. P. Blewett. *"200-GeV Intersecting Storage Accelerators"*. In: *eConf* C710920 (1971), p. 501.

[74]   ALICE Collaboration. *"The ALICE experiment at the CERN LHC"*. In: *JINST* 3 (2008), S08002. DOI: `10.1088/1748-0221/3/08/S08002`.

[75]   *"The ATLAS Experiment at the CERN Large Hadron Collider"*. In: *JINST* 3 (2008), S08003. DOI: `10.1088/1748-0221/3/08/S08003`.

[76]   *"The CMS Experiment at the CERN LHC"*. In: *JINST* 3 (2008), S08004. DOI: `10.1088/1748-0221/3/08/S08004`.

[77]   LHCb Collaboration. *"The LHCb Detector at the LHC"*. In: *JINST* 3 (2008), S08005. DOI: `10.1088/1748-0221/3/08/S08005`.

[78]   CMS Collaboration. CMS Physics. CERN-LHCC-2006-001, CMS-TDR-8-1. 2006. URL: `https://cds.cern.ch/record/922757`.

[79]   M. Shopova et al. *"Performance of Resistive Plate Chambers installed during the first long shutdown of the CMS experiment"*. In: arXiv: `1605.06798`.

[80]   *"CMS Technical Design Report for the Level-1 Trigger Upgrade"*. In: CERN-LHCC-2013-011. CMS-TDR-12 (2013). URL: `https://cds.cern.ch/record/1556311`.

[81]   John C. Collins and Davison E. Soper. *"The Theorems of Perturbative QCD"*. In: *Ann. Rev. Nucl. Part. Sci.* 37 (1987), pp. 383–409. DOI: `10.1146/annurev.ns.37.120187.002123`.

[82]   Andy Buckley et al. *"General-purpose event generators for LHC physics"*. In: *Phys. Rept.* 504 (2011), pp. 145–233. DOI: `10.1016/j.physrep.2011.03.005`. arXiv: `1101.2599`.

[83]   GEANT4 Collaboration. *"GEANT4: A Simulation toolkit"*. In: *Nucl. Instrum. Meth.* A506 (2003), pp. 250–303. DOI: `10.1016/S0168-9002(03)01368-8`.

[84] CMS Collaboration. *"Study of the underlying event in top quark pair production in* pp *collisions at 13 TeV"*. In: *Eur. Phys. J.* C79.2 (2019), p. 123.
DOI: 10.1140/epjc/s10052-019-6620-z. arXiv: 1807.02810.

[85] CMS Collaboration. *"Particle-flow reconstruction and global event description with the CMS detector"*. In: *JINST* 12.10 (2017), P10003.
DOI: 10.1088/1748-0221/12/10/P10003. arXiv: 1706.04965.

[86] CMS Collaboration. *"Description and performance of track and primary-vertex reconstruction with the CMS tracker"*. In: *JINST* 9.10 (2014), P10009.
DOI: 10.1088/1748-0221/9/10/P10009. arXiv: 1405.6569.

[87] R. Frühwirth. *"Application of Kalman filtering to track and vertex fitting"*. In: *Nucl. Instrum. Meth.* A262 (1987), pp. 444–450.
DOI: 10.1016/0168-9002(87)90887-4.

[88] Kenneth Rose. *"Deterministic annealing for clustering, compression, classification, regression, and related optimization problems"*. In: *Proceedings of the IEEE* 86 (Dec. 1998), pp. 2210–2239.
DOI: 10.1109/5.726788.

[89] R Frühwirth, Wolfgang Waltenberger, and Pascal Vanlaer. Adaptive Vertex Fitting. CMS-NOTE-2007-008. Geneva: CERN, Mar. 2007. URL: https://cds.cern.ch/record/1027031.

[90] CMS Collaboration. *"Performance of Electron Reconstruction and Selection with the CMS Detector in Proton-Proton Collisions at $\sqrt{s} = 8$ TeV"*. In: *JINST* 10.06 (2015), P06005.
DOI: 10.1088/1748-0221/10/06/P06005. arXiv: 1502.02701.

[91] S. Baffioni et al. *"Electron reconstruction in CMS"*. In: *The European Physical Journal C* 49.4 (Mar. 2007), pp. 1099–1116.
DOI: 10.1140/epjc/s10052-006-0175-5.

