IEKP-KA/2017-16

Karlsruher Institut für Technologie

Establishing the Standard Model Higgs Boson in the Decay Channel $H \to \tau\tau$ with LHC Run II data

Zur Erlangung des akademischen Grades eines
DOKTORS DER NATURWISSENSCHAFTEN (Dr. rer. nat)
von der KIT-Fakultät für Physik des
Karlsruher Instituts für Technologie (KIT)
angenommene Dissertation

DISSERTATION

von

Dipl.-Phys. Raphael Friese
aus Ostfildern

Tag der mündlichen Prüfung: 26. 05. 2017

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

Contents

Introduction

The subject of this thesis is the mechanism through which leptons obtain their masses. This is done via the search for the decay of Higgs bosons into pairs of the heaviest kind of leptons, the tau leptons.

In this first chapter, the theoretical backgrounds are shortly recalled. Chapter 2 summarizes the analyses that led to the discovery of the Higgs boson and presents a description of the CMS experiment. Chapter 3 presents the progress achieved in the reconstruction of undetectable particles. The last chapter presents in detail the $H \to \tau\tau$ analysis and its results, addressing the initial question of the lepton mass generating mechanism.

Chapter 1

The Higgs Mechanism in the Standard Model of Particle Physics

This chapter is a brief introduction in the Higgs mechanism and its couplings to bosons and fermions. It introduces the Higgs field leading to the postulation of the Higgs boson. It also gives an overview on the Higgs boson production processes and branching ratios. For a deeper study see the textbooks [1–3].

1.1 The Standard Model of Particle Physics

The Standard Model of Particle Physics (SM) is a renormalisable quantum field theory. Excitations of fields are interpreted as particles. There are two kinds of particles. Bosons with integer spin are force carriers, and are responsible for the three fundamental interactions that are described by the Standard Model. Fermions have half-integer spin and are the matter particles. The Higgs boson is the only particle with zero spin.

The dynamics of fields can be described by means of the Lagrange formalism and a Lagrangian density $\mathcal{L}_{\mathrm{SM}}$. This Lagrangian density is a function of quantum fields from which the dynamics of all known particles can be derived from. $\mathcal{L}_{\mathrm{SM}}$ has some intrinsic degrees of freedom that do not change the observables.

The concept of formulating classical mechanics with a Lagrangian has already been developed in the late 18th century. In essence, it is a reformulation of Newton's mechanics. One of its features is the ability to formulate a problem in so-called *generalized coordinates* that exploits intrinsic symmetries of a given problem. In this way, solving the system of equations can become easier.

The objects of interest in particle physics are usually Lorentz Vectors, also called four-vectors. Lorentz Vectors are Lorentz invariant, which means that they keep their magnitude under relativistic transformations. This is important, because in particle physics experiments most particles move at relativistic velocities. A four-vector describes a four-momentum that combines both energy and momentum.

In classical mechanics the minimization of action is done with the Euler-Lagrange equation. Quantum mechanics uses the path integral formulation. This formulation is a generalization that accounts for the fact that in quantum mechanics not longer localized particles move through space and time. Instead, fields as functions of space and time are introduced. The Lagrangian density in quantum mechanics then is a function of these fields $\phi(x^\mu)$ and their derivatives $\partial_\mu\phi$:

$$\mathcal{L}\left(\phi, \partial_\mu\phi\right) \text{ with } \partial_\mu\phi = \frac{\partial\phi}{\partial x^\mu} \tag{1.1}$$

Integration of the Lagrangian density gives the action

$$S = \int \mathrm{d}^4x \mathcal{L}\left(\phi, \partial_\mu\phi\right) \tag{1.2}$$

The appropriate Lagrangian for a certain problem is not derived but stated axiomatically. By deriving other equations from it, it has to prove that it is suitable for the problem. Three kinds of Lagrangian densities for three different kinds of fields are postulated below.

- The Lagrangian density of a complex scalar field (spin 0) is given by

  $$\mathcal{L} = \frac{1}{2}\left(\partial_\mu\phi^*\right)\left(\partial^\mu\phi\right) - \frac{1}{2}m^2\phi^*\phi \tag{1.3}$$

  with the complex conjugate field ϕ^*. Applying the Euler-Lagrange equation this leads to the Klein-Gordon equation,

  $$\left(\partial_\mu\partial^\mu - m^2\right)\phi = 0 \tag{1.4}$$

  which describes the dynamics of free scalar fields.

- The Lagrangian density of fermion fields ψ with spin $\frac{1}{2}$ requires the fields to be spinors. It is given by

  $$\mathcal{L} = i\bar{\psi}\gamma^\mu\partial_\mu\psi - m\bar{\psi}\psi \tag{1.5}$$

  with the abbreviation $\bar{\psi} = \psi^\dagger\gamma^0$ while γ^μ are the Dirac or gamma matrices. This Lagrangian density leads to the Dirac equation, which describes the behavior and motion of a fermionic field:

  $$\left(i\gamma^\mu\partial_\mu - m\right)\psi = 0 \tag{1.6}$$

- A vector field A^μ has a spin 1 and the following Lagrangian:

$$\mathcal{L} = -\frac{1}{4}F^{\mu\nu}F_{\mu\nu} + \frac{1}{2}m^2A^\nu A_\nu \tag{1.7}$$

with the field strength tensor $F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu$. The corresponding equation of motion are the source-free Maxwell equations

$$\partial_\mu F^{\mu\nu} + m^2A^\nu = 0 \tag{1.8}$$

1.2 Local Gauge invariance

Physical observables are always invariant under phase transformations of the underlying field. This is reflected in the Lagrange mechanism e.g. in the Dirac Lagrangian (equation 1.5), by the fact that a transformation like

$$\psi \to e^{i\theta}\psi \tag{1.9}$$

cancel out, since ψ only occurs in the combination $\bar{\psi}\psi$. This is called a global gauge transformation. A local phase transformation is defined as $\theta \to \theta(x_\mu)$. The Dirac Lagrangian contains the derivative of ψ, which transforms as:

$$\partial_\mu\left(e^{i\theta}\psi\right) = i\left(\partial_\mu\theta\right)e^{i\theta}\psi + e^{i\theta}\partial_\mu\psi \tag{1.10}$$

This leads to a modified Lagrangian

$$\mathcal{L} \to \mathcal{L} - \left(\partial_\mu\theta\right)\bar{\psi}\gamma^\mu\psi \tag{1.11}$$

By replacing of θ with a quantity called λ:

$$\lambda(x) = -\frac{1}{q}\theta(x), \tag{1.12}$$

a constant q appears that later will be identified as the charge. Written that way, the Lagrangian becomes

$$\mathcal{L} \to \mathcal{L} + \left(q\bar{\psi}\gamma^\mu\psi\right)\partial_\mu\lambda \tag{1.13}$$

This new term is a result of a local gauge transformation. To compensate it, a field A_μ is added to the Dirac Lagrangian, leading to

$$\mathcal{L} = \left[i\bar{\psi}\gamma^\mu\partial_\mu\psi - m\bar{\psi}\psi\right] - \left(q\bar{\psi}\gamma^\mu\psi\right)A_\mu \tag{1.14}$$

By requiring that A_μ transforms as

$$A_\mu \to A_\mu + \partial_\mu \lambda \tag{1.15}$$

it is easy to see that this newly introduced gauge field A_μ exactly compensates the new term in equation 1.13. We already know from equation 1.7 how the Lagrangian for a vector field looks like from where the dynamics can be derived from. The first part, $\frac{1}{4}F^{\mu\nu}F_{\mu\nu}$, is invariant under transformation 1.15 since $F^{\mu\nu}$ itself is invariant. This isn't the case for the second term $\frac{1}{2}m_A^2 A^\nu A_\nu$. This requires $m_A = 0$ to keep gauge invariance.

1.3 The Higgs mechanism

The theory derived above stands in obvious contradiction to the observation of the $W^\pm$ and the Z bosons with masses of $m_W = 80.4\,\mathrm{GeV}/c^2$ and $m_\mathrm{Z} = 91.2\,\mathrm{GeV}/c^2$. A mechanism that keeps local gauge invariance but can explain these high masses is introduced in the following section. The mechanism might also provides an explanation of the fermion masses and comes with predictions of the branching ratio of a newly postulated boson.

1.3.1 The Higgs mechanism

The Brout-Englert-Higgs mechanism [4–8] does not add an explicit mass term to the Lagrangian but postulates a scalar field ϕ with a characteristic potential $V(\phi)$. The important thing about this potential is, that it has to be symmetric around the origin under phase transformation and it has to have a global minimum not sharing this symmetry. The simplest choice to fulfill this is

$$V(\phi) = -\mu^2 \phi\phi^* + \lambda^2 \left(\phi\phi^*\right)^2 \tag{1.16}$$

where $\phi(x)$ is a complex scalar field

$$\phi(x) = \frac{1}{\sqrt{2}} \left(\phi_1(x) + i\phi_2(x)\right). \tag{1.17}$$

The potential V has its minimum at

$$|\phi_0| = \sqrt{\frac{\phi_1^2 + \phi_2^2}{2}} = \frac{v}{\sqrt{2}} \text{ with } v = \frac{\mu}{\lambda} \tag{1.18}$$

The choice of phase is arbitrary, since the potential has rotational symmetry. The potential has the characteristic shape of a *Mexican hat*. See Figure 1.1 for a visualization.

The potential $V(\phi)$ can be developed around the ground state with the ansatz

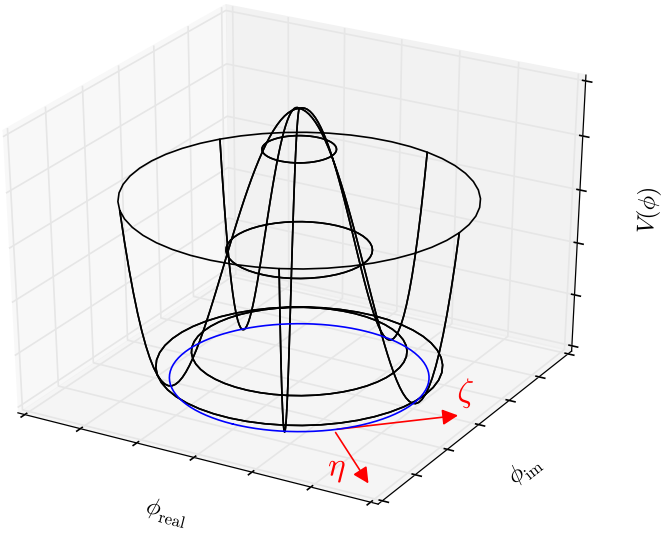


Figure 1.1: The characteristic Higgs boson potential $V(\phi)$ with $\mu \in \mathbb{R}$ has the shape of a *Mexican hat*. The blue circle marks the locus of the global minima. They all have the same distance $|\phi_0|$ to the unstable local maximum in the center at $\phi = 0 + i \cdot 0$ but differ in phase. The radial direction η and tangential direction ζ is marked. They are later used in the development around the minimum of the potential $V(\phi)$.

$$\phi(x) = \frac{1}{\sqrt{2}}\left(v + \eta(x) + i\zeta(x)\right) \tag{1.19}$$

The Lagrangian for the Higgs field and an arbitrary vector field is

$$\mathcal{L} = (D_\mu\phi)(D^\mu\phi)^* + \mu^2(\phi\phi*) - \frac{\mu^2}{v^2}(\phi\phi*)^2 - \frac{1}{4}F_{\mu\nu}F^{\mu\nu}. \tag{1.20}$$

D_μ is the covariant derivative $D_\mu = \partial^\mu + iqA^\mu$. Inserting the potential $V(\phi)$ and neglecting the higher order terms, one gets

$$\begin{aligned}\mathcal{L} \approx &\left[\frac{1}{2}(\partial_\nu\eta)(\partial^\nu\eta) - \mu^2\eta^2\right] + \left[\frac{1}{2}(\partial_\mu\zeta)(\partial^\mu\zeta)\right] \\ &- \frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \frac{1}{2}q^2v^2A_\mu A^\mu + qvA_\mu(\partial^\mu\zeta)\end{aligned} \tag{1.21}$$

This Lagrangian is gauge invariant, so a local gauge transformation can be performed

$$\chi(x) = -\frac{1}{qv} \cdot \zeta(x) \tag{1.22}$$

leading to a transformed Field A':

$$A'_\mu(x) = A_\mu(x) + \partial_\mu \left(\frac{1}{qv} \zeta(x) \right) \tag{1.23}$$

inserting this into equation 1.21 one gets

$$\mathcal{L} = \left[\frac{1}{2} (\partial_\nu \eta)(\partial^\nu \eta) - \mu^2 \eta^2 \right] - \frac{1}{4} F_{\mu\nu} F^{\mu\nu} + \frac{1}{2} q^2 v^2 A'_\mu A'^\mu + \dots \tag{1.24}$$

This means that the ζ field can be canceled out just by the choice of an appropriate gauge. The η field stays, and it is massive. There is also a massive vector field A with the mass $m = qv$.

1.3.2 The generation of mass for W and Z bosons

The simplest way to explain the mass generating mechanism for the gauge bosons of the weak interaction is by introducing a SU(2) doublet field

$$\Phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} \tag{1.25}$$

The Higgs and elektroweak gauge boson field Lagrangian then reads as

$$\mathcal{L} = (D^\mu \Phi)^\dagger (D_\mu \Phi) \underbrace{+\mu^2 (\Phi^\dagger \Phi) - \lambda^2 (\Phi^\dagger \Phi)}_{-V(\Phi^\dagger, \Phi)} \underbrace{-\frac{1}{4} F^i_{\mu\nu} F^{i\mu\nu} - \frac{1}{4} f_{\mu\nu} f^{\mu\nu}}_{\mathcal{L}_{\text{gauge bosons}}} \tag{1.26}$$

and is gauge invariant under local SU(2) transformations as well as local U(1) transformations. With the ansatz from equation 1.19 and already knowing that the ζ term can be eliminated, the field $\Phi(x)$ around the ground state is

$$\Phi(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + \eta(x) \end{pmatrix}. \tag{1.27}$$

The covariant derivative is

$$D_\mu = \partial_\mu + i \frac{g}{2} \boldsymbol{\tau} \cdot \boldsymbol{W}_\mu + i \frac{g'}{2} \cdot B_\mu \tag{1.28}$$

with the SU(2) generating Pauli-Matrices $\boldsymbol{\tau}$, the fields $\boldsymbol{W}_\mu$ and the vector field B_μ, linked to the gauge boson Lagrangian via

$$\begin{aligned} F^i_{\mu\nu} &= \partial_\mu W^i_\mu - \partial_\nu W^i_\mu - g\epsilon_{ijk} W^j_\mu W^k_\nu \\ f_{\mu\nu} &= \partial_\mu B_\nu - \partial_\nu B_\mu \end{aligned} \tag{1.29}$$

Performing the substitution of the covariant derivative in the Lagrangian in equation 1.26 one gets

$$\begin{aligned} \mathcal{L} = &\left[\frac{1}{2}(\partial_\nu \eta)(\partial^\nu \eta) - \mu^2 \eta^2\right] - \frac{1}{4} F^i_{\mu\nu} F^{i\mu\nu} - \frac{1}{4} f_{\mu\nu} f^{\mu\nu} \\ &+ \frac{1}{2} \cdot \frac{g^2 v^2}{4} \left(\left| W^{(+)}_\mu \right|^2 + \left| W^{(-)}_\mu \right| \right) + \frac{1}{2} \cdot \frac{v^2}{4} \left| g' B_\mu - g W_{3\mu} \right|^2 \end{aligned} \tag{1.30}$$

From this Lagrangian one can read off the mass of the η field that can be identified as the Higgs field:

$$m_H = \sqrt{2}\mu. \tag{1.31}$$

The Higgs boson mass used to be the only free parameter of the Standard Model until until the discovery in 2012 (see next chapter). The W-boson terms also appear with a mass

$$m_W = \frac{gv}{2}. \tag{1.32}$$

The characteristic parameter v is linked with the Higgs field vacuum expectation value by the Fermi constant G_F

$$\frac{G_F}{\sqrt{2}} = \frac{g^2}{8 m_W^2} = \frac{v^2}{2} \;\Rightarrow\; v \approx 246\,\text{GeV} \tag{1.33}$$

The last term can be transformed so that it gives the link between the Z and W boson masses:

$$m_Z = \frac{gv}{2\cos\theta_w} = \frac{m_W}{\cos\theta_W} \tag{1.34}$$

with the Weinberg angle θ_W.

1.3.3 Fermion masses

The Yukawa coupling Lagrangian, linking the scalar Higgs boson field with the fermion fields is

$$\mathcal{L}_{\text{Yukawa}} = -\widetilde{g_e} \left[\bar{R}(\Phi^\dagger L) + (\bar{L}\Phi) R \right] \tag{1.35}$$

L is the left-handed electron-neutrino doublet and R is the right-handed electron singlet. e_L and e_R are the left- and right-handed components of the electron field, ν_e is the field of the left-handed neutrino.

$$L \equiv \begin{pmatrix} \nu_e \\ e_L \end{pmatrix} \text{ and } R \equiv e_R \tag{1.36}$$

Putting these into the Yukawa Lagrangian with the vacuum expectation value of the Higgs boson field one gets

$$\begin{aligned} \mathcal{L}_{\text{Yukawa}} &= -\widetilde{g_e}\left[e_R(0, v/\sqrt{2})\begin{pmatrix} \nu_e \\ e_L \end{pmatrix} + ((\nu_e, e_L)\begin{pmatrix} 0 \\ v/\sqrt{2} \end{pmatrix})e_R\right] \\ &= -\widetilde{g_e}\frac{v}{\sqrt{2}}\left[e_R e_L + e_L e_R\right] \\ &= -\frac{\widetilde{g_e}\cdot v}{\sqrt{2}}\bar{e}e \end{aligned} \tag{1.37}$$

where one can identify the mass term

$$m_e = \frac{\widetilde{g_e}\cdot v}{\sqrt{2}} \tag{1.38}$$

with the corresponding coupling

$$\widetilde{g_e} = \frac{m_e}{v/\sqrt{2}} \tag{1.39}$$

The finding that the coupling of the Higgs boson is proportional to the mass of the fermion has consequences for the Higgs boson production and decay. This is discussed in the following section.

1.4 Higgs boson production and decay

With m_{H} being a free parameter of the Standard Model Higgs boson, one can calculate all of its properties depending on this mass. This is done in the following, explaining the possible production modes at the LHC and the expected branching ratios.

1.4.1 Higgs boson production

There are four main production modes contributing at the LHC.

The Higgs boson production via **gluon-gluon-fusion** ($gg \to H$) with a cross section of 48.5 pb at 13 TeV is the dominant one (figure 1.2). Two gluons produce a Higgs boson via a ferminoic loop. Since the Higgs boson coupling to fermions is proportional to their masses, the top quark loop has the largest contribution. The $pp \to H + \text{jet}$ production where the top loop additionally emits a gluon is an important event signature. It causes the Higgs boson to be boosted. This increases the reconstruction efficiency, since in the $H \to \tau\tau$ analysis, the tau decay products are required to exceed a certain threshold of transverse momentum. Boosted topologies improve the mass resolution and can therefore be better separated from the backgrounds. The cross section for the $pp \to H + 1\text{jet}$ production with a transverse momentum of $p_{\mathrm{T}}^{\mathrm{H}} > 100\,\mathrm{GeV}/c$ is 2 pb and 0.35 pb for a momentum $p_{\mathrm{T}}^{\mathrm{H}} > 200\,\mathrm{GeV}/c$ at 13 TeV [9].

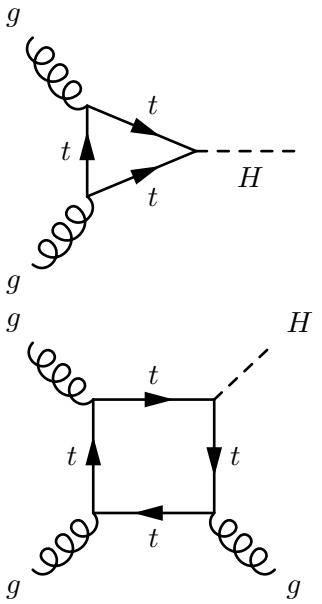


Figure 1.2: Higgs boson production from gluons via a loop of heavy quarks

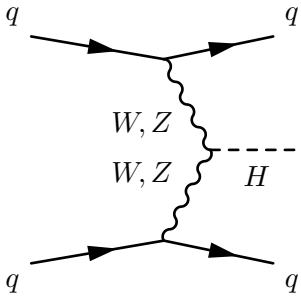


Figure 1.3: Higgs boson production via vector boson fusion

The **V**ector **B**oson **F**usion ($qq \to H$) has a cross section of 3.8 fb. Two quarks radiate vector bosons, fusing into a Higgs boson, see the Feynman graph in figure 1.3. The remaining quarks hadronize, leading to a very specific event topology with two quark jets, separated by a large gap in pseudorapidity. The VBF jets are also good for probing the Higgs boson spin or the interference by the CP-even and CP-odd couplings by their signed angular difference [10], [11].

In the **associated production** mode two quarks annihilate to produce a W or Z boson, radiating a Higgs boson. The subsequent decays of the W or Z bosons can also be used for a special selection targeting the associated production. The reconstruction and identification of the associated bosons are essential for the $H \to b\bar{b}$ analysis since the QCD background is otherwise overwhelming. The cross section for the production in association with a W boson is 1.3 pb and 0.87 pb in association with a Z boson. See the Feynman graph in figure 1.4.

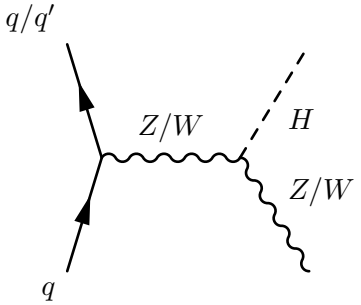


Figure 1.4: Higgs boson production is association with a W or Z boson.

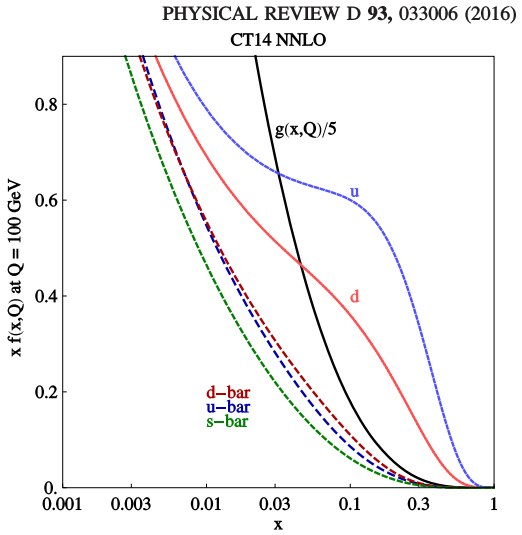


Figure 1.6: Parton distribution function at the scale of $Q = 100\,\text{GeV}$ which is close to the search range of m_H from the global fit to the CT14 NNLO ensemble [12]. The plot shows the probability density functions for observing a valence quark (u, d) or anti-quark ($\bar{d}$,$\bar{u}$,$\bar{s} = s$) as well as a gluon g (divided by a factor of 5) as a function of x. The PDFs are derived from the CT14 dataset containing the HERA, TEVATRON and LHC Run I measurements.

With only about 1 of 100 Higgs bosons being produced in the $\mathbf{t\bar{t}H}$ mode, roughly only 3000 events have been expected to be produced in Run I. Nevertheless, the presence of the two $t\bar{t}H$ jets is exploited in many analyzes.

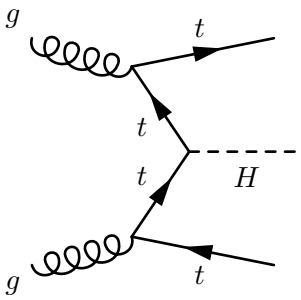


Figure 1.5: $\mathbf{t\bar{t}H}$ production mode Feynman graph

The fact that gluon-gluon fusion is the dominant production mode can be explained through the parton density function. At the scale around $Q = 100\,\text{GeV}$ (see figure 1.6) it is a lot more probable to find gluons carrying a small fraction x of the proton energy than finding a quark. To produce a Higgs boson, a total energy of 125 GeV is required plus the energy for the boost, if there is any. To produce an unboosted Higgs boson, it is sufficient that both partons have a very small $x \approx 0.01$. In this region, gluons are dominant. Only at higher $x > 0.2$ quarks start to take over. It is very unlikely to find two quarks of large x.

There are also self-coupling terms of the Higgs boson that have not been derived in the previous section. The $H \to HH$ trilinear coupling constant is [13]

$$f_{H\to HH} = i\frac{3m_H^2}{v} \tag{1.40}$$

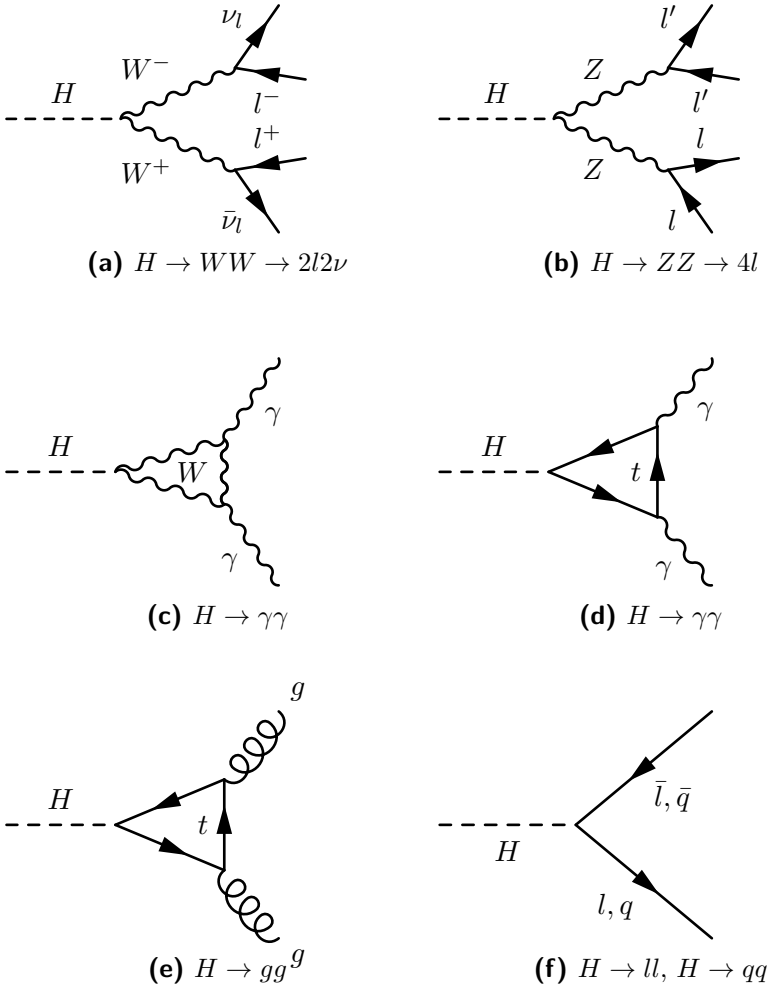
The expected cross section at next-to-leading order for di-Higgs production via the dominant process gluon-gluon fusion is 33.89 fb at a center-of-mass energy of 14 TeV, which is three orders of magnitude lower than the cross section for single

Higgs production. The di-Higgs production is not yet experimentally accessible with the data taken up to now at the LHC, since probably only a handful of such events have occurred during the first five years of data taking of the LHC.

1.4.2 Higgs boson decay

The Higgs boson branching ratios can be expressed as functions of the single parameter m_{H}. For masses up to two times the W mass and Z mass, the decay to $\mathrm{b\bar{b}}$ is dominant while $\mathrm{BR}(H \to \tau\tau)$ is in the order of 9%. Assuming m_{H} to be above the threshold of two times the top mass, the $H \to \mathrm{t\bar{t}}$ is the dominant decay to fermions.

The leading Feynman graphs can be found in figure 1.7 for both the leptonic and fermionic decays. The branching ratios can be found in figure 1.8. The bosonic decay modes have, in general, a much smaller branching ratios. However, the have the advantage of a better mass resolution in the decay to photons and Z bosons, having no $\not{E}_{\mathrm{T}}$ in the final state. The decay to quarks and gluons is experimentally hard to select upon.



(a) $H \to WW \to 2l2\nu$ **(b)** $H \to ZZ \to 4l$

(c) $H \to \gamma\gamma$ **(d)** $H \to \gamma\gamma$

(e) $H \to gg$ **(f)** $H \to ll$, $H \to qq$

Figure 1.7: The main bosonic decay modes (a to e) with the important subsequent decays of the bosons and the leading fermionic one (f). l always denotes a leptonic decay, $l = e,\mu,\tau$. (a) The $H \to 2l2\nu$ final state with its clear signature. (b) $H \to ZZ$: the two Z can either decay to lepton pairs having the same or different flavor, leading to different background compositions and different analysis strategies. The $H \to \gamma\gamma$ decay happens only via a W loop (c) where the photons couple to the W charge or a fermionic loop (d), where the largest contribution comes from the top quark, since its the heaviest. The decay of the Higgs boson to gluons over a lepton loop is experimentally uninteresting, having a large QCD background and neither a large branching ratio such as $H \to b\bar{b}$, nor a clean signature like the other bosonic channels.

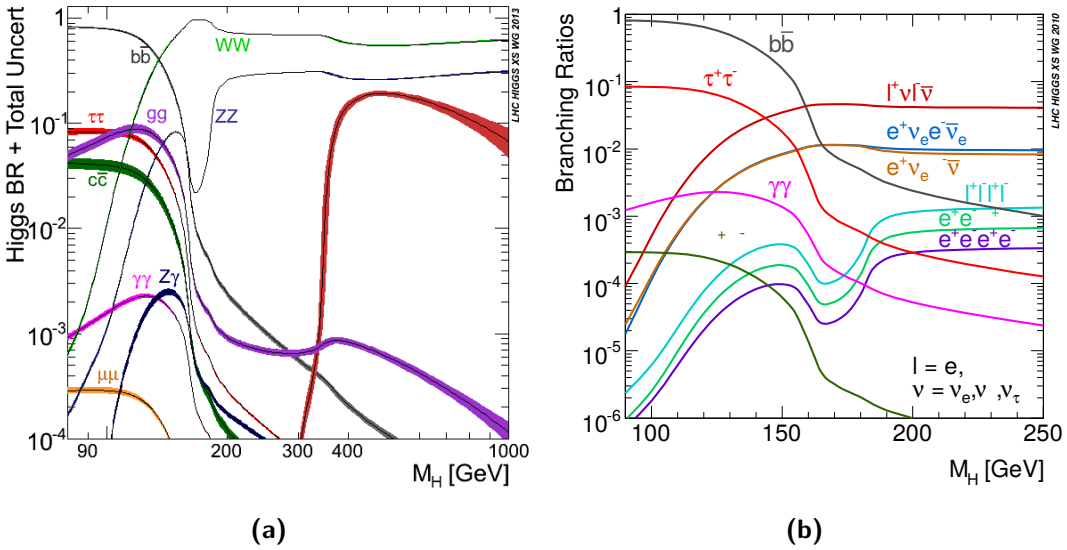


Figure 1.8: (a) Expected Higgs boson branching ratios with as a function of m_{H} in the search range from 80 GeV to 1 TeV. (b) Expected production modes times branching ratio as a function of m_{H} with a range from 90 GeV to 250 GeV. [14]

Chapter 2

The Higgs Boson Discovery at the CMS experiment

2.1 The LHC

The **L**arge **H**adron **C**ollider [15] is a ring accelerator for protons and heavy ions with a design center-of-mass energy of 14 TeV while up to now operating at 7, 8 and 13 TeV. It is located at the French-Swiss border near Geneva at the European Organization for Nuclear Research (CERN). It has been built in the nearly ring-shaped tunnel that originally has been created for the **L**arge **E**lectron-**P**ositron collider [16]. The center-of-mass energy reached by the LHC is only determined by two numbers: The circumference and the strength of the magnetic field of the dipole magnets. These magnets can deliver a field of up to 8.33 T, keeping bunches of protons on their circular track. In the electron-positron collider LEP the main limitation for higher beam energies was the energy loss due to synchrotron radiation, given by the power P:

$$P_{sync} = \frac{q^2}{6\pi R^2} \left(\frac{E}{m}\right)^4 \tag{2.1}$$

for the charge q, radius R and energy E. Even though the particle energy increased from 104.5 GeV (LEP) to 8 TeV (LHC), due to the high proton mass of 2000 times the electron mass, the synchrotron radiation went down by a factor of $O(10^6)$ and is therefore negligible at the LHC.

The LHC is part of a large acceleration complex and filled by a chain of preceding smaller particle accelerators, see figure 2.1. Hydrogen Ions get accelerated by a linear accelerator (LINAC2) to 50 MeV. In the *booster*, they are brought up to 1.4 GeV. In the Super Proton Synchrotron, they are accelerated to 450 GeV before being finally injected into the LHC. The acceleration of up to 14 TeV is performed by eight radio-frequency cavities. This technique requires the proton beam to be split up into bunches, of which the LHC can theoretically hold up to 2808. At this bunch spacing, there is a proton bunch at nearly the speed of light around every

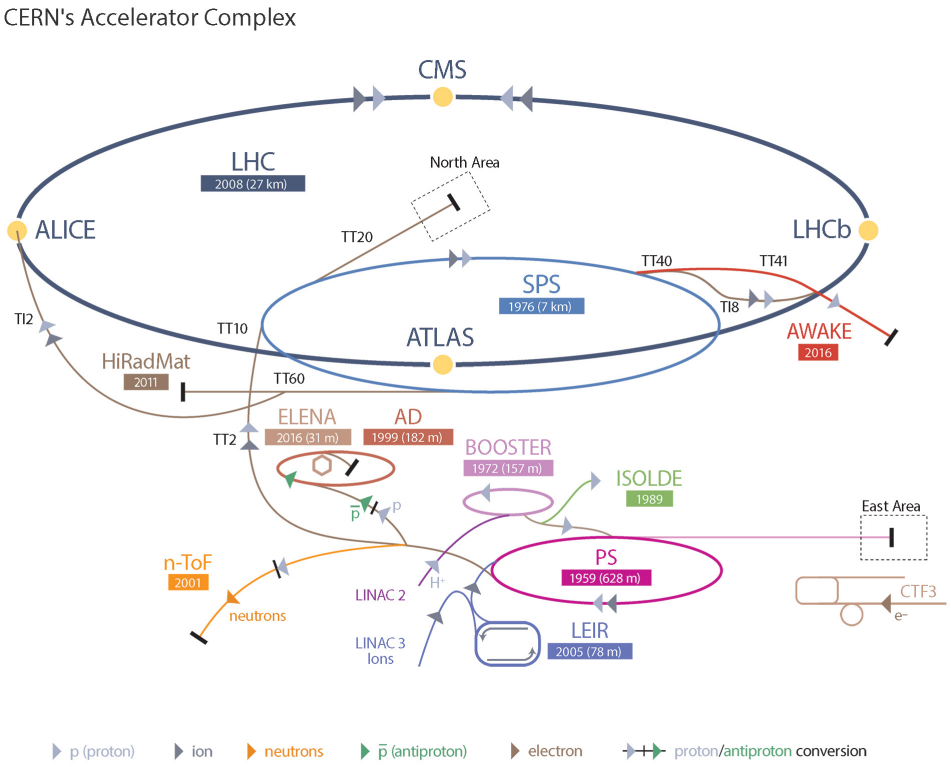


Figure 2.1: The LHC accelerator complex. Many different colliders are involved until the final beam energy of 7 TeV is reached. Ionized Hydrogen get accelerated by the only linear accelerator in the chain, the LINAC2. The protons follow a chain of the Booster, **P**roton **S**ynchrotron and **S**uper **P**roton **S**ynchrotron while the last step is the LHC. [18]

9 m. Since mid of 2015, the LHC is operated with its design collision rate of 40 MHz. If the instantaneous luminosity sinks down below a certain level or in case of an exception, the beam gets *dumped* with all the stored energy of nearly 100 kWh to a block made of graphite with a front-surface of 0.5 m and a length of 7 m. The block is water cooled and can absorb the proton energy safely, while an uncontrolled beam dump might cause serious damage to the machine.

The recorded luminosity of the LHC in its second run was even higher than expected. In 2016 more collisions happened than in all proceeding years together, see Figure 2.2. After a longer end-of-the year stop in the beginning of 2016 in which some detector upgrades took place, the Run II will continue until end of 2018, targeting an integrated luminosity of 150 fb^{-1} [17].