[92] W Adam et al. *"Reconstruction of electrons with the Gaussian-sum filter in the CMS tracker at the LHC"*. In: *Journal of Physics G: Nuclear and Particle Physics* 31.9 (July 2005), N9–N20.
DOI: 10.1088/0954-3899/31/9/n01.

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

[94] Gavin P. Salam. *"Towards Jetography"*. In: *Eur. Phys. J.* C67 (2010), pp. 637–686.
DOI: 10.1140/epjc/s10052-010-1314-6. arXiv: 0906.1833.

[95] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. *"The anti-k_t jet clustering algorithm"*. In: *JHEP* 04 (2008), p. 063.

DOI: `10.1088/1126-6708/2008/04/063`. arXiv: `0802.1189`.

[96]   CMS Collaboration. Jet algorithms performance in 13 TeV data. CMS-PAS-JME-16-003. Geneva: CERN, 2017. URL: `https://cds.cern.ch/record/2256875`.

[97]   CMS Collaboration. *"Determination of Jet Energy Calibration and Transverse Momentum Resolution in CMS"*. In: *JINST* 6 (2011), P11002.
DOI: `10.1088/1748-0221/6/11/P11002`. arXiv: `1107.4277`.

[98]   CMS Collaboration. Identification of b quark jets at the CMS Experiment in the LHC Run 2. CMS-PAS-BTV-15-001. Geneva: CERN, 2016. URL: `https://cds.cern.ch/record/2138504`.

[99]   *"Heavy flavor identification at CMS with deep neural networks"*. In: (Mar. 2017). URL: `http://cds.cern.ch/record/2255736`.

[100] CMS Collaboration. *"Reconstruction and identification of τ lepton decays to hadrons and ν_τ at CMS"*. In: *JINST* 11.01 (2016), P01019.
DOI: `10.1088/1748-0221/11/01/P01019`. arXiv: `1510.07488`.

[101] CMS Collaboration. *"Performance of reconstruction and identification of τ leptons decaying to hadrons and ν_τ in pp collisions at $\sqrt{s} = 13$ TeV"*. In: *JINST* 13.10 (2018), P10005.
DOI: `10.1088/1748-0221/13/10/P10005`. arXiv: `1809.02816`.

[102] CMS Collaboration. *"An embedding technique to determine $\tau\tau$ backgrounds in proton-proton collision data"*. In: *JINST* 14.06 (2019), P06032.
DOI: `10.1088/1748-0221/14/06/P06032`. arXiv: `1903.01216`.

[103] CMS Collaboration. *"Measurement of the $Z\gamma^* \to \tau\tau$ cross section in pp collisions at $\sqrt{s} = 13$ TeV and validation of τ lepton analysis techniques"*. In: *Eur. Phys. J.* C78.9 (2018), p. 708.
DOI: `10.1140/epjc/s10052-018-6146-9`. arXiv: `1801.03535`.

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

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

[106] Paolo Nason. *"A New method for combining NLO QCD with shower Monte Carlo algorithms"*. In: *JHEP* 11 (2004), p. 040.
DOI: `10.1088/1126-6708/2004/11/040`. arXiv: `hep-ph/0409146`.

[107] Stefano Frixione, Paolo Nason, and Carlo Oleari. *"Matching NLO QCD computations with Parton Shower simulations: the POWHEG method"*. In: *JHEP* 11 (2007), p. 070.

DOI: `10.1088/1126-6708/2007/11/070`. arXiv: `0709.2092`.

[108] Simone Alioli et al. *"NLO Higgs boson production via gluon fusion matched with shower in POWHEG"*. In: *JHEP* 04 (2009), p. 002.
DOI: `10.1088/1126-6708/2009/04/002`. arXiv: `0812.0578`.

[109] Simone Alioli et al. *"A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX"*. In: *JHEP* 06 (2010), p. 043.
DOI: `10.1007/JHEP06(2010)043`. arXiv: `1002.2581`.

[110] Simone Alioli et al. *"Jet pair production in POWHEG"*. In: *JHEP* 04 (2011), p. 081.
DOI: `10.1007/JHEP04(2011)081`. arXiv: `1012.3380`.