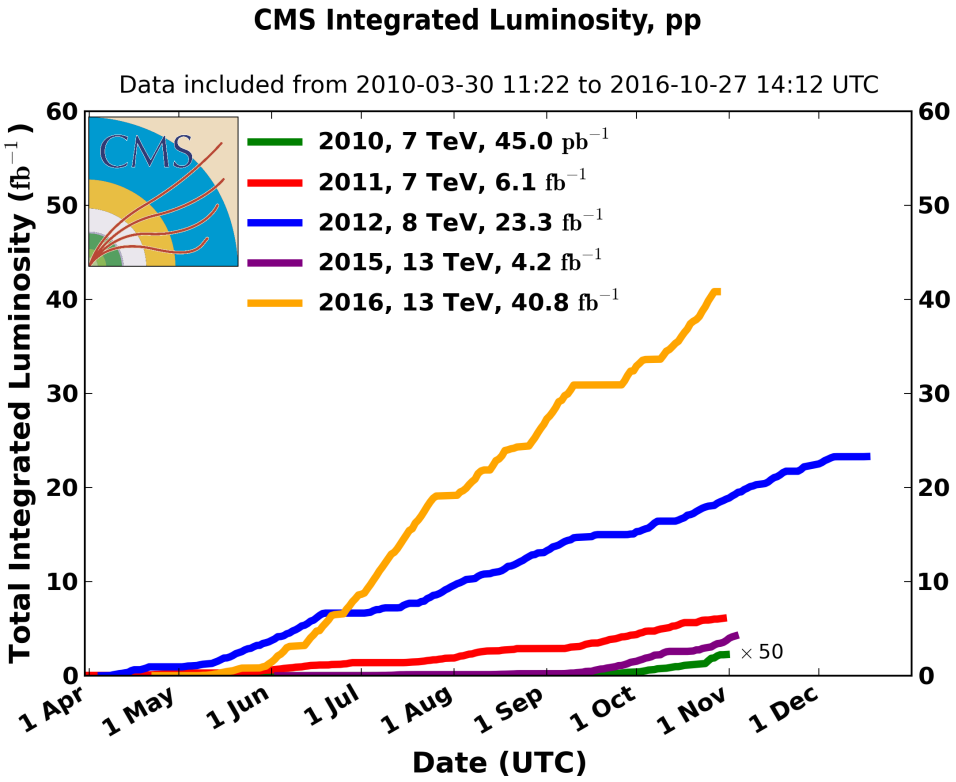


Figure 2.2: Integrated, recorded luminosity of the CMS detector in five years of operation [19]. The Analysis in this thesis is based on the first 12.6 fb^{-1} of integrated luminosity from 2016 at 13 TeV center of mass energy, representing the dataset used for the ICHEP 2016 conference.

2.2 The CMS Experiment

The **C**ompact **M**uon **S**olenoid detector is one of the two general purpose detectors observing proton-proton collisions at the LHC. The CMS collaboration consists of more than 3500 scientists, engineers and students from all over the world [20].

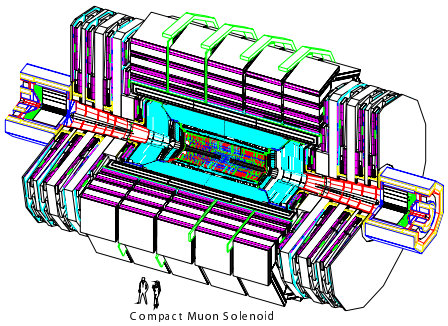


Figure 2.3: The CMS detector with two persons for scale.

Due to its multi-purpose nature, the CMS detector has to fulfill a wide range of specifications. To distinguish prompt decays from secondary ones, a good spatial resolution in the track reconstruction system is necessary. Combined with the high resolution in the calorimeters, pile-up mitigation techniques, sub-jet analysis as well as flavor tagging become possible and are done.

For a perfect coverage of the interaction point, the optimal shape of the detector would be a ball or cigar. Because this is technically not possible, most subdetectors have two parts. One is the **barrel** region, being cylindric and covering the region transverse to the beam direction. The other part covers the **forward** or **endcap** region that is shaped in disks with a hole in the middle for the beam pipe, being large enough keep radiation damage at a reasonable level while still being traversed by as many particles as possible.

CMS conventions

The CMS experiment uses a right-handed coordinate system with the origin at the nominal interaction point. The x-axis points towards the center of the LHC, the Y-axis upwards and the Z-axis is parallel to the beam axis, pointing in the anti-clockwise direction. The coordinates are either given in the Cartesian coordinate system which is e.g. used internally in the CMS software. Or, in physical equations, the usual particle physics coordinate system is used where coordinates are expressed in the mass m, the momentum in the transverse plane p_{T}, the azimuthal angle ϕ and the pseudorapidity η

$$\eta \equiv -\ln\left(\tan\left(\frac{\theta}{2}\right)\right) \tag{2.2}$$

with the angle θ defined as the polar angle. ϕ is the azimuthal angle.

The main or leading primary vertex is the vertex where the quadratic sum of

transverse momenta of all particles associated to it

$$\sum_{i \in PV} p_{\mathrm{T},i}^2 \tag{2.3}$$

is the highest. The CMS detector is able to separate vertices down do a distance of 0.5 mm.

The **angular distance** between two particles i and j called ΔR is defined as

$$\Delta R_{ij} \equiv \sqrt{(\eta_i - \eta_j)^2 + (\phi_i - \phi_j)^2}. \tag{2.4}$$

The **I**mpact **P**arameter is defined as the shortest distance between the reconstructed vertex and the linear prolongation of a reconstructed track. All CMS analyses have to be developed *blinded*, meaning the analyst is not allowed to perform any searches in phase-spaces where the signal of interest is assumed. The choice of the "signal region" is however not universally defined. The CMS $H \to \tau\tau$ working group agreed on my proposal at the end of 2015 to use a significance-based bin-wise signal yield y_ϵ of

$$y_\epsilon = \frac{s}{\sqrt{b + (\epsilon b)^2}} \tag{2.5}$$

as parameter to determine which bins have to be blinded.

The parameter ϵ is analysis-dependent, reflecting the overall systematic uncertainty. In all bins with a signal yield of $y_\epsilon \leq 0.1$, data points are removed for both plotting and statistical inference.

With the high LHC design instantaneous luminosity of $\mathcal{L} = 10^{34}\,\mathrm{cm}^2\,\mathrm{s}^{-1}$, in each interaction several soft scatter processes happen apart from the hard scatter processes that are usually the physics of interest. These soft interactions are called **pile-up** interactions. The handling and suppression of pile-up interactions is an important topic for all analyses. The number of pile-up interactions $(N_{PU})_i$ in a single event i is given by

$$(N_{PU})_i = \frac{\mathcal{L}_i \cdot \sigma_{\text{min. bias}}}{n_b \cdot f} \tag{2.6}$$

with the instantaneous luminosity $\mathcal{L}_i$ of this specific bunch crossing, the number of bunches n_b and the revolution frequency $f = 11.2\,\mathrm{kHz}$. The minimum-bias cross section $\sigma_{\text{min. bias}}$ has been measured to be 69.2 mb.

2.2.1 The Solenoid

The superconducting solenoid is built around the inner detector. Inside it provides a homogeneous magnetic field of 3.8 T parallel to the beam pipes. The magnet is

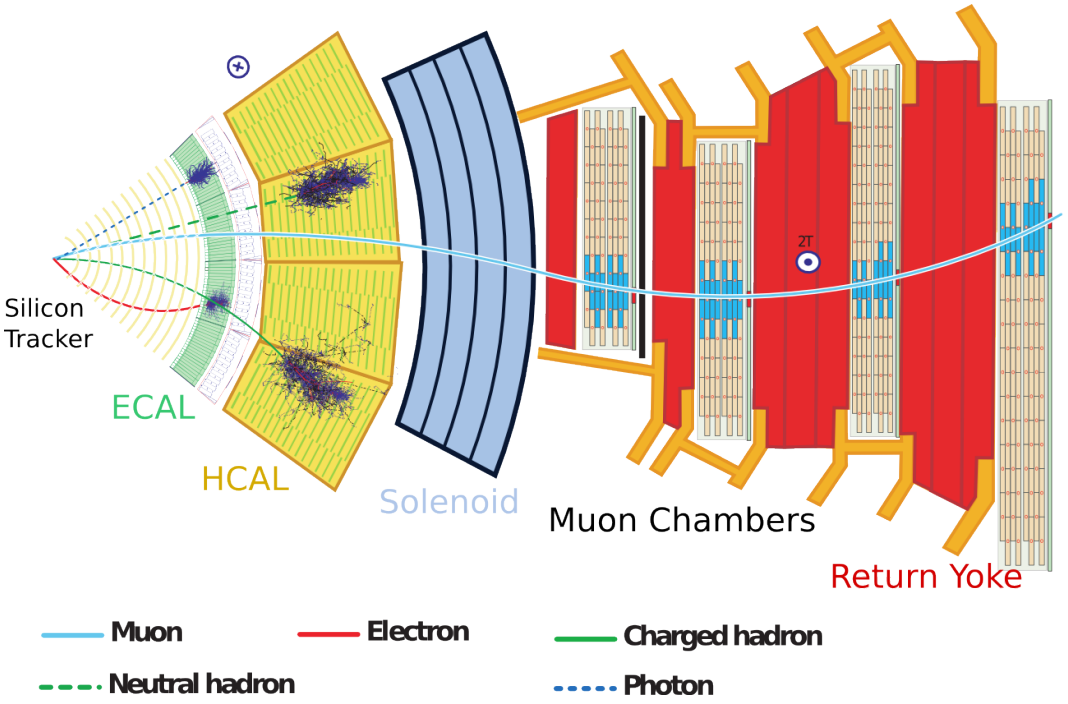


Figure 2.4: Part of a slice through the CMS detector, adjusted from [21]. Most stable particles are stopped in the corresponding systems, leaving calorimeter deposits. Only muons are not stopped, neutrinos stay undetected. Only charged particles leave a *track* (solid lines) that allows their assignment to a primary vertex.

made out of NbTi, being superconducting at the operation temperature of 4 K. The outer support structure serves as a return yoke for the magnetic field. The magnetic flux density there is around 2.4 T.

2.2.2 The Tracker

The *tracker* is there to reconstruct tracks of charged particles [22]. It surrounds the interaction point cylindrically. The whole tracker is based on silicon semiconductor technology, delivering the required radiation hardness at an affordable price. The pixels are p-doped on the downside and n-doped on the upside. Charged particles ionize the depleted region, which causes a measurable current. The main issue of reduction of depletion by radiation damage can be addressed by the application of a gradually increased bias voltage of up to 300 V.

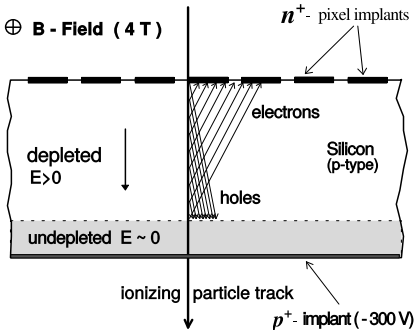


Figure 2.5: Working principle of a pixel module. Particles ionize the depleted silicon, with a bias voltage separating the charge that can be registered as a *hit*.

The innermost part is a pixel detector with pixel sizes of 100 µm × 150 µm and the closest modules at a distance of 40 mm to the primary vertex. The overlapping structure of the pixel modules in each layer ensures full coverage over its full extend of 1 m, covering in total a surface of $0.78\,m^2$ in three layers. With around 1000 tracks per bunch crossing, the occupancy of the 66 million pixels in the detector is around $O(10^{-4})$.

The silicon strip detector consists of ten layers around the pixel detector. It covers an area of about $200\,m^2$ and is operated at temperatures around $-20\,°C$ to limit radiation damage. The inner strips have a size of 10 cm × 180 µm while in the outer region they are prolonged to 25 cm. In total they deliver 9.6 million strip channels, showing an occupancy $O(10^{-3})$.

The spatial resolution depends on the angle of the reconstructed particle. The tracker spatial resolution is 20 to 65 µmetre [23].

The CMS Tracker allows to reconstruct tracks of charged objects down to a p_T of 150 MeV/c in a region under $|\eta| < 2.5$ and up to several hundred GeV/c. See figure 2.6 for a schematic view of the tracking system.

2.2.3 Electromagnetic Calorimeter

The **E**lectromagnetic **CAL**orimeter [25] consists of 75,000 crystals made of $PbWO_4$, each exposing only a surface of $4\,cm^2$ to the particles and covering a range in η up

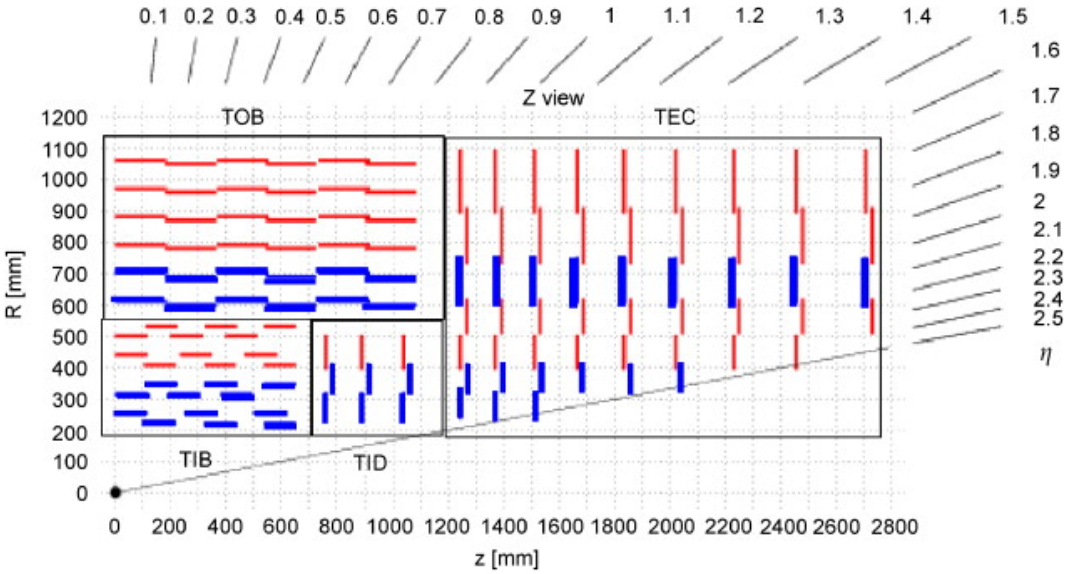


Figure 2.6: One quadrant in the longitudinal view of the CMS tracker. The blue modules are double-sided while the red ones are single-sided. The transition between the inner and the outer region is between $0.9 < \eta < 1.4$, the endcap coverage is up to $|\eta| < 2.5$. The strip detector parts start with a *T* for tracker, *I* stands for inner, O for outer. The last character shows the orientation of the modules. *B* stands for cylindrical architecture, called barrel. *D* stands for disk, *EC* for endcap. [24]

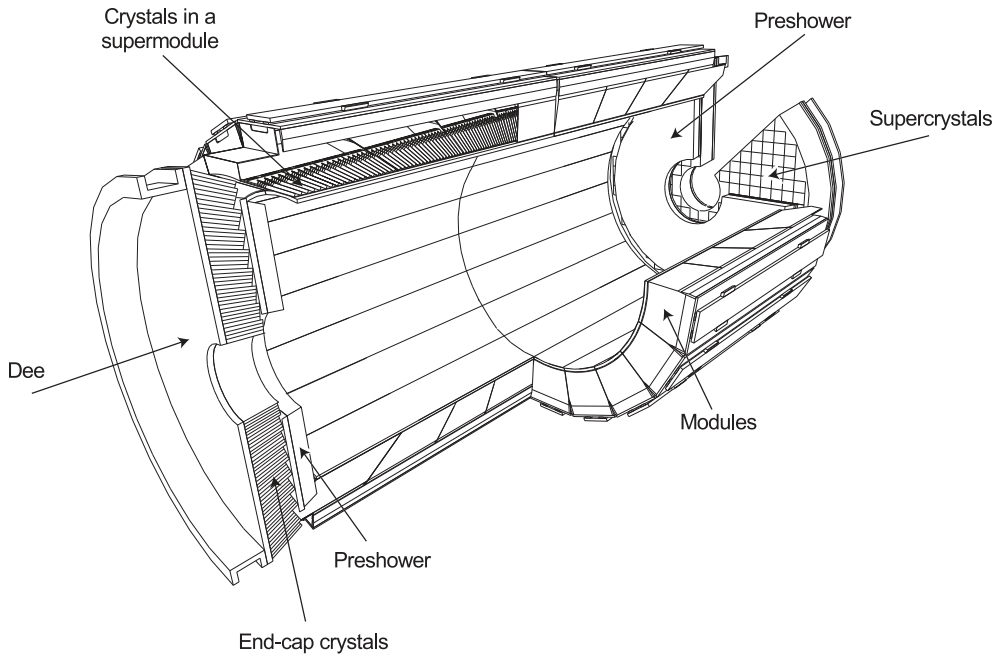


Figure 2.7: Sectional view of the ECAL. All modules point towards the nominal interaction point in both the barrel and the endcap region. The preshower with its thin lead radiators of 2 and $1X_0$ causes most of the photons to shower [25].

to 3.0. Lead tungstate has a short radiation length of $X_O = 8.9\,\text{mm}$ and a high density of $\rho = 8.28\,\text{g}\ /\text{cm}^3$, which makes it an ideal material for an electromagnetic calorimeter for the CMS experiment with its size limitation to fit everything within the solenoid. The modules in the barrel have a length of 23 cm, which corresponds to $25.8X_0$. The spatial resolution is $22\,\text{mm} \times 22\,\text{mm}$ corresponding to 0.0174×0.0174 in $\eta - \phi$. The fast light emittance can keep up with the collision rate of 40 MHz. The, in average, 4500 photoelectrons per GeV/c are read-out by avalanche photodiodes in the barrel and vacuum phototriodes in the endcaps. The radiation damage causes wavelength-dependent absorption processes, resulting in a reduced number of registered photons. This effect is measured with an injection laser and can be calibrated out.

While also having both a barrel and an endcap module, the ECAL also provides a **preshower** detector. It helps to identify neutral pions, to discriminate between electrons and minimal ionizing particles. Also, it provides additional spatial information about electrons and photons. This is especially necessary since photons have a high probability to convert to e^+e^- pairs within the tracker endcaps due to the large material budget in the barrel region (see figure 2.8).

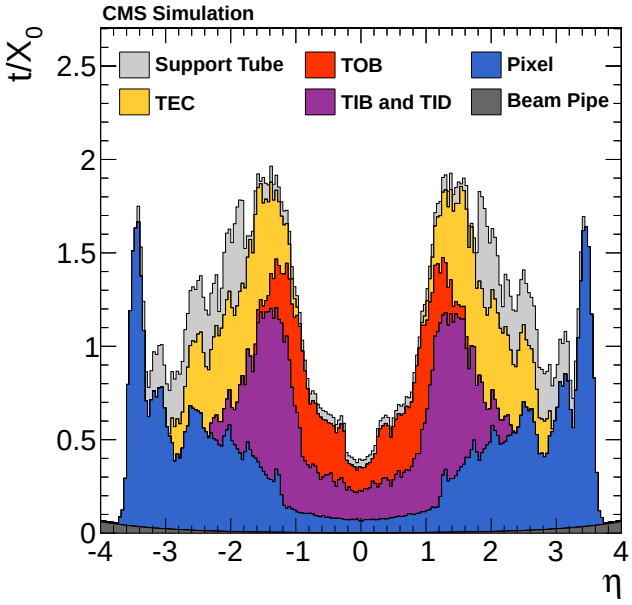


Figure 2.8: The material budget in front of the ECAL in units of interaction length X_0 that a particle must traverse. Most of the material is actually not active sensor material but rather there as support structure, cooling, read-out electronics etc. The large material budget between $|\eta| > 1$ and $|\eta| < 2.5$ is the region covered by the tracker endcaps. [26]

2.2.4 Hadron Calorimeter

Quarks and gluons from hard interactions can not be measured directly. Even though initially they can have several hundred GeV/c of momentum, the strong force potential favors the emergence of softer quarks, until energies are reached when quarks start hadronizing. This can be observed as a collimated stream of particles that due to its shape is referred to as a **jet**. The hadron calorimeter measures the energy of these jets, aiming to reconstruct the properties of the initial quark or gluon.

The **H**adron **CAL**orimeter surrounds the ECAL, with a distance of 1.7 m to the beam line, see figure 2.9. It is made out of brass and scintillators, having an interaction length of $\lambda_I = 16.42\,\text{cm}$. It has the same coverage as the ECAL, but with 0.087×0.087 in $\eta - \phi$ a spatial resolution 25 times higher than the ECAL. The plastic scintillators are connected with fibers to hybrid photodiodes. Depending on the region, hadrons are exposed to a minimum of $5.8\lambda_I$ up to more than $11\lambda_I$. Because for centrally produced, high-energetic hadrons a non-negligible fraction of energy is not deposited in the inner HCAL, an outer part surrounding the solenoid improves shower containment.

A forward hadron calorimeter (HF) extends the coverage of the hadron measurement up to $|\eta| < 5.0$. The HF has special needs to the radiation hardness, absorbing more than 7 times the radiation than the rest of CMS does. Steel is being used as absorber material with quartz fibers being the active material, giving a signal for charged particles above the Cherenkov threshold (190 keV for electrons). The fibers

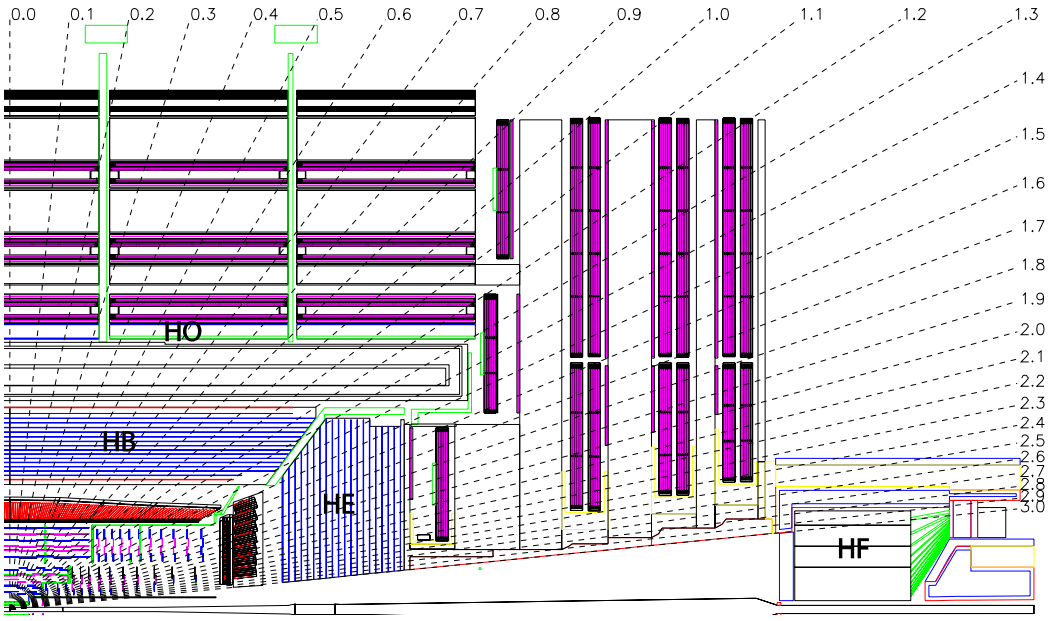


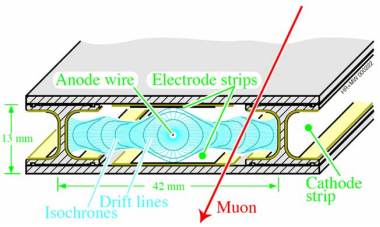
Figure 2.9: Sectional view of a fourth of CMS with the tracker and ECAL unlabeled and the HCAL components in the barrel (HB), endcap (HE), forward (HF) and outer (HO) [24] direction.

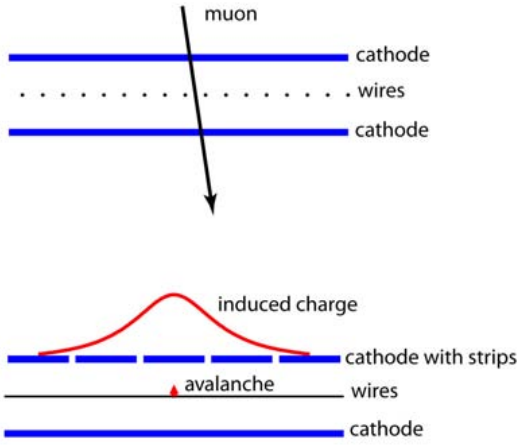
transport the emitted light to photomultipliers further away from the beam line to reduce radiation damage.

2.2.5 Muon system

The muon system is placed outside the solenoid, exposed to the return field between 1.8 and 2.5 T. Although not mandatory for the reconstruction of most muons, it helps in the identification. It extends the charge and momentum measurement to the TeV range, where the silicon tracker would only reconstruct nearly straight tracks resulting in degraded resolution. The muon system consists of three different types of gas detectors, see figure 2.10 [24, 25].

The **D**rift **T**ubes are made of aluminum, filled with gas and a wire in the middle. Ions in the gas are separated by a voltage of 3.6 kV on the wire and −1.2 kV on the cathodes. Different orientations of the tubes allow a combined measurement in the $r-\phi$ direction with a resolution of 100 µm and 150 µm in the $r-z$ plane. Traversing muons ionize the gas, while the applied voltage attracts these ions, resulting in a measurable current.





The **C**athode **S**trip **C**hamber modules are placed in the forward direction, covering $0.9 < \eta < 2.4$. CSCs are able to handle a large particle flux. In each chamber, seven strips run at a constant ϕ while six anode wires are mounted azimuthally. The CSC system is huge, covering a total area of 5000 m^2 with 2 million wires providing a spatial resolution of 2 mm for the L1-trigger and 75 µm in the $r-\phi$ plane for offline reconstruction. The muon system is completed by the **R**esistive **P**late **C**hambers. The RPCs are gaseous parallel-plate detectors with one plate being positively charged and the other one negatively charged. Muons hitting the gas atoms in between the plates cause an avalanche of electrons. The plates themselves are transparent to the electron avalanche, wich is detected by coarse strip system. The RPCs provide a very high time resolution of a few nanoseconds and are a valuable part of the CMS trigger system.

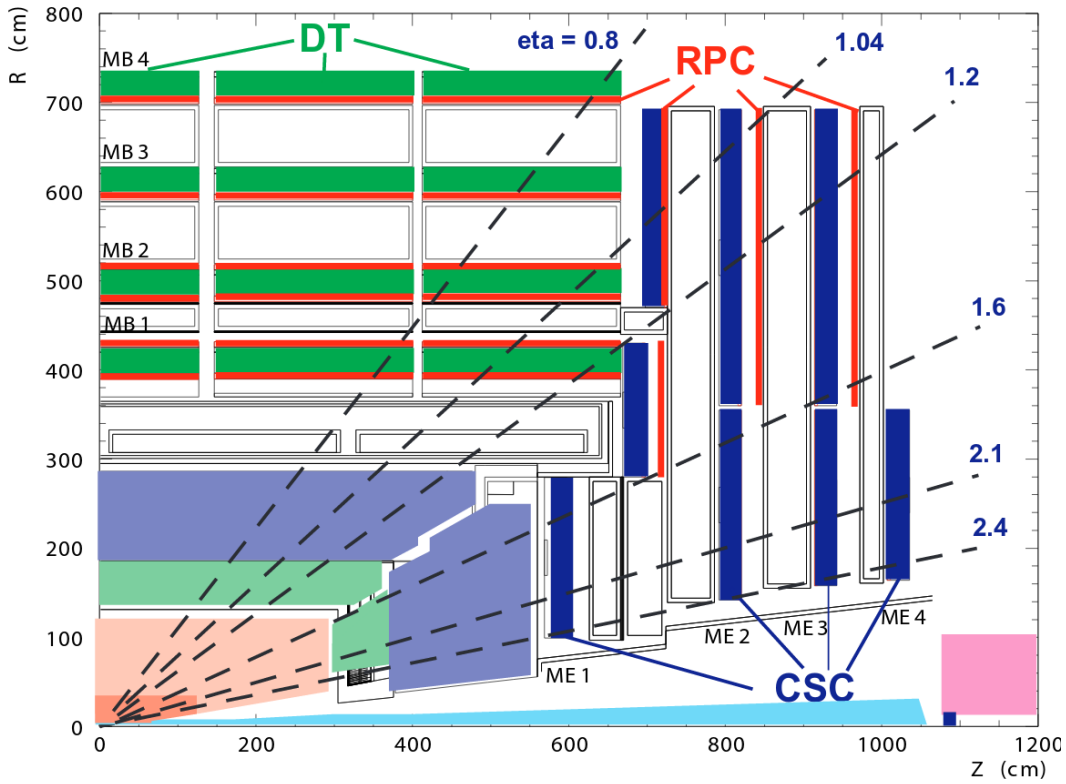


Figure 2.10: A quarter of the CMS muon system consisting of the drift tubes (DT), resistive plate chambers (RPC) and the cathode strip chambers (CSC). [27]

2.2.6 The Trigger System

The cross section of **minimum-bias** events is several orders of magnitude higher than the one of hard interactions. The term minimum-bias refers to any bunch-crossing leading to reconstructed particles in the detector. While the expected event rate is around a billion inelastic events per second, only a tiny fraction of it is subject to the LHC physics program. The physical motivation to preselect only interesting events is supported by the need to reduce the sheer amount of data the LHC produces. With millions of output channels delivering data at 40 MHz it is just impossible to save the full information from the detector at all times.

This gives rise to the need of a sophisticated **Trigger System** in CMS. It consists of two stages. The first one is the so-called **L1** Trigger. It is made by programmable electronics depending on information from the muon system and the calorimeters. It has to decide within microseconds, if an event is potentially interesting in terms of physics. This decision is based on the occurrence of an electron, muon, photon, jet or missing transverse energy. Since CMS is designed as a lossless system, any event recorded by the subdetector can be stored - no matter how long the trigger decision takes. A long waiting time for sure causes the buffers of subdetectors to saturate. In this case, the **T**rigger **T**hrotteling **S**ystem is activated, stopping all subdetectors from further measurements. Once the trigger decisions are there, the system continues as usual.

The maximum L1 rate is around 100 kHz. Events with that rate go into the **H**igh **L**evel **T**rigger system, a computer farm consisting of standard hardware with around 13.000 CPUs promptly doing a full event reconstruction. With the full tracker information available and already the identification and isolation algorithms being run, higher-level objects like taus can be used as trigger objects. It is also possible to combine several objects as a trigger requirement. The HLT rejects 99.9% of events, resulting in a manageable rate of $O(100\,\text{Hz})$. Exceeding this rate would mean that the output can not be stored fast enough, causing the TTS to pause and cause deadtime.

How many collisions per second really happen is not under the control of CMS but of the LHC team. During a run of typically a few hours, the event rate drops since the number of protons per bunch decreases exponentially with protons leaving the beam in the collisions. The trigger system compensates this, allowing its operator the choice of different *trigger menus* that are linked to so-called *pre-scales*. A pre-scale of e.g. 2 means, that only every second event fulfilling the requirements of a HLT is kept and stored. Lowering the pre-scales allows to increase the HLT rate if it falls below 100 Hz.

2.2.7 Data acquisition and data quality monitoring

The CMS experiment is controlled by a sophisticated **D**ata **Acq**uisition system. It takes control of all subdetectors and steers their data taking. The whole run is being monitored by several stages of **D**ata **Q**uality **M**onitoring, where the functionality of CMS is concurrently monitored with a set of control plots, giving the possibility to spot problems already during data taking.

2.3 Stable Particle Reconstruction

The CMS experiment follows a **P**article **F**low approach [28] for the reconstruction of an event. Particle flow means that measurements of different subdetectors are combined in an algorithm for the best possible reconstruction of individual stable particles by following the trajectory of each stable particle through the detector. This is a challenging concept, since in each event hundreds of individual stable particles are being reconstructed with many of them only carrying a few GeV/c of momentum. This reconstruction of individual particles allows advanced methods in terms of object reconstruction, cleaning and identification that is presented in the next sections.

The core component of the PF algorithm is the high efficiency in the tracker combined with a low misidentification rate. This was achieved by an iterative tracking algorithm. It starts with the reconstruction of tracks, fulfilling tight requirements on the track quality. The related hits are removed. Then, requirements are loosened and the reconstruction is ran again. This leads to both a high efficiency as well as a low fake rate. The muon efficiency reaches 99.5 % and the charged hadron efficiency 90 % in three iterations. Starting from the fourth iteration, an origin outside a small cone around the beam axis is allowed to also reconstruct tracks from secondary vertices up to 50 cm away from the beam axis.

The calorimeter deposits are clustered. Stable particles can deposit their energy in several calorimeter cells, which is accounted for by the clustering. Calorimeter clusters are summarized to particle flow clusters, being the starting point for the reconstruction of a particle flow object.

The **P**ile **U**p **P**er **P**article **I**dentification algorithm [29] is an approach to clean the Particle Flow collection from pile-up. The cleaned PF collection can then be used to run the usual reconstruction explained below, requiring minimal modifications. This is in contrast to other pile-up handling approaches, where quantities are corrected for pile-up influences.

The PUPPI algorithm does a re-scaling of each particle four-momentum. Particles identified as pile-up get ideally re-scaled to 0 and are therefore removed. Particles from the hard interaction get a weight of 1.0 and are kept with full momentum. The core of PUPPI is a local shape variable α_i assigned to each particle flow candidate i:

$$\begin{aligned} \alpha_i &= \log \sum_{j \in event} \xi_{ij} \times \Theta\left(R_{min} \leq \Delta R_{ij} \leq R_0\right) \\ \text{with } \xi_{ij} &\equiv \frac{p_{\mathrm{T}_j}}{\Delta R_{ij}} \\ \text{and } \Theta\left(R_{min} \leq \Delta R_{ij} \leq R_0\right) &\equiv \Theta\left(\Delta R_{ij} - R_{min}\right) \times \Theta\left(R_0 - \Delta R_{ij}\right) \end{aligned} \tag{2.7}$$

with the Heaviside step function Θ, the distance from formula 2.4 and the particle j lepton transverse momentum $\frac{p_{\mathrm{T}_j}}{\Delta R_{ij}}$. The parameter R_0 is a cone around the particle i, being a measure for the distance in which neighboring particles contribute. R_{min} excludes particles too close to ensure collinear safety, see chapter 2.3.3. The tracking information in the central region is used in a sense that it delivers a truth information for charged particles to have its origin in the leading primary vertex, leading to a weight $w_i = 0$ for pile-up and $w_i = 1$ for charged particles from the leading primary vertex.

A χ^2-like quantity is introduced

$$\chi_i^2 = \Theta\left(\alpha_i - \bar{\alpha}_{PU}\right) \cdot \frac{\left(\alpha_i - \bar{\alpha}_{PU}\right)^2}{\sigma_{PU}^2} \tag{2.8}$$

meaning values of α_i below the median $\bar{\alpha}_{PU}$ are considered pile-up. $\bar{\alpha}_{PU}$ and σ_{PU}^2 are extracted from the region where tracking information is available, on an event-by-event basis. The χ^2 quantity leads to the weight

$$w_i = F_{\chi^2,\mathrm{NDF}=1}\left(\chi^2\right) \tag{2.9}$$

with F_{χ^2} being the cumulative distribution function of the χ^2 distribution. The weight distribution in the central region of the detector and the resulting mean weight over p_{T} can be found in figure 2.11.

Once the weight α_i is there, the PF candidate 4-vectors are re-scaled with α_i. Particles with a weight or transverse momentum below a certain threshold are dropped, e.g. $w_{cut} < 0.1$ and $p_{\mathrm{T},cut} < 0.1\,\mathrm{GeV}/c$, with both parameters being dependent on the number of reconstructed primary vertices n_{PU}.

The PUPPI Algorithm nevertheless needs intense parameter tuning (R_0, R_{min}, w_{cut}, $p_{\mathrm{T},cut}$...). Using it in context for the $\not{E}_{\mathrm{T}}$ calculation has the necessity of two different PUPPI PF candidate collections, one incorporating identified leptons and the other dropping it. This results from the fact that in the vicinity of leptons, all particles get a high weight, which leads to an imbalance of the $\not{E}_{\mathrm{T}}$ towards the lepton, showing up as a reduced response. For the PUPPI $\not{E}_{\mathrm{T}}$ performance see chapter 3.