[111] E. Bagnaschi et al. *"Higgs production via gluon fusion in the POWHEG approach in the SM and in the MSSM"*. In: *JHEP* 02 (2012), p. 088.
DOI: `10.1007/JHEP02(2012)088`. arXiv: `1111.2854`.

[112] Johan Alwall et al. *"MadGraph 5: Going Beyond"*. In: *JHEP* 06 (2011), p. 128.
DOI: `10.1007/JHEP06(2011)128`. arXiv: `1106.0522`.

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

[114] Torbjorn Sjostrand, Stephen Mrenna, and Peter Z. Skands. *"A Brief Introduction to PYTHIA 8.1"*. In: *Comput. Phys. Commun.* 178 (2008), pp. 852–867.
DOI: `10.1016/j.cpc.2008.01.036`. arXiv: `0710.3820`.

[115] NNPDF Collaboration. *"Parton distributions for the LHC run II"*. In: *JHEP* 04 (2015), p. 040.
DOI: `10.1007/JHEP04(2015)040`. arXiv: `1410.8849`.

[116] CMS Collaboration. *"Event generator tunes obtained from underlying event and multiparton scattering measurements"*. In: *Eur. Phys. J. C* 76.3 (2016), p. 155.
DOI: `10.1140/epjc/s10052-016-3988-x`. arXiv: `1512.00815`.

[117] Kirill Melnikov and Frank Petriello. *"Electroweak gauge boson production at hadron colliders through $\mathcal{O}(\alpha_s^2)$"*. In: *Phys. Rev. D* 74 (2006), p. 114017.
DOI: `10.1103/PhysRevD.74.114017`. arXiv: `hep-ph/0609070`.

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

[119] Nikolaos Kidonakis. *"Top Quark Production"*. In: *Proceedings, Helmholtz International Summer School on Physics of Heavy Quarks and Hadrons (HQ 2013): JINR, Dubna, Russia, July 15-28, 2013*. 2014, pp. 139–168.

DOI: `10.3204/DESY-PROC-2013-03/Kidonakis`. arXiv: `1311.0283`.

[120] John M. Campbell, R. Keith Ellis, and Ciaran Williams. *"Vector boson pair production at the LHC"*. In: *JHEP* 07 (2011), p. 018.
DOI: `10.1007/JHEP07(2011)018`. arXiv: `1105.0020`.

[121] T. Gehrmann et al. *"W^+W^- Production at Hadron Colliders in Next to Next to Leading Order QCD"*. In: *Phys. Rev. Lett.* 113.21 (2014), p. 212001.
DOI: `10.1103/PhysRevLett.113.212001`. arXiv: `1408.5243`.

[122] CMS Collaboration. *"Performance of the CMS muon detector and muon reconstruction with proton-proton collisions at $\sqrt{s} = 13$ TeV"*. In: *JINST* 13.06 (2018), P06015.
DOI: `10.1088/1748-0221/13/06/P06015`. arXiv: `1804.04528`.

[123] CDF Collaboration. *"Search for neutral MSSM Higgs bosons decaying to tau pairs in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV"*. In: *Phys. Rev. Lett.* 96 (2006), p. 011802.
DOI: `10.1103/PhysRevLett.96.011802`. arXiv: `hep-ex/0508051`.

[124] Simon S Haykin. Neural networks : A comprehensive foundation. 3ed. Prentice Hall, 2008.

[125] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. Deep Learning. `www.deeplearningbook.org`. MIT Press, 2016.

[126] Stuart Russell and Peter Norvig. Artificial Intelligence: A Modern Approach. 3rd. Prentice Hall Press, 2009.

[127] Warren Mcculloch and Walter Pitts. *"A Logical Calculus of Ideas Immanent in Nervous Activity"*. In: *Bulletin of Mathematical Biophysics* 5 (1943), pp. 127–147.

[128] Chigozie Nwankpa et al. *"Activation Functions: Comparison of trends in Practice and Research for Deep Learning"*. In: *CoRR* (2018). arXiv: `1811.03378`. URL: `http://arxiv.org/abs/1811.03378`.

[129] Stefan Wunsch et al. *"Identifying the relevant dependencies of the neural network response on characteristics of the input space"*. In: *Comput. Softw. Big Sci.* 2.1 (2018), p. 5.
DOI: `10.1007/s41781-018-0012-1`. arXiv: `1803.08782`.