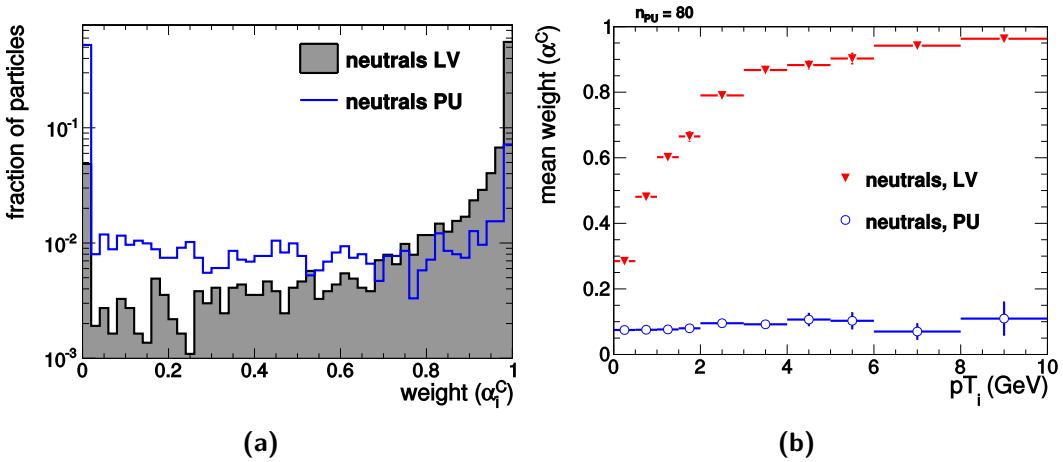


Figure 2.11: (a) PUPPI weight for a many events for neutral particles with $p_T > 1\,\text{GeV}/c$. Many particles get a clear assignment to either a way 0 or 1. (b) The mean weight over many events of neutral particles from the leading vertex (red) and pile-up (blue) as a function of the particle p_T. Low-transverse momentum particles are harder to classify than the ones with higher transverse momentum.

2.3.1 Muon reconstruction

Since muons interact with all subdetectors, there are several ways for their reconstruction. CMS follows two ways: The *outside-in* and *inside-out* reconstruction, getting the name from the system the reconstruction was started from.

The outside-in reconstruction starts with the hits in the muon system, combining tracks reconstructed in the muon system with tracks reconstructed in the tracker using a Kalman-filter technique [26]. This is especially efficient at high transverse momenta. The downside is that when the muon reaches the muon systems, energy loss might have already happened leading to a degraded resolution.

The inside-out muon track reconstruction starts with all tracks of at least $p_T > 0.5\,\text{GeV}/c$ and a total momentum of $2.5\,\text{GeV}/c$. These tracks are extrapolated to the muon system under consideration of the magnetic field and interactions with the detector material leading to elastic and inelastic scattering. If only one segment from the muon system matches the extrapolated tracker track, it is identified as a muon. Because only one hit in the muon systems is required, the efficiency is higher than in the outside-in approach, where three hits are required to do an extrapolation towards the tracker.

Muons from the inside-out approach are matched with the ones from the outside-in one. The matching muons get a special quality flag and called *global muons*. Muons only reconstructed from outside-in tracks not having a corresponding inside-out track are usually dropped. The global track has to fulfill certain χ^2 requirements to be

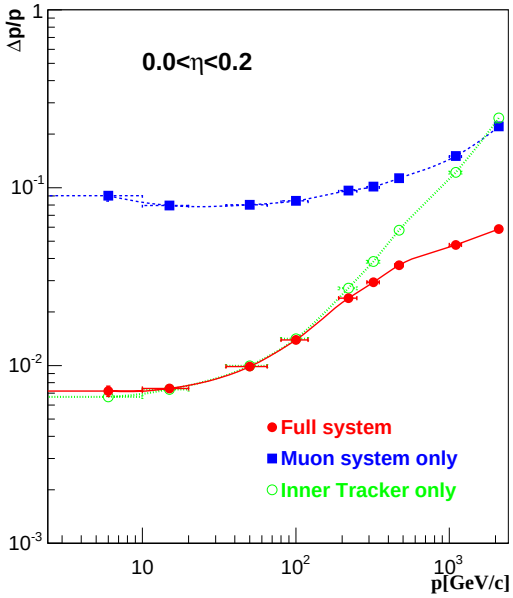
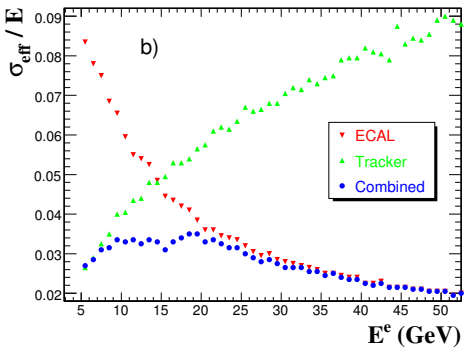


Figure 2.12: Relative muon resolution in central regions. Even though the muon system has inferior resolution in all momentum regions compared to the tracker, a combination of both improves the resolution significantly starting from $p_T > 200\,\mathrm{GeV}/c$ [27].

accepted. See Figure 2.12 for a comparison of the resolution of both reconstruction approaches and their combination.

2.3.2 Electron and Photon reconstruction

Electrons as well as photons are mainly deposits in the ECAL with electrons potentially having a track. Due to the high material budget in front of the ECAL (figure 2.8) electrons emit bremsstrahlung and may underly photon conversion. In the reconstruction of electrons, there are two cases. The first one is the one with the best possible resolution. An unconverted electron enters the ECAL starting to shower. 97% of its energy is stored within 5×5 ECAL crystals, making it possible to sum up the deposited energy. The same is true for unconverted photons.



Fractional resolution over generated energy. The ECAL measurement is designed to be precise for high energies with the tracker keeping precision down to energies of $5\,\mathrm{GeV}/c^2$ [27].

The emission of bremsstrahlung and the conversion spreads the initial electron energy among several ECAL cells among the ϕ direction due to the strong

magnetic field within the CMS detector. A two-step clustering algorithm combines energy deposits to clusters and assigns them with the measured electron track to reconstructed electrons. Corrections are applied that account for geometric differences of the ECAL, like the different material budget traversed by the electrons. Very high energetic particles above 1.7 TeV in the barrel and 3 TeV in the endcap may saturate the Pre-Amplifier, a limitation that is moderated by several algorithms based on a fit done on neighboring cells.

2.3.3 Jet reconstruction

Unlike muons and electrons, that are well localized and may only have some radiation that has to be assigned to them, jets from quarks or gluons have a complex shape and energy distribution. This is due to the parton showering, where an initial quark or gluon produces additional particles to satisfy the color confinement. One is usually only interested in the initial quark or gluon and wants to reconstruct its properties. The jet clustering algorithm plays the role of ensuring that the reconstructed jet is independent of the details of the showering itself. The two requirements to meet are

- *collinear safety*: The resulting jet properties should not change whether a particle in a jet is there itself or if it has decayed in e.g. two particles carrying half of the momentum each.
- *infrared safety*: The hadronization can go so far that even if the initial quark or gluon had a high momentum, some of its decay products might carry a small momentum or even stay undetected at all. The jet algorithm has to make sure that the jet reconstruction does not change if these infinitely soft particles are added or left out.

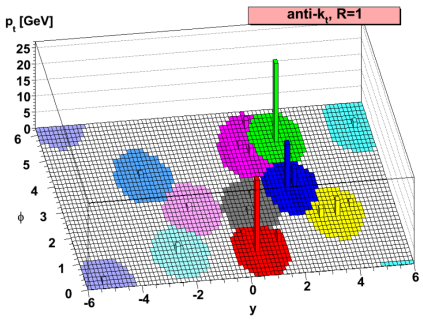
The standard algorithm for jet reconstruction in CMS is the anti-k_T algorithm [30]. It is a modification of the k_t [31] and Camebridge-Aachen [32] algorithms with the following distance definitions:

$$\begin{aligned} d_{ij} &= min\left(p_{\mathrm{T},i}^{-2}, p_{\mathrm{T},j}^{-2}\right)\frac{\Delta R_{ij}^2}{R^2}, \\ d_{iB} &= p_{\mathrm{T}i}^{-2} \end{aligned} \tag{2.10}$$

with the usual definition of ΔR used in particle physics from formula 2.4.

The anti-k_T algorithm does, in contrast to previous algorithms, choose a negative exponent -2 for the distance metric d_{ij} between two particles. d_{iB} can be seen as a distance to the beam line. R is a parameter to be chosen, reflecting the jet size. This was 0.5 in the LHC Run I and 0.4 in Run II. Special cone sizes for jet substructure analyses can go up to 0.8.

The jet clustering algorithm compares for each particle combination i and j if d_{ij} is smaller than d_{iB}. If this is the case, the four-vectors of both are added, fulfilling the collinear safety requirement. If not, the particle i is considered a jet and no further merging is done. This is iteratively done until no further merging is possible. What happens is, that first harder particles are merged with the neighboring ones. If a cluster consists only out of soft particles, they are merged together until they may reach the threshold to be accepted as a jet. The order does not play a role in this case, making the algorithm infrared-safe.



Resulting jets from the anti-k_T algorithm. Even though not explicitly required, the jets are mostly cone-shaped.

The anti-k_T algorithm is ran on the collection of PF candidates. To suppress pile-up, the **C**harged **H**adron **S**ubtraction technique is being used, meaning charged particles not associated to the primary vertex are subtracted before running the jet clustering algorithm.

Even though PF candidates might already have been identified as candidates for being leptons, they are also kept for the clustering. Many of them carry high momentum, letting them act as starting points for jets. This results in the fact that the same PF candidate can end up e.g. being a candidate for an electron, jet and even a tau. The analyses have to takes care of this ambiguity by removing jets matching e.g. an identified electron from the jet collection.

Since the anti-k_T algorithm theoretically would have a complexity of $O(N^{3/2})$ the **FastJet** [33] implementation managed to keep $O(N \ln N)$ complexity for N in the usual range for the CMS experiment.

Jet Energy Corrections

The determination of the energy of an initial parton can only be done in an indirect way, see figure 2.14. This indirect measurement has several biases, which in order to measure the initial parton as well as possible, need to be addressed. The **J**et **E**nergy **C**alibration is done in a multiplicative way with the scalar correction factor $\mathscr{C}$ so that the initial four-momentum p_{jet}^{raw} becomes p_{jet}^{cor}.

$$p_{jet}^{cor} = \mathscr{C} p_{jet}^{raw} \tag{2.11}$$

The correction factor consists of four factors that are introduced in the following.

- **Offset corrections**: This first step removes contributions from the jet not

caused by the hard interaction but by one of the pile-up interactions as well as the contributions from detector noise. The pile-up contribution is estimated with the *Jet area* method. This method calculates an effective p_{T} density ρ. Under the assumption that it is uniformly distributed, the contribution to a jet is proportional to its area.

- **Monte Carlo Calibration**: Various effects cause the reconstructed jet momentum to be different from the generated one. A response variable $\mathscr{R} = p_{\mathrm{T}}^{reco}/p_{\mathrm{T}}^{gen}$ is introduced to compensate this effect, caused by particles leaving the jet (*out-of-cone*) or reconstruction efficiencies. The Monte Carlo Calibration is applied on both data and simulation, hence not taking care of potentially biasing differences in simulation and data. The inverse of the expectation value of $\mathscr{R}$ is considered as the Monte Carlo correction factor $\mathscr{R}_{MC}(p_{\mathrm{T}}^{reco})$.

- **Relative Jet Energy Scale**: Using a special di-jet selection, differences in the jet response relative to η can be estimated. In the barrel region they are small, but non-negligible when going to the forward region, see Figure 2.13. The relative Jet Energy scale also does the residual corrections, making sure the jets are described comparably in both data and simulation.

- **Absolute Jet Energy Scale**: The final calibration of the absolute jet energy scale is done using objects with a better resolution than jets: photons, electrons or muons from the Drell-Yan process, having no neutrinos involved. There are two established, complementary methods. The **M**issing $\not{E}_{\mathrm{T}}$ **P**rojection **F**raction assumes a perfect balancing between the true transverse momentum of the reconstructed di-muon system and the recoiling jet. Each object has its response and every mismeasurement is assigned to the $\not{E}_{\mathrm{T}}$, leading to the response $\mathscr{R}_{MPF}$

$$\mathscr{R}_{MPF} = 1 + \frac{\not{E}_{\mathrm{T}} \cdot \vec{p}_{\mathrm{T}}{}^{z}}{\left(p_{\mathrm{T}}^{2}\right)^{2}} \tag{2.12}$$

  The p_{T} balance method on the other hand defines the response as the ratio of transverse momenta of the Z boson and the jet:

$$\mathscr{R}_{bal} = \frac{p_{\mathrm{T}}^{jet1}}{p_{\mathrm{T}}^{Z}} \tag{2.13}$$

Jet tagging and assignment

The jet reconstruction does not give information on the nature of the jet at all. A jet may originate from light or heavy quarks, gluons or hadronically decaying taus

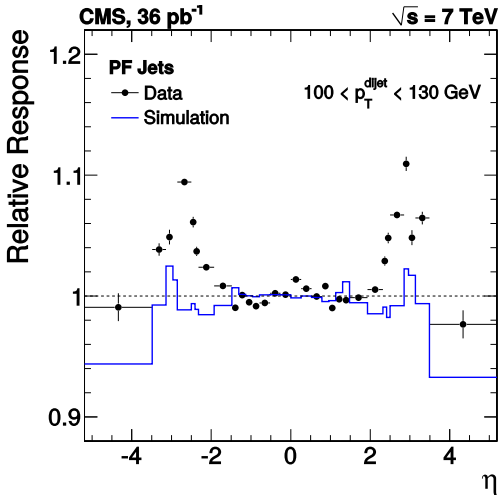


Figure 2.13: Relative response of a di-jet system with one jet in the central region and the other at higher η [34].

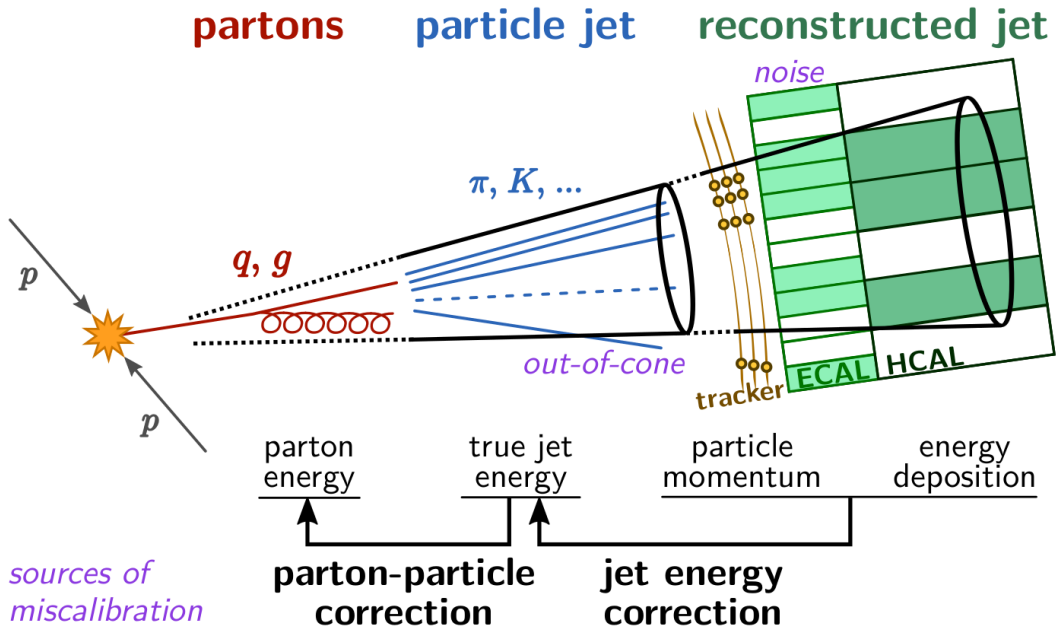


Figure 2.14: Schematic overview of a jet measurement and the effects to be calibrated [35]

and it might have its origin in the hard scatter of the event or it may originate from pile-up collisions.

The identification of tau jets is discussed in the next section. The most important *flavor tagging* algorithm is the **b-tagging** one called **C**ombined **S**econdary **V**ertex. It aims at the identification of jets originating from a b-quark. The b-quark has a relatively long lifetime of $\tau = 1.5 \times 10^{-12}$ s. With the γ factor

$$\gamma = \frac{1}{\sqrt{1 - v^2/c^2}} = \frac{E}{m_0 c^2} + 1 \tag{2.14}$$

with the kinetic Energy E in the GeV range and the rest mass m, a resulting decay length of $\gamma \cdot 450\,\mu\text{m}$ is in the order of millimetres, leading to a high impact parameter and a displaced secondary vertex. b-jet decay is also more likely to be accompanied by semileptonic decays, meaning there are electrons or muons with high transverse momenta. The CSV algorithm combines these quantities, providing a discriminator where higher values relate to a higher probability of the jet to originate from a b-quark. The b-tagging reaches 85 % efficiency with a misidentification rate of 10 % for the medium working point and 70 % efficiency for the tight one with a misidentification rate of only 1.5 % [36]. Differences in the b-tagging efficiency are compensated using b-tag scale factors.

The identification of pile-up jets is especially important for high-pile up regimes. The **MVA Pile-up jet ID** [37] is an approach to discriminate jets originating from the hard scatter from pile-up jets. In four $|\eta|$ regions, jets are classified with a BDT-based discriminator, depending on the pile-up environment, the jet radius profile and the charged and neutral multiplicities. Also, the compatibility of tracks within the jet contains information on the origin of the jet.

2.3.4 Tau reconstruction and identification

The tau lepton decays with its short lifetime of $\tau = 2.9 \times 10^{-13} s$, leading to a decay length of $c\tau = 87.11\,\mu\text{m}$. Even though the tau decays electroweak, the lifetime is so short that it is a challenge for the CMS tracker to resolve its secondary vertex. A tau lepton decays in most cases hadronically, see Table 2.1.

In the case of electrons and muons originating from prompt decays (e.g. $Z \to \mu\mu$) or as well leptonic tau decays, they are expected to be isolated, meaning there is no or few other hadronic activity in their vicinity. This is different from in-flight decays of e.g. pions, kaons or heavy flavor jets where the leptons are part of the collimated particle stream. This is distinguished by an isolation variable I^L with $L = e, \mu$, defined as

$$I^L = \sum_{\text{charged, PV}} p_\text{T} + max\left(0, \sum_{\text{neutral}} p_\text{T} - \Delta\beta\right) \tag{2.15}$$

where the sums are performed within a cone of $\Delta R = 0.4$ around the lepton. The $\Delta\beta$ term corresponds to the pile-up correction. The energy deposit from charged particles can be tracked and therefore only the these charged particles from the lepton's primary vertex are considered. The neutral sum however incorporates also pile-up. Since it has been estimated that both charged and neutral particles in the hadronization process of inelastic proton-proton scattering carry about the same energy in the final states of pile-up interactions, the neutral transverse momentum deposit around the lepton is corrected by half of the transverse momentum deposit of the charged particles in the cone, meaning that

$$\Delta\beta = \frac{1}{2} \sum_{charged,PU} p_{\mathrm{T}}. \tag{2.16}$$

Hadronic tau decays are usually denoted with the symbol τ_{had} or even only τ_h. Hadronically decaying taus are reconstructed with the so-called **H**adron **P**lus **S**trips [38] algorithm that is based on the PF algorithm.

The HPS algorithm starts with each anti-$kT4$ jet, being possibly a π^0 from a hadronically decaying tau. Electron/Positron pairs from conversion photons are bended in the magnetic field. This causes a broadening of the calorimeter signature of neutral pions in the azimuthal direction. The result are **strips** out of electromagnetic particles. They get identified in an iterative way, starting with the most energetic electromagnetic particle. The strip is then defined by the second most energetic particle close to the first one. The momentum is re-calculated using all particles contained in that strip. This procedure is repeated until no new particles can be assigned to the strip, going down to a threshold of $1\,\mathrm{GeV}/c$ per particle.

The HPS algorithm can distinguish between the four decay topologies single hadron ($h^{\pm}$), hadron plus one ($h^{\pm}\pi^0$) and two strips ($h^{\pm}\pi^0\pi^0$) and three hadrons ($h^{\pm}h^{\mp}h^{\pm}$), see table 2.1 for a summary. Figure 2.15 shows the reconstructed mass $m_{vis}^{\tau_h}$ in comparison between data and simulation in a certain decay channel of the $H \to \tau\tau$ analysis.

The reconstructed four-momentum and mass has to be consistent with the intermediate meson resonances. The reconstructed tau collection is finally nearly identical with the jet collection. A tau can never be reconstructed unambiguously, since there are three sources caus-

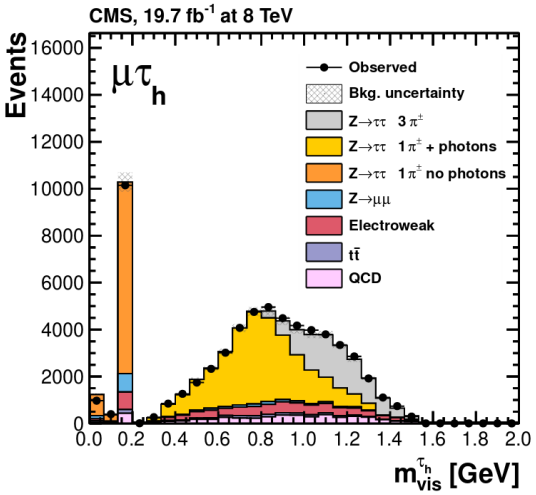


Figure 2.15: Reconstructed visible τ_h mass. The selection is introduced later in section 2.4.6 [39]

ing a signature close to the one of the tau:

- $jet \rightarrow \tau_h$: This background is caused by jets from other processes, e.g. quark or gluon jets. They also decay to charged and neutral hadrons. Two rejection methods have been developed, a cut-based and a MVA based approach. The cut-based approach using the isolation within the angular distance of $\Delta R = 0.5$

$$I^{\tau} = \sum_{\text{charged, PV}} p_{\text{T}} + max\left(0, \sum_{\gamma} p_{\text{T}} - \Delta\beta\right) \tag{2.17}$$

  and a modified pile-up suppressing $\Delta\beta$ correction

$$\Delta\beta = 0.46 \sum_{charged,PU} p_{\text{T}} \tag{2.18}$$

  within a cone of $\Delta R = 0.8$. The MVA based approach uses this information together with impact parameter variables and re-fitted distances and significances of the primary and secondary tau vertex. A large distance of the primary and secondary tau vertex is a hint for a non-prompt decay.

- $\mu \rightarrow \tau_h$: Muons have a good chance to be reconstructed in the $h^{\pm}$ decay mode. A loose anti-muon discrimination is to ask for at most one track segment in the muon systems to found compatible with the τ_h or to require at least 20% of its energy deposited in the calorimeters. Another, more strict approach additionally vetoes all hits in the muon systems compatible with the hadronic tau candidate.

- $e \rightarrow \tau_h$: Electrons, in addition to the muons, potentially radiate photons from bremsstrahlung converting to π^0, that might end up being reconstructed in the $h^{\pm}\pi^0$ mode. Since electron separation is more challenging, an MVA-based anti-electron discriminator is used, taking ECAL, HCAL deposits as well as geometrical information on the decay and electron reconstruction quality into account.

2.3.5 Variable Regression with Gradient Boosting

Boosted decision trees (BDTs) are widely used in the particle physics community. They provide a binary classifier, allowing the distinction of two classes as a function of input vectors. The most common implementation is part of ROOT and called Toolkit for Multivariate Analysis (TMVA, [40]). The popularity of BDTs is driven by the robustness of the algorithm concerning the input vectors and training parameters,

Table 2.1: Branching fractions of the dominant hadronic decay modes. K decays may also occur, but are not reconstructed separately.

Decay mode	**Resonance**	**Mass** MeV/c^2	**Branching fraction** (%)
$\tau^- \to \pi^- \nu_\tau$			11.6
$\tau^- \to \pi^- \pi^0 \nu_\tau$	ρ	770	26.0
$\tau^- \to \pi^- \pi^0 \pi^0 \nu_\tau$	a_1	1200	9.5
$\tau^- \to \pi^- \pi^+ \pi^- \nu_\tau$	a_1	1200	9.8
$\tau^- \to \pi^- \pi^+ \pi^- \pi^0 \nu_\tau$			4.8

the fast training of usually only a few hours on a single CPU and the low tendency to overtraining. The rejection of electrons being misidentified as hadronically decaying τ_h is an example for the usage of BDTs or the classification of the signal in the Run I $H \to \tau\tau$ analysis in the $\mu\mu$ final state[39, 41].

The regression of continuous values however is not yet that common. The concept of the **G**radient **B**oosted **R**egression **T**rees is introduced in the following.

A regression takes a vector of input variables $\vec{X} = (x_1, x_2, ..., x_i)$ and assigns it a prediction $f(\vec{X})$. The GBRT algorithm is a supervised learning technique, meaning both the input vector $\vec{x}$ as well as the true value or target Y is known.

While the linear regression uses linear functions as f, the function used here is

$$f(\vec{x}) = \sum_{m=1}^{M} r_m I(\vec{x} \in A_m) \tag{2.19}$$

with A_m standing for a region in the phase space covered by $\vec{x}$ and the mean value of the targets in the phase-space region r_m. The function $I(\vec{x} \in A_m)$ is zero if $\vec{x}$ is contained in A_m and 0 otherwise. The estimation therefore is defined by the definition of the sub-spaces A_m. These regions are here represented with rectangular cuts, hence they can be represented by a decision tree.

The metric to probe the quality of a set of cuts is to calculate the mean squared error between the target value of each input vector Y_i and its regressed value, being the same as the mean value of targets r_m in the associated phase-space A_m is a quadratic loss function $\mathcal{S}$. The separation power by adding an additional split

$$\frac{\mathcal{S}_P - (\mathcal{S}_L + \mathcal{S}_R)}{\mathcal{S}_P} \tag{2.20}$$

is aimed to be minimized. This is done by calculating the separation gain of each input variable in the vector $\vec{x}$ with a given granularity. Figure 2.16 shows an exemplary dataset and the categorization for the first tree.

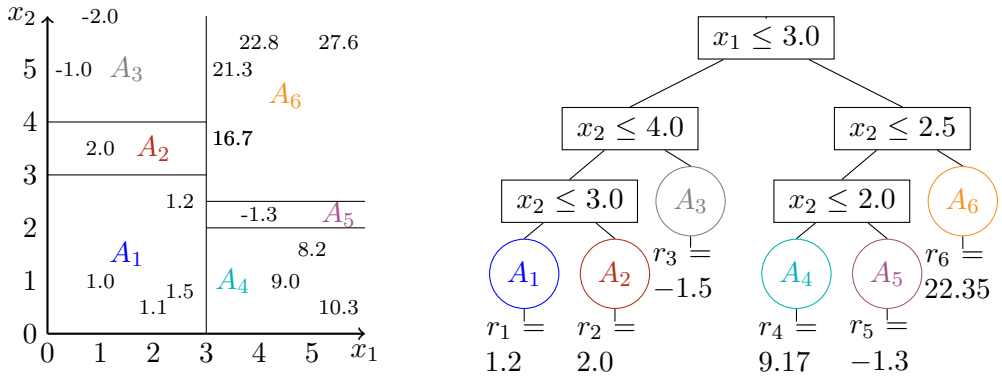


Figure 2.16: An exemplary dataset with 15 two-dimensional input vectors $\vec{X} = (x_1, x_2)^T$ and the 15 target values. The borders of the cut search are marked with lines and the resulting phase-space region are labels with $A_m, m \in \{1..6\}$. The tree has a depth of three. In a real training, one usually requires a minimal number of events per leaf, which is here left out for simplicity. On the right the corresponding representation in a tree is shown with the regressed value at the end.

After the first tree is trained, the boosting algorithm takes over. For binary BDTs a re-weighting of the input vector is done, leading to a higher importance of mis-classified vectors. This is very different for gradient boosting.

A loss function is defined, giving both a penalty to vectors away from the target value without biasing the result too much in case of outliers. A common choice is the Huber loss function [42]

$$\mathcal{L}(F(\vec{X}), Y) = \begin{cases} \frac{1}{2}(Y - F(\vec{X}))^2 & \text{for} (Y - F(\vec{X})) \leq \delta \\ \delta(\left|Y - F(\vec{X})\right| - \delta/2) & \text{for} (Y - F(\vec{X})) > \delta \end{cases} \tag{2.21}$$

for each event i a so-called pseudo-residual r_{it} is calculated:

$$r_{it} = -\left(\frac{\partial \mathcal{L}\left(F(\vec{X}), Y\right)}{\partial F(\vec{X})}\right)_{F(\vec{X})=F_{t-1}(\vec{X})} \tag{2.22}$$

The training of the second and all consecutive trees $f_t(\vec{X})$ uses these pseudo residuals as regression targets. The idea is, that they tend to point towards the true value. The step size towards the target, called shrinkage or learning rate ν, can be chosen. A small shrinkage leads to a slow convergence towards the true value. A high shrinkage has the risk of overtraining and its results may fluctuate. Therefore, a low learning rate combined with a high number of trees us generally favorable.

With the single tree $f_t(\vec{X})$ trained, the forest is updated to

$$F_t(\vec{X}) = F_{t-1}(\vec{X}) + \nu f_t(\vec{X}) \tag{2.23}$$

There are several stopping criteria of which one of them has to be true to stop the training process. One is reaching the maximum number of trees. This is the rather suboptimal case, since it means that either the training is not converging at all or that additional trees would improve the regression. The other criterion is the separation falling below a certain threshold. The setting of this threshold is important: If it is too high, the training stops very early and did not yet reach optimal performance. If it is too low, the forest gets sensitive to statistical fluctuations in the training dataset.

The implementation used for the $\not{E}_{\mathrm{T}}^{\mathrm{MVA}}$ (see chapter 3) is part of the CMSSW framework and has been implemented by Joshua Bendavid (California Institute of Technology). As a pre-processing step it transforms all input variables to Gaussian distributions. This is beneficial for an efficient cut search strategy, since the effect of potential outliers becomes negligible. The cut search itself is parallelized with OpenMP[1]. This is trivially possible, since the cut search is performed for each variable independently.

[1] http://www.openmp.org/

2.4 The Discovery of the Higgs Boson

After being predicted already for decades, many experiments like LEP and TEVATRON have already been searching for the Higgs Boson. Since the only missing parameter has been its mass, scanning the mass range was the most crucial part in all searches. Due to known exclusion limits from previous experiments and from theoretical calculations, masses in the range between $m_{\mathrm{H}} = 110\,\mathrm{GeV}/c^2$ and 1 TeV have been considered. This resulted in a wide range of searches, each targeting its own mass range, see figure 1.8.

On 4$^{\mathrm{th}}$ of July 2012, the discovery of a new particle compatible with the Standard Model Higgs boson was announced from both the CMS and ATLAS experiments [43, 44]. Already one year later, Peter Higgs and François Englert have been awarded the Nobel price in physics. In the following I will only discuss the final results of the LHC Run I of the CMS experiment containing the full datasets at 7 TeV and 8 TeV. CMS has covered five main decay channels, each summarized in the following.

2.4.1 Statistical Method

The statistical inference is based on the classical hypothesis test, meaning a comparison of the signal plus background expectation $(s + b)$ with the background-only expectation (b) is done. The Likelihood $\mathcal{L}_{s+b}$ and $\mathcal{L}_b$ are calculated with the following formulas [2]:

$$\begin{aligned} \mathcal{L}_{s+b} &= \prod_{k=1}^{N} \left(\frac{(s_k + b_k)^{n_k}}{n_k!} e^{-(s_k+b_k)} \cdot \prod_{j=1}^{n_k} \frac{s_k S_{kj} + b_k B_{kj}}{s_k + b_k} \right) \\ \mathcal{L}_b &= \prod_{k=1}^{N} \left(\frac{b_k^{n_k}}{n_k!} e^{-b_k} \cdot \prod_{j=1}^{n_k} \frac{b_k B_{kj}}{b_k} \right). \end{aligned} \tag{2.24}$$

With the index k representing all independent measurements. Independent means, that there is no overlap in measured events, be it because the run periods are different, the decay channel or the categorization, in question made sure by special vetoes. n_k is the number of observed events, s_k and b_k the number of predicted signal and background events. S_{kj} and B_{kj} are the probability functions to observe an event k in the bin j, while S_{kj} depends in the predicted Higgs boson mass m_{H}: $S_{kj}(m_{\mathrm{H}})$.

The test statistic is a likelihood ratio Q, given by

$$\lambda = \frac{\mathcal{L}_{s+b}}{\mathcal{L}_b} = \prod_{k=1}^{N} \left(e^{-s_k} \cdot \prod_{j=1}^{n_k} \frac{s_k S_{kj} + b_k B_{kj}}{b_k B_{kj}} \right) \tag{2.25}$$

Systematic uncertainties are taken into account by nuisance parameters. They modify the expected yields s_k and b_k. E.g. the normalization of all background

processes depends on the integrated luminosity. Therefore the background yield is modified by a function $f(\theta_k, \sigma_k, x)$

$$b_k \to b'_k \cdot f(\theta_k, \sigma_k, x) = \begin{cases} \frac{b_k}{\sqrt{2\pi}x\sigma_k} e^{-(\ln(x)-\theta_k)^2/\sigma_k^2} & x > 0 \\ 0 & x \leq 0 \end{cases} \tag{2.26}$$

with θ_k being the best known value of b_k and its uncertainty σ_k. x is integrated out by a numerical integration using effectively a *toy based* method which will be explained later. To turn the likelihood ratio into a minimization problem, usually the negative log likelihood is used, $q(\lambda)$:

$$q(\lambda) = -2 \ln \lambda. \tag{2.27}$$

This leads to an effectively weighted sum of events, according to the expectation of the corresponding bin, to be $s + b$ or b-like.

$$q = -2 \sum_{k=1}^{N} \left(s_k - \sum_{j=1}^{n_k} \ln \left(1 + \frac{s_k S_{kj}}{b_k B_{kj}} \right) \right). \tag{2.28}$$

The combination of all experiments leads to a measured value q. The uncertainties are taken into account by a large number of pseudo-experiments (*toy study*), each one estimating how q would have looked under the hypothesis of the random variable x. The random variable x transforms to a value of b'_k (equation 2.26), an approach usually called numeric integration. By design, the overall normalization stays untouched, since equation 2.26 is a probability distribution function. All correlated uncertainties of the combined measurement are evaluated using the same random variable x. This is valid for all b_k in case of e.g. the luminosity uncertainty and a subset in case of the cross section uncertainties. Thus, in this model parameters can only either be fully correlated or not correlated at all when using a different random variable.

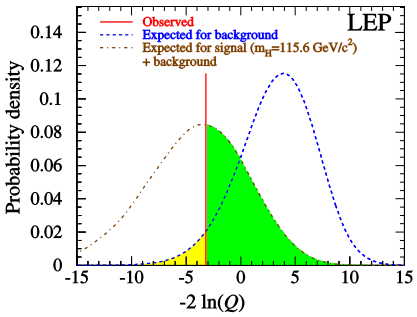


Figure 2.17: Example for $\mathcal{P}_{s+b}$ (brown curve) and $\mathcal{P}_b$ (blue curve) at LEP for a Higgs mass hypothesis of $m_{\mathrm{H}} = 115.6\,\mathrm{GeV}/c^2$. The yellow area corresponds to $1 - CL_b$, the green to CL_{s+b}. [45]

The median for both the $(s + b)$ and (b) hypothesis for large numbers of pseudo-experiments define the value q. By combining all pseudo-experiments, the probability

density functions $\mathcal{P}_{s+b}$ and $\mathcal{P}_b$ are obtained. With them, the conclusions one is interested in can be obtained.

To what confidence level is the observation q_{obs} caused by the known backgrounds?

The number in question is called the CL_b confidence, given by the integral

$$CL_b = \int_{q_{obs}}^{+\infty} \mathcal{P}_b. \tag{2.29}$$

How large is the probability that the observation is caused by a background fluctuation?

This is just the integral from on the opposite part of the probability density function, by construction equivalent to $1 - CL_b$:

$$1 - CL_b = \int_{-\infty}^{q_{obs}} \mathcal{P}_b \tag{2.30}$$

$1 - CL_b$ is called the p-value that can be converted into a significance Z by

$$Z = \Phi^{-1}(1 - p) \tag{2.31}$$

with Φ as the quantile of the standard Gaussian, corresponding to multiples of the standard deviation σ [46]. It is a measure for the probability that the observation is purely caused by a fluctuation of the background. The commonly agreed p-values are 0.27% or 3σ which is called *evidence* and 2.87×10^{-7} or 5σ called an *observation*.