[130] K. Hornik, M. Stinchcombe, and H. White. *"Multilayer Feedforward Networks Are Universal Approximators"*. In: *Neural Netw.* 2.5 (1989), pp. 359–366.
DOI: `10.1016/0893-6080(89)90020-8`.

[131] Steve Baker and Robert D. Cousins. *"Clarification of the Use of Chi Square and Likelihood Functions in Fits to Histograms"*. In: *Nucl. Instrum. Meth.* 221 (1984), p. 437.
DOI: `10.1016/0167-5087(84)90016-4`.

[132] Xavier Glorot and Yoshua Bengio. *"Understanding the difficulty of training deep feedforward neural networks"*. In: *Proceedings of the thirteenth international conference on artificial intelligence and statistics*. 2010, p. 249.

[133] Diederik P. Kingma and Jimmy Ba. *"Adam: A Method for Stochastic Optimization"*. In: (2014). arXiv: `1412.6980`.

[134] A. N. Tikhonov. *"Solution of incorrectly formulated problems and the regularization method"*. In: *Soviet Math. Dokl.* 4 (1963), p. 1035.

[135] Nitish Srivastava et al. *"Dropout: A Simple Way to Prevent Neural Networks from Overfitting"*. In: *Journal of Machine Learning Research* 15 (2014), pp. 1929–1958. URL: `http://jmlr.org/papers/v15/srivastava14a.html`.

[136] Trevor Hastie, Robert Tibshirani, and Jerome Friedman. The Elements of Statistical Learning. New York, NY, USA: Springer, 2001.

[137] J. R. Andersen et al. *"Les Houches 2015: Physics at TeV Colliders Standard Model Working Group Report"*. In: *9th Les Houches Workshop on Physics at TeV Colliders (PhysTeV 2015) Les Houches, France, June 1-19, 2015*. 2016. arXiv: `1605.04692`. URL: `http://lss.fnal.gov/archive/2016/conf/fermilab-conf-16-175-ppd-t.pdf`.

[138] CMS Collaboration. *"CMS luminosity - Public results"*. In: (2019). URL: `https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults`.

[139] CMS Collaboration. *"Measurement of the differential cross section for top quark pair production in pp collisions at $\sqrt{s} = 8\,TeV$"*. In: *Eur. Phys. J.* C75.11 (2015), p. 542.
DOI: `10.1140/epjc/s10052-015-3709-x`. arXiv: `1505.04480`.

[140] Keith Hamilton et al. *"NNLOPS simulation of Higgs boson production"*. In: *JHEP* 10 (2013), p. 222.
DOI: `10.1007/JHEP10(2013)222`. arXiv: `1309.0017`.

[141] J. S. Conway. *"Incorporating Nuisance Parameters in Likelihoods for Multisource Spectra"*. In: *Proceedings, PHYSTAT 2011 Workshop on Statistical Issues Related to Discovery Claims in Search Experiments and Unfolding, CERN,Geneva, Switzerland*. 2011, pp. 115–120.
DOI: `10.5170/CERN-2011-006.115`. arXiv: `1103.0354`.

[142] CMS Collaboration. CMS Luminosity Measurements for the 2016 Data Taking Period. CMS-PAS-LUM-17-001. Geneva: CERN, 2017.

[143] CMS Collaboration. CMS Luminosity Measurement for the 2017 Data-Taking Period at $\sqrt{s} = 13$ TeV. CMS-PAS-LUM-17-004. Geneva: CERN, 2018.

[144] Ye Li and Frank Petriello. *"Combining QCD and electroweak corrections to dilepton production in FEWZ"*. In: *Phys. Rev. D* 86 (2012), p. 094034.

DOI: `10.1103/PhysRevD.86.094034`. arXiv: `1208.5967`.

[145] CMS Collaboration. *"Cross section measurement of t-channel single top quark production in pp collisions at $\sqrt{s} = 13$ TeV"*. In: *Phys. Lett. B* 772 (2017), p. 752.
DOI: `10.1016/j.physletb.2017.07.047`. arXiv: `1610.00678`.

[146] CMS Collaboration. *"Measurement of the WZ production cross section in pp collisions at $\sqrt{s} = 13$ TeV"*. In: *Phys. Lett. B* 766 (2017), p. 268.
DOI: `10.1016/j.physletb.2017.01.011`. arXiv: `1607.06943`.

[147] ALEPH, DELPHI, L3, OPAL, LEP Working Group for Higgs Boson Searches Collaboration. *"Search for neutral MSSM Higgs bosons at LEP"*. In: *Eur. Phys. J.* C47 (2006), pp. 547–587.
DOI: `10.1140/epjc/s2006-02569-7`. arXiv: `hep-ex/0602042`.

[148] CDF Collaboration. *"Search for Higgs bosons predicted in two-Higgs-doublet models via decays to tau lepton pairs in 1.96-TeV p anti-p collisions"*. In: *Phys. Rev. Lett.* 103 (2009), p. 201801.
DOI: `10.1103/PhysRevLett.103.201801`. arXiv: `0906.1014`.

[149] D0 Collaboration. *"Search for neutral Higgs bosons in the multi-b-jet topology in $5.2fb^{-1}$ of $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV"*. In: *Phys. Lett.* B698 (2011), pp. 97–104.
DOI: `10.1016/j.physletb.2011.02.062`. arXiv: `1011.1931`.

[150] D0 Collaboration. *"Search for Higgs bosons decaying to $\tau\tau$ pairs in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV"*. In: *Phys. Lett.* B707 (2012), pp. 323–329.
DOI: `10.1016/j.physletb.2011.12.050`. arXiv: `1106.4555`.

[151] CDF Collaboration. *"Search for Higgs Bosons Produced in Association with b-quarks"*. In: *Phys. Rev.* D85 (2012), p. 032005.
DOI: `10.1103/PhysRevD.85.032005`. arXiv: `1106.4782`.

[152] CMS Collaboration. *"Search for neutral MSSM Higgs bosons decaying into a pair of bottom quarks"*. In: *JHEP* 11 (2015), p. 071.
DOI: `10.1007/JHEP11(2015)071`. arXiv: `1506.08329`.

[153] CMS Collaboration. *"Search for beyond the standard model Higgs bosons decaying into a $b\bar{b}$ pair in pp collisions at $\sqrt{s} = 13$ TeV"*. In: *JHEP* 08 (2018), p. 113.
DOI: `10.1007/JHEP08(2018)113`. arXiv: `1805.12191`.

[154] ATLAS Collaboration. *"Search for the neutral Higgs bosons of the Minimal Supersymmetric Standard Model in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector"*. In: *JHEP* 02 (2013), p. 095.
DOI: `10.1007/JHEP02(2013)095`. arXiv: `1211.6956`.

[155] CMS Collaboration. *"Search for neutral MSSM Higgs bosons decaying to $\mu^+\mu^-$ in pp collisions at $\sqrt{s} = 7$ and 8 TeV"*. In: *Phys. Lett.* B752 (2016), pp. 221–246.

DOI: `10.1016/j.physletb.2015.11.042`. arXiv: `1508.01437`.

[156] ATLAS Collaboration. *"Search for additional heavy neutral Higgs and gauge bosons in the ditau final state produced in $36\,fb^{-1}$ of pp collisions at $\sqrt{s} = 13\,TeV$ with the ATLAS detector"*. In: *JHEP* 01 (2018), p. 055.
DOI: `10.1007/JHEP01(2018)055`. arXiv: `1709.07242`.

[157] CMS Collaboration. *"Search for neutral MSSM Higgs bosons decaying to a pair of tau leptons in pp collisions"*. In: *JHEP* 10 (2014), p. 160.
DOI: `10.1007/JHEP10(2014)160`. arXiv: `1408.3316`.

[158] ATLAS Collaboration. *"Search for neutral Higgs bosons of the minimal supersymmetric standard model in pp collisions at $\sqrt{s} = 8\,TeV$ with the ATLAS detector"*. In: *JHEP* 11 (2014), p. 056.
DOI: `10.1007/JHEP11(2014)056`. arXiv: `1409.6064`.

[159] ATLAS Collaboration. *"Search for Minimal Supersymmetric Standard Model Higgs bosons H/A and for a Z' boson in the $\tau\tau$ final state produced in pp collisions at $\sqrt{s} = 13\,TeV$ with the ATLAS Detector"*. In: *Eur. Phys. J.* C76.11 (2016), p. 585.
DOI: `10.1140/epjc/s10052-016-4400-6`. arXiv: `1608.00890`.