To what confidence level can an observation be explained by the signal plus background hypothesis?

$$CL_{s+b} = \int_{q_{obs}}^{+\infty} \mathcal{P}_{s+b} \tag{2.32}$$

CL_{s+b} represents the fraction of toys having higher q than the observation, thus giving a measure for the confidence in the $(s + b)$ hypothesis. This corresponds to the probability to miss a discovery by falsely assigning it to the b hypothesis. The commonly agreed threshold here is only 95% or 1.64σ, because the a-priori probability to find a signal without a certain prediction is very low. Nevertheless, there is a 5% probability to miss the signal when calculating exclusion limits.

How can one exclude a signal hypothesis?

In the case of largely overlapping probability functions $\mathcal{P}_b$ and $\mathcal{P}_{s+b}$, fluctuations in either of them can cause CL_{s+b} quickly to also fluctuate below 0.05. Fluctuations can never be fully avoided and the method should be robust to them. For that reason, the exclusion judgment is rather done on the ratio between the confidence for the signal plus background and the background confidence level [45].

$$CL_s = \frac{CL_{s+b}}{CL_b} \tag{2.33}$$

CL_s is by definition always larger or equal to CL_{s+b} and therefore a more conservative measure.

How to probe the cross section?

While the previously explained methods have been used at LEP, at LHC a modified likelihood ratio depending on a **signal strengh modifier** $\mu = \sigma/\sigma_{SM}$ was used with the expected Standard Model cross section σ_{SM} [46]. It is introduced by the substitution

$$s_k \to \mu \cdot s_k. \tag{2.34}$$

While by the signal strength modifier is zero for the non-existence of a signal and positive otherwise, one demands that $\mu \leq 0$. Modifying equation 2.25 with μ, the best fit value $\hat{\mu}$ with the corresponding nuisance parameters $\hat{\hat{\theta}}$ while $0 \leq \hat{\mu} < \mu$ to guarantee one-sided confident intervals as needed for upper limits.

$$\begin{aligned} \mathcal{L}_{s+b} &\to \mathcal{L}_{s+b}(\mu, \hat{\theta}) \\ \mathcal{L}_b &\to \mathcal{L}_b(\hat{\mu}, \hat{\hat{\theta}}) \end{aligned} \tag{2.35}$$

we get the negative log-likelihood ratio

$$q_\mu = -2 \ln \lambda(\mu) = \frac{\mathcal{L}\left(data|\mu, \theta(\mu)\right)}{\mathcal{L}\left(data|\hat{\mu}, \hat{\hat{\theta}}\right)} \tag{2.36}$$

quantifying the main parameter of interest, μ. $\hat{\mu}$ and $\hat{\hat{\mu}}$ correspond to the global maximum of the likelihood functions. $\theta(\mu)$ on the other hand maximizes the likelihood for a given parameter of interest μ.

Using this definition, the likelihood ratio becomes a profile likelihood function that can be used in asymptotic χ^2 formulas.

Further, the CL_s value was used to quantify the *exclusion sensitivity* of an analysis, meaning for which signal strength modifier μ the analysis is able to give a statement on the absence of the searched particle. The p-value is obtained by requiring $\mu = 0$, hence it gives a statement about the compatibility with the observation in the absence of the signal.

2.4.2 The Higgs boson decay to photons

The decay of a Higgs boson to two photons is very rare with a branching ratio of only 0.23% at $m_{\mathrm{H}} = 125\,\mathrm{GeV}/c^2$. With the total production cross section of 19.3 pb in the gluon-gluon fusion and 1.6 pb in the VBF production mode one expects only about 50 interesting events per integrated luminosity of $1\,\mathrm{fb}^{-1}$. The sensitivity therefore was achieved by the good and refined mass resolution in the reconstruction of photons combined with a high signal acceptance and reconstruction efficiency.

The leading photon had the requirement of having at least $p_{\mathrm{T}} > 33\,\mathrm{GeV}/c$, the trailing one $25\,\mathrm{GeV}/c$. The acceptance was kept to a fiducial region of $|\eta| < 2.5$, excluding the transition region from the barrel to the endcaps. The with time degrading ECAL resolution was precisely monitored and modeled in simulated data by the application of smearing on the energy resolution.

The $Z \to ee$ background could be removed by requiring an anti-electron veto, excluding all events where the ECAL superclusters are matched to a track.

The photon identification to distinguish between photons from jets and prompt photons was performed using a BDT. The BDT was trained on shape information, isolation information, median energy density as a measure for pile-up as well as the η and p_{T} variables of the photons, knowing that the identification is momentum-dependent. In the performance plot for the photon-identification, a slight excess of the data above the simulated backgrounds is visible that is compatible with the simulated Higgs boson events.

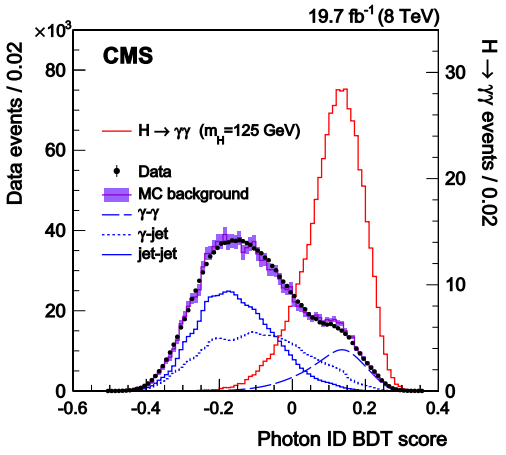


Figure 2.18: Performance of the photon-identification. [47]

For a precise di-photon mass reconstruction the angle between the two photons is important. Unfortunately, one can not assign a track to photons and therefore also no primary vertex can be assigned to a photon. The usual primary vertex definition has the problem that formula 2.3 is lacking the, in this case domi-

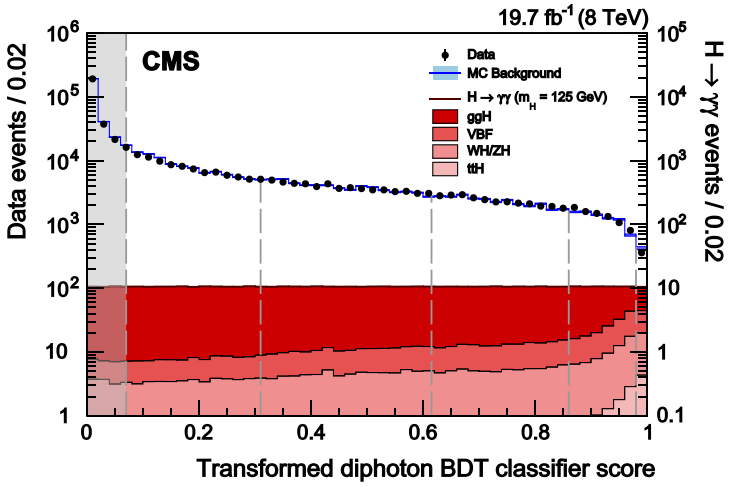


Figure 2.19: Transformed BDT score comparison for simulation and data. The leftmost area in gray is neglected due to its low $(S/(S+B))$ ratio. The dashed lines show the borders between the untagged event classes [47].

nant, electrical neutral components, possibly leading to the wrong primary vertex assignment. With a BDT-based method, combining the usual primary vertex definition with modified ones together and a variable testing the compatibility of the recoil with the identified photons, an independent study found an overall correct assignment of the primary vertex of 80%.

The event classification was purely based on BDTs. What the BDT should not do is to select anything correlated with the mass, because the mass variable is reserved as final discriminating variable. Rather, the classifying BDT should select events having a good mass resolution and high signal probability. The independence of the mass was made sure e.g. by dividing the photon momenta by the reconstructed di-photon mass. Other variables are the vertex information and other kinematic variables. The BDT output has been rescaled such that the expected signal yield is constant over all bins.

Events fulfilling certain tag requirements have been evaluated using special BDTs. The tags are targeting objects from the production process, like the VBF tag by two jets with a large gap in η. Each of these tagged events have been classified, in the same manner as the untagged events, using a BDT while neglecting the lowest class.

The total efficiency times acceptance in the di-photon search was nearly 50%. The combined results can be found in figure 2.20.

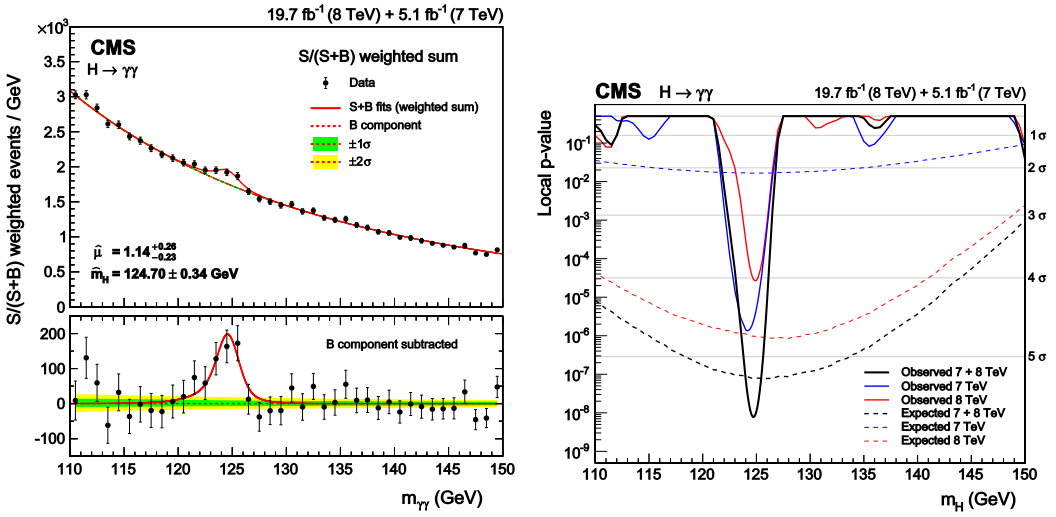


Figure 2.20: Final mass distribution, $s/(s+b)$ weighted *(left)* showing an excess with a best-fit value of $m_{\mathrm{H}} = 124.70\,\mathrm{GeV}/c^2$. *right*: the local p-value for the 7 and 8 TeV analysis and their combination[47].

2.4.3 Higgs to $ZZ \to 4l$

The $H \to ZZ \to 4l$ channel is considered as one of the *golden channels* for several reasons. The clear signature of four leptons in the final state with high transverse momentum has a high identification probability. The four leptons in the final state are reconstructed with high precision of the 4-lepton mass with $1-2\%$ resolution. Also important is that there are few processes in the SM also leading to the same final state, meaning there is few background.

Above $m_{\mathrm{H}} > 200\,\mathrm{GeV}/c^2$ the $H \to ZZ$ channel would have had also a relatively good branching fraction, see figure 1.8a. At at $m_{\mathrm{H}} = 125\,\mathrm{GeV}/c^2$ only about 1 out of 10000 Higgs bosons decay decay via the $H \to ZZ \to 4l$ mode, see figure 1.8b. The analysis therefore relies very much on high reconstruction efficiencies.

The analysis chose pairs of isolated leptons with opposite charge, leading to two Z candidates. Final state radiation photons were only added if bringing the reconstructed Z mass closer to its nominal value. The four-lepton search range of m_{4l} was restricted to be $> 100\,\mathrm{GeV}/c^2$.

The overall reconstruction efficiency was found to be between 62% and 30%. The very low signal event rate does not allow a sophisticated categorization. Instead, a one, two and three-dimensional likelihood function was used to estimate the signal significance (figure 2.21). The analysis expected a significance of 6.7σ and observed the Higgs boson with 6.8σ [48].

PHYSICAL REVIEW D 89, 092007 (2014)

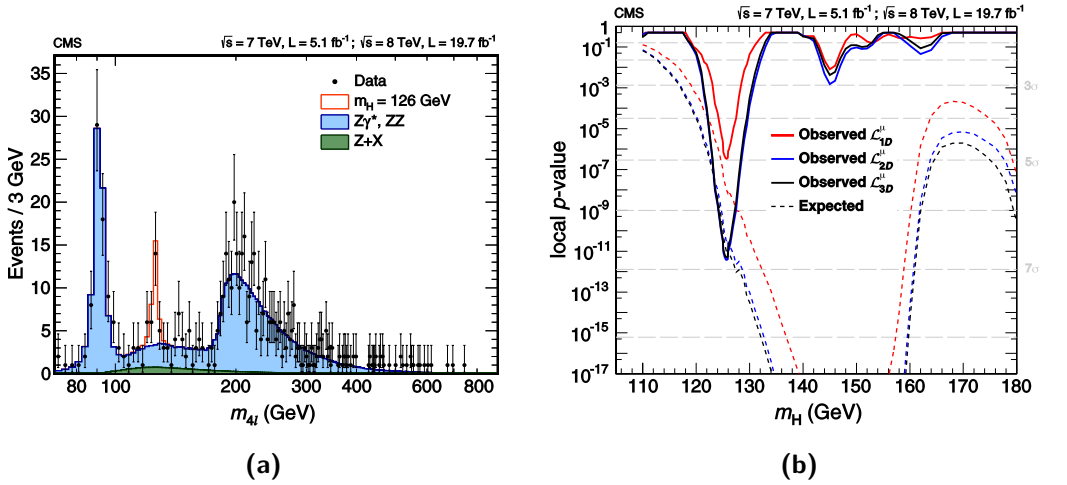


Figure 2.21: Final mass distribution of the $H \to ZZ \to 4l$ analysis (a) and the observed and expected p-values giving the probability of the absence of the Higgs boson(b) [48].

2.4.4 The Higgs boson decay to W bosons

The decay channel $H \to WW \to 2l2\nu$ is very different from the $H \to \gamma\gamma$ and $H \to ZZ$ channels. The energy carried by the two neutrinos makes it impossible to fully reconstruct m_{H} and comes with a mass resolution of only $O(20\%)$ [47]. The advantage of the $H \to WW$ channel is the relatively high branching ratio of 1/100. This allows the additional exploitation of signatures from the production process like VBF and associated production and also the categorization depending on the final state as well as in terms of jet multiplicities. The background rejection against backgrounds is done by requirements on the di-lepton mass and transverse momentum as well as the transverse mass m_{T}, defined as

$$m_{\mathrm{T}} = \sqrt{2p_{\mathrm{T}}^{\mathrm{LL}} \not{E}_{\mathrm{T}} \left(1 - \cos\phi(ll, \vec{\not{E}}_{\mathrm{T}})\right)} \tag{2.37}$$

A special $Z + jet$ veto is applied when a jet is compatible with the di-lepton system. The assumption of a spin 0 leads to a preference of low ϕ_{ll} compared to non-resonant WW production. The final discriminating variable is a two-dimensional template of 9 bins in m_{ll} and 14 bins in m_T, leading to an expected significance of 5.4σ, while 4.7σ have been observed. An alternative is a counting experiment, observing 4.3σ. The spin-parity hypothesis $J^P = 0^+$ is favored. Both expectations are shown in figure 2.23, the final fit for the signal strength modifier μ is displayed to the right.

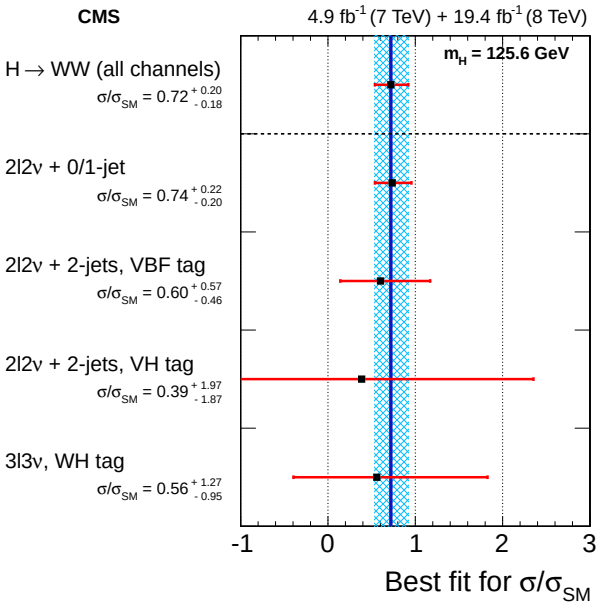
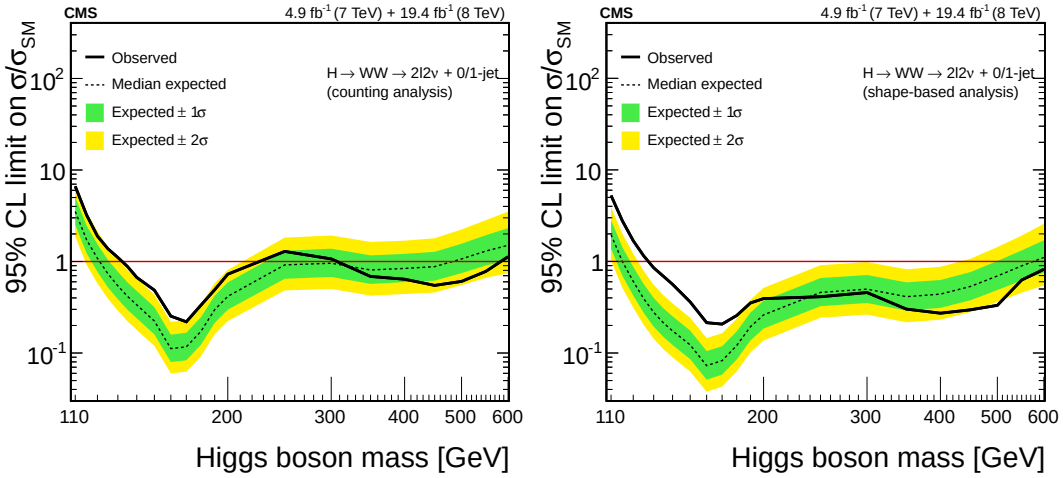


Figure 2.22: Observed σ/σ_{SM} for m_{H} =125.6 GeV/c^2 mass hypothesis [49].



(a) 95% CL level from the counting analysis. **(b)** 95% CL level from the shape analysis.