[160] CMS Collaboration. *"Search for Neutral MSSM Higgs Bosons Decaying to Tau Pairs in pp Collisions at $\sqrt{s} = 7\,TeV$"*. In: *Phys. Rev. Lett.* 106 (2011), p. 231801.
DOI: `10.1103/PhysRevLett.106.231801`. arXiv: `1104.1619`.

[161] CMS Collaboration. *"Search for neutral Higgs bosons decaying to tau pairs in pp collisions at $\sqrt{s} = 7\,TeV$"*. In: *Phys. Lett.* B713 (2012), pp. 68–90.
DOI: `10.1016/j.physletb.2012.05.028`. arXiv: `1202.4083`.

[162] CMS Collaboration. Summary results of high mass BSM Higgs searches using CMS run-I data. Tech. rep. CMS-PAS-HIG-16-007. Geneva: CERN, 2016. URL: `https://cds.cern.ch/record/2142432`.

[163] Torbjörn Sjöstrand et al. *"An introduction to PYTHIA 8.2"*. In: *Comput. Phys. Commun.* 191 (2015), p. 159.
DOI: `10.1016/j.cpc.2015.01.024`. arXiv: `1410.3012`.

[164] Emanuele Bagnaschi et al. *"Resummation ambiguities in the Higgs transverse-momentum spectrum in the Standard Model and beyond"*. In: *JHEP* 01 (2016), p. 090.
DOI: `10.1007/JHEP01(2016)090`. arXiv: `1510.08850`.

[165] Emanuele Bagnaschi and Alessandro Vicini. *"The Higgs transverse momentum distribution in gluon fusion as a multiscale problem"*. In: *JHEP* 01 (2016), p. 056.
DOI: `10.1007/JHEP01(2016)056`. arXiv: `1505.00735`.

[166] Robert V. Harlander, Hendrik Mantler, and Marius Wiesemann. *"Transverse momentum resummation for Higgs production via gluon fusion in the MSSM"*. In: *JHEP* 11 (2014), p. 116.

DOI: `10.1007/JHEP11(2014)116`. arXiv: `1409.0531`.

[167] M. Wiesemann et al. *"Higgs production in association with bottom quarks"*. In: *JHEP* 02 (2015), p. 132.
DOI: `10.1007/JHEP02(2015)132`. arXiv: `1409.5301`.

[168] A. D. Martin et al. *"Parton distributions for the LHC"*. In: *Eur. Phys. J.* C63 (2009), pp. 189–285.
DOI: `10.1140/epjc/s10052-009-1072-5`. arXiv: `0901.0002`.

[169] A. D. Martin et al. *"Uncertainties on alpha(S) in global PDF analyses and implications for predicted hadronic cross sections"*. In: *Eur. Phys. J.* C64 (2009), pp. 653–680.
DOI: `10.1140/epjc/s10052-009-1164-2`. arXiv: `0905.3531`.

[170] Thomas Junk. *"Confidence Level Computation for Combining Searches with Small Statistics"*. In: *Nucl. Instrum. Meth. A* 434 (1999), p. 435.
DOI: `10.1016/S0168-9002(99)00498-2`. arXiv: `hep-ex/9902006`.

[171] Alexander L. Read. *"Presentation of search results: The CL_s technique"*. In: *J. Phys. G* 28 (2002), p. 2693.
DOI: `10.1088/0954-3899/28/10/313`.

[172] ATLAS and CMS Collaborations. Procedure for the LHC Higgs boson search combination in Summer 2011. ATL-PHYS-PUB 2011-11, CMS NOTE 2011/005. CERN, 2011. URL: `http://cdsweb.cern.ch/record/1379837`.

[173] Glen Cowan et al. *"Asymptotic formulae for likelihood-based tests of new physics"*. In: *Eur. Phys. J.* C71 (2011), p. 1554.
DOI: `10.1140/epjc/s10052-011-1554-0`. arXiv: `1007.1727`.

[174] Lorenzo Bianchini et al. *"Reconstruction of the Higgs mass in $H \to \tau\tau$ Events by Dynamical Likelihood techniques"*. In: *J. Phys. Conf. Ser.* 513 (2014), p. 022035.
DOI: `10.1088/1742-6596/513/2/022035`.

Curriculum Vitae

Dipl.-Ing. Markus Spanring

Contact: markusspanring@gmail.com

Education

2016 – today	Doctoral Program in Technical Sciences Vienna University of Technology, Austria Technical Physics
2013 – 2015	Master Program Vienna University of Technology, Austria Technical Physics: Degree Dipl.-Ing./M.Sc (passed with distinction) Thesis: Study of potential improvements of the CMS H → ττ analysis using artificial neural networks with multiple layers
2009 – 2013	Bachelor Program Vienna University of Technology, Austria Technical Physics: Degree B.Sc. Thesis: Pressure dependence of the breakdown voltage between two neutron mirrors.
2008	Military Service, Amstetten, Austria
2003 – 2008	Reife- und Diplomprüfung Höhere Technische Bundeslehr- und Versuchsanstalt, Waidhofen/Ybbs, Austria Electronics and Information Technology: passed with distinction Thesis: GPS based track simulation for indoor cycling

Experience

Research

2016 – today	Research for doctoral thesis, Institute of High Energy Physics, Vienna, Austria Research of the Higgs boson decay to tau leptons with the CMS experiment at the Large Hadron Collider
2014 – 2015	Research for master thesis, Institute of High Energy Physics, Vienna, Austria Research of the Higgs boson decay to tau leptons with focus on the implementation and application of deep learning for signal extraction
2014	Project assistant, Institute of Solid State Physics, Vienna University of Technology, Austria Precision measurement of the temperature on a platinum resistance temperature detector
2013	Project work, Atominstitut, Vienna, Austria Calibration of the electronic data acquisition and control system "LogicBox"

Curriculum Vitae

Teaching

2017 – 2019   Co-supervision of master students, Institute of High Energy Physics, Vienna, Austria

2013   Tutor, Vienna University of Technology, Austria
Teaching of the exercise "Experimental Exercises in Physics III"

2012 – 2013   Tutor, Vienna University of Technology, Austria
Teaching of the exercise "Experimental Exercises in Physics II"

Work

2009, 2015 – 2016, 2019   Engineer, E-Wolf Automatisierungstechnik GmbH, Opponitz, Austria

2005 – 2014   Miscellaneous summer internships, Austria

Publications

- CMS Collaboration. Measurement of Higgs boson production and decay to the $\tau\tau$ final state CMS-PAS-HIG-18-032 (2019). https://cds.cern.ch/record/2668685.

- CMS Collaboration. Search for additional neutral MSSM Higgs bosons in the $\tau\tau$ final state in proton-proton collisions at 13 TeV. JHEP 09(2018), 007.

- CMS Collaboration. Measurement of the $Z \rightarrow \tau\tau$ cross section in pp collisions at 13 TeV and validation of τ lepton analysis techniques. Eur.Phys.J. C78(2018),708.

- M. Spanring for the CMS Collaboration. Search for additional neutral MSSM Higgs bosons in the di-tau final state. PoS(ALPS2018) 054.

Contributions to Conferences

- Couplings and properties measurements with fermionic Higgs decays and searches for second/first generation decays at the LHC
LHCP2019: 7th Conference of Large Hadron Collider Physics, 20-25 May 2019, BUAP, Puebla (Mexico)

- Search for additional neutral MSSM Higgs bosons in the di-tau final state at CMS
Alps2018: Third Alpine LHC Physics Summit, 15-20 Apr 2018, Obergurgl (Austria)

- CMS Higgs Physics
BSM-2017: Beyond Standard Model: From Theory To Experiment, 17-21 Dec 2017, Hurghada (Egypt)

- H $\rightarrow \tau\tau$ measurements with CMS
12th-VS: 12th Central European Seminar on Particle Physics and Quantum Field Theory "Physics of LHC Run 2", 1-2 Dec 2016, Natural History Museum Vienna, Vienna (Austria)

- Measurement of Higgs and Z boson decays to a pair of tau leptons with CMS
OePG 2016: 66th Annual Meeting of the Austrian Physical Society, 27-29 Sep 2016, University of Vienna, Vienna (Austria)