Figure 2.23: Expected and observed 95% CL upper limits on the $H \to WW$ production cross section relative to the Standard Model expectation as a function of the Higgs boson mass hypothesis m_{H}. (a) is obtained from the counting experiment, (b) from the shape analysis. The deviation from the background-only hypothesis is in a broad mass range way above $m_{\mathrm{H}} = 125\,\mathrm{GeV}/c^2$ [49].

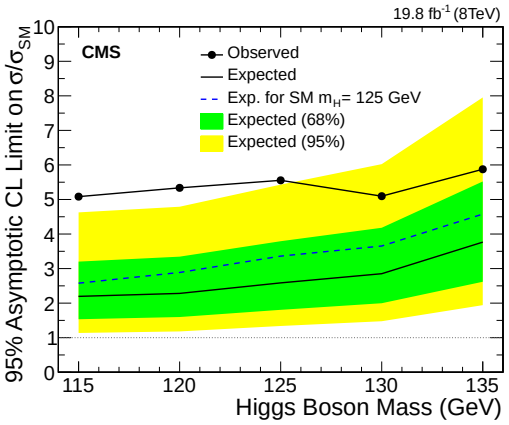


Figure 2.25: Observed and expected 95% confidence level limits for the VBF $H \to b\bar{b}$ analysis as a function of m_{H} [51].

2.4.5 The Higgs boson decay to b-quarks

Already because the majority of Higgs bosons with $m_{\mathrm{H}} = 125\,\mathrm{GeV}/c^2$ decay to pairs of b-quarks, the $H \to b\bar{b}$ channel is worth analyzing. The challenging situation in the $H \to b\bar{b}$ analysis is the two b decaying into two jets, a final state shared with the QCD multijet production.

The associated production mode with W or Z is therefore a possibility, using the electron, muon and tau triggers. The $Z \to \nu\nu$ decay can cause the $\not{E}_{\mathrm{T}}$ trigger to fire and was used since even having a low resolution, the larger branching fraction compensates the disadvantages. Depending on the decay mode and transverse momentum of the associated vector boson, 6 decay channels and 14 exclusive event categories have been defined. A BDT further dividing these categories depending on their signal contribution was applied, splitting some of the categories further up. The final discriminating variable was a BDT, trained to separate the production process VH from all production processes [50]. The significance of the analysis was a lot smaller than the ones of the bosonic channels, expecting and observing both a significance of 2.1σ for a Higgs boson of $m_{\mathrm{H}} = 125\,\mathrm{GeV}/c^2$.

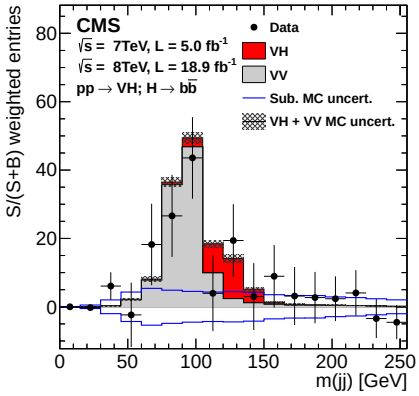


Figure 2.24: Observed σ/σ_{SM} for m_{H} =125.6 GeV/c^2 mass hypothesis [50].

In case the Higgs boson originates from VBF, the production signature adds another two jets. There are two possibilities to trigger these events: the first one is the general-purpose VBF trigger. This trigger fires, having registered two jets with $p_{\mathrm{T}} > 30\,\mathrm{GeV}/c$, a different sign in η, a di-jet mass $m_{jj} > 700\,\mathrm{GeV}/c$ and a difference

in $\Delta\eta_{jj} > 3.5$. The other possibility is a three-jet trigger with high p_{T} thresholds for the three jets of 64-68 GeV/c, 44-48 GeV/c and 24-32 GeV/c, depending on the instantaneous luminosity. All events are classified in two classes with the first one having stricter requirements on all jets from the $qqH \to qqb\bar{b}$ process and the second one having stricter requirements in the VBF topology. Controlled by a Z boson analysis, the Higgs boson signal extraction is performed by a multivariate method, dividing the di-b-jet mass m_{bb} into seven categories. m_{bb} is corrected by a regression technique, improving the mass resolution. The overwhelming QCD background however causes very low signal yields $s/(s+b)$ causing the VBF $H \to b\bar{b}$ analysis to not be sensitive for the Higgs boson Standard Model cross section at the Run I dataset[51].

2.4.6 The Higgs boson decay to tau-leptons

The channel $H \to \tau\tau$ is the most promising channel for a direct measurement of the coupling of leptons to the Higgs boson. The branching ratio of 6.3% at $m_{\mathrm{H}} = 125\,\mathrm{GeV}/c^2$ is the second highest of all branching ratios, leading to an expectation of 35000 $H \to \tau\tau$ decays in Run I[39].

The Run I analysis covered all six possible final states noted as *decay channels*, summarized in table 2.2. In all channels, 3^{rd} lepton vetoes are applied and jets matching identified electrons are removed. This allows the addition of special categories, targeting the tag from production processes. These are

- **VBF Tag**: Two jets with a large difference in η and a high di-jet mass and no hadronic activity in between are called the VBF-tagged regions. In the final states $\mu\tau_h$, $e\tau_h$ and $e\mu$ they are additionally divided into a more tight and more loose selection.
- **Associated production with a Z**: Theses analyses make use of events where the Z decays into pairs of electrons or muons, the final state is then called $ll + LL'$. To prevent overlap with the high-resolution $H \to ZZ \to 4l$ analysis, the leptonic same-flavor $H \to \tau\tau$ final states do not have a associated production category.
- **Associated production with a W**: The associated production with a W is performed in case the W decays to $e\nu^e$ or $\mu\nu^\mu$ in the channels listed in table 2.2.

channel	discriminator ggH + qqH	discriminator ass. Z	discriminator ass. W	Sub-dominant background	branching ratio
$\tau_h\tau_h$	$m_{\tau\tau}^{SVFit}$	$m_{\tau\tau}^{SVFit}$	$m_{\tau\tau}^{vis}$	QCD	42%
$e\mu$	$m_{\tau\tau}^{SVFit}$	$m_{\tau\tau}^{SVFit}$	$m_{\tau\tau}^{vis}$	$t\bar{t}$	3.1%
$e\tau_h$	$m_{\tau\tau}^{SVFit}$	$m_{\tau\tau}^{SVFit}$	$m_{\tau\tau}^{vis}$	W + Jets	11.5%
$\mu\tau_h$	$m_{\tau\tau}^{SVFit}$	$m_{\tau\tau}^{SVFit}$	$m_{\tau\tau}^{vis}$	W + Jets	11.3%
ee	BDT	x	x	$Z \to ee$	3.2%
$\mu\mu$	BDT	x	x	$Z \to \mu\mu$	3.0%

Table 2.2: Summary of the decay channels in the $H \to \tau\tau$ analysis. The dominant background for qqH and qqH is $Z \to \tau\tau$, while for VH di-Boson production is the largest irreducible one.

The lepton identified is described in section 2.3. For the rejection of non-prompt or misidentified leptons, the relative lepton isolation R^L is defined as

$$R^L = \frac{I^L}{p_{\mathrm{T}}^{\mathrm{L}}} \tag{2.38}$$

with the absolute isolation I^L from equation 2.15 and the lepton transverse momentum $p_{\mathrm{T}}^{\mathrm{L}}$, being valid for electrons, muons and taus, while for each of them the PF candidates assigned to the lepton are removed from the isolation sums.

For the reduction of the sub-dominant backgrounds events fulfilling the following criteria are rejected

- $\mu\tau_h$ and $e\tau_h$: Transverse lepton mass

$$m_{\mathrm{T}} < 30\,\mathrm{GeV}/c^2 \tag{2.39}$$

  with the transverse momentum of the lepton ${p_{\mathrm{T}}}^l$ and difference $\Delta\phi = \phi^l - \phi\not{E}_{\mathrm{T}}$.

- $e\mu$: A BDT reduces the $t\bar{t}$ background by combining information from the kinematics of the di-lepton system, the $\not{E}_{\mathrm{T}}$, distance of closest approach between the leptons and the primary very as well as b-tagging information on the reconstructed jets

- $l+\tau_h\tau_h$: A BDT is trained to reduce the QCD, $W+$Jets and Z + jets production based on the lepton and $\not{E}_{\mathrm{T}}$ kinematics.

- $\not{E}_{\mathrm{T}} > 30\,\mathrm{GeV}/c$ for $e\tau_h$

- all channels: If an identified b-tagged jet with a transverse momentum $p_{\mathrm{T}} > 20\,\mathrm{GeV}/c$ is present, the event is rejected in order to suppress the $t\bar{t}$ background

The final discriminating variables: $m_{\tau\tau}^{vis}$, $m_{\tau\tau}^{SVFit}$ and a multivariate classifier

The best separating variable for the signal extraction turned out to be the reconstructed mass. The $H \to \tau\tau$ analysis uses different ways to make use of this information. The main criteria for the decision is the energy carried away by neutrinos.

In the analysis targeting associated production with a W boson, its decay is itself a source of neutrinos and therefore missing energy. This makes it impossible to reconstruct the momentum carried away by the neutrinos in the di-tau system coming from the Higgs boson decay. Therefore, the associated production with W channels use the only neutrino-independent information: the visible mass $m_{\tau\tau}^{vis}$ defined as the mass of the four-vector sum of the reconstructed di-tau system.

All other analyses use a more sophisticated mass reconstruction method, called SVFIT. This algorithm guesses unconstrained parameters by evaluating possible values using a likelihood approach.

A hadronic tau decay is fully specified by

- The boost to its rest frame (3 parameters)
- The polar and azimuthal angles of the visible decay products (2 parameters)
- The invariant mass of the τ_h decay products (1 parameter)

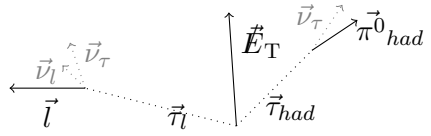


Figure 2.26: Sketch of the measurable components (solid black: visible decay products, $\not{E}_{\mathrm{T}}$) and the ones to be reconstructed (dashed and gray).

The leptonic tau decay has two neutrinos in its final state. This does not add another three parameters because we are not interested in the full reconstruction of them. It adds only one parameter $m_{\nu\nu}$ that is assumed to be 0 in the case of a hadronically decaying tau. What is experimentally accessible is the four-vector of the visible τ decay products in the laboratory frame, leaving two to three parameters unconstrained. The two components of the $\vec{\not{E}}_{\mathrm{T}}$ add another two constraints. The $\not{E}_{\mathrm{T}}$ resolution is worse than the lepton resolution but with an event-by-event covariance matrix estimation. For more details see 3.

The unconstrained parameters are then chosen to be

- The fraction of energy carried by the visible decay products x
- The azimuthal angle of the tau lepton direction in the laboratory frame
- In case of a leptonically decaying tau: the two-neutrino mass $m_{\nu\nu}$

denoted as the decay parameters $\vec{a_1} = (x_1, \phi_1, m_{\nu\nu,1})$.

A likelihood function $f(\vec{\not{E}}_{\mathrm{T}}, \vec{y}, \vec{a_1}, \vec{a_2})$ with the measured four-momenta $\vec{y} = (p_1^{vis}, p_2^{vis})$.

The probability for a mass hypothesis $m_{\tau\tau}^i$ is

$$P(m_{\tau\tau}^i) = \int \delta\left(m_{\tau\tau}^i - m_{\tau\tau}(\vec{y}, \vec{a_1}, \vec{a_2})\right) f(\vec{\not{E}}_{\mathrm{T}}, \vec{y}, \vec{a_1}, \vec{a_2}) \mathrm{d}\vec{a_1} \mathrm{d}\vec{a_2} \tag{2.40}$$

consisting of basically three parts. Two give the likelihood for the decay parameters, the third one quantifies the compatibility of the hypothesis with the measured $\vec{\not{E}}_{\mathrm{T}}$. The likelihood functions for the decay parameters differ whether the decay is leptonically or hadronically. The Likelihood function for the leptonic tau decay is

$$\mathcal{L}_{\tau\to l} = \frac{\mathrm{d}\Gamma}{\mathrm{d}x\mathrm{d}m_{\nu\nu}\mathrm{d}}\phi \propto \frac{m_{\nu\nu}}{4m_\tau^2}\left[(m_\tau^2+2m_{\nu\nu}^2)(m_\tau^2-m_{\nu\nu}^2)\right]\mathcal{L}_{\tau_h} \quad \frac{\mathrm{d}\Gamma}{\mathrm{d}\,x\mathrm{d}\phi} \propto \frac{1}{1-(m_{\tau\tau}^{vis})^2/m_\tau^2} \tag{2.41}$$

x is constrained, in case of the leptonic decay, to $0 \leq x \leq 1$. The two-neutrino mass constraint is $0 \leq m_{\nu\nu} \leq m_\tau\sqrt{1-x}$. In the hadronic case, all visible decay products are treated as a single system and the constraint on x becomes

$$\frac{(m_{\tau\tau}^{vis})^2}{m_\tau^2} \leq x \leq 1 \tag{2.42}$$

It is assumed that neutrinos from tau decays are the only source of missing transverse energy. Therefore, based on the decay parameters $\vec{a_1}$ and $\vec{a_2}$, the likelihood as a function of the $\vec{I\!\!E}_{\mathrm{T}}$ is

$$\mathcal{L}_\nu(\vec{I\!\!E}_{\mathrm{T}}) = \frac{1}{2\pi\sqrt{|V|}}exp\left(-\frac{1}{2}\left(\begin{pmatrix}I\!\!E_x - \sum p_x^\nu\\ I\!\!E_y - \sum p_y^\nu\end{pmatrix}\right)V^{-1}\left(\begin{pmatrix}I\!\!E_x - \sum p_x^\nu\\ I\!\!E_y - \sum p_y^\nu\end{pmatrix}\right)\right) \tag{2.43}$$

where the covariance matrix V is estimated on an event-by-event basis by the $I\!\!E_{\mathrm{T}}^{\mathrm{MVA}}$algorithm.

In the same-flavor leptonic channels ee and $\mu\mu$, the dominant background is the $Z \to ee$ respectively the $Z \to ll$ background. With four neutrinos in the final state, the mass resolution degrades so much that using $m_{\tau\tau}^{SVFit}$ as final discriminating variable is not sufficient to extract any significant contribution in the combination. This already becomes obvious in the comparison of the expected $s/(s+b)$ ratio across channels. Where in $\mu\tau_h$, $e\tau_h$ and $e\mu$ it is over 0.3 to 0.5 in the VBF categories, even the VBF category has $O(10^{-6})$. This is why a two-stage BDT approach in the $H \to \tau\tau \to \mu\mu$ channel has been developed[41]. A multivariate approach based on two specialized BDTs is used. They are both based on kinematic variables of the di-lepton system, the $I\!\!E_{\mathrm{T}}$, the distance of closest approach between the leptons and in the 2-jet categories m_{jj} and $\Delta\eta_{jj}$. The first BDT was trained to discriminate prompt from non-prompt muon decays, separating $Z \to ee/\ Z \to ll$ from the $Z \to \tau\tau$ background. The second BDT was targeting the different production processes, separating $H \to \tau\tau$ from $Z \to \tau\tau$. After several studies it was found that a combined training for 0- and 1-jet and an extra training for 2-jet with the 2-jet variables available leads to the best results.

The combination of the two BDT classifiers B_1 and B_2 is done using the two-dimensional joint probability f_{sig}

$$D = \int_{-\inf}^{B_1}\int_{-\inf}^{B_2} f_{sig}\left(B_1^{'},B_2^{'}\right)\mathrm{d}B_1^{'}\mathrm{d}B_2^{'} \tag{2.44}$$

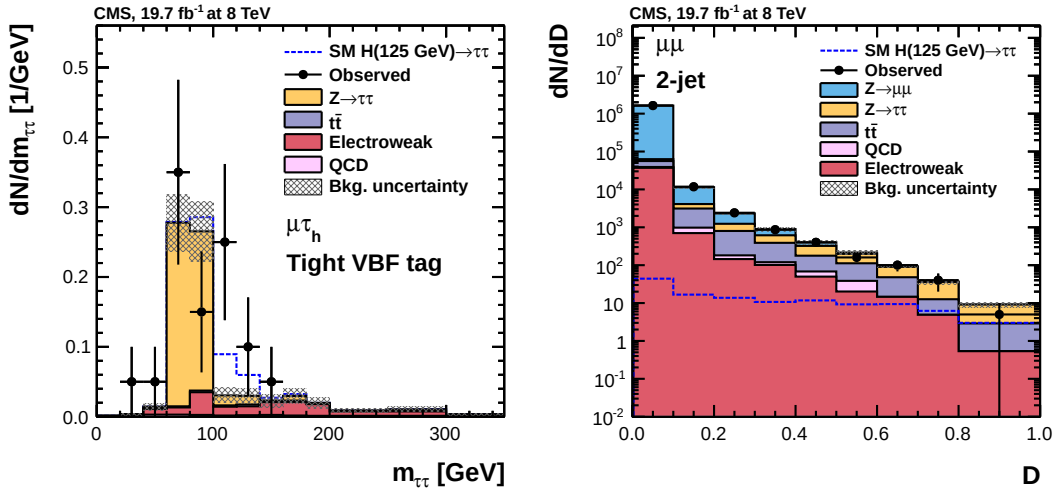


Figure 2.27: Comparison of final discriminating variables. Left: $m_{\tau\tau}^{SVFit}$ in the VBF tight category of the $\mu\tau_h$ channel. 20 events are observed in this category with a signal expectation of 2.4 events. Right: The final discriminating variable D in the $H \to \tau\tau \to \mu\mu$ analysis in the VBF category. The prompt $Z \to \mu\mu$ decays get classified with a low score. The relative signal contribution increases from $O(10^{-5})$ at low values of D to 9 expected background in 2 expected signal events in the high-D region.

Two exemplary distributions of final discriminating variables can be found in figure 2.27.

Categorization

The reconstructed and pre-selected events are categorized in mutually exclusive event categories with the aim to maximize sensitivity in the search range of m_{H} from $110\,\mathrm{GeV}/c^2$ to $145\,\mathrm{GeV}/c^2$.

The number of reconstructed jets is defines the first set of categories. The number of jets with a transverse momentum of $p_{\mathrm{T}} > 30\,\mathrm{GeV}/c$, $|\eta| < 4.7$ and a distance in ΔR between the jet and any identified leptons of 0.5.

The *VBF tag* makes use of the VBF production process where the jets are separated by a large gap in η. If there are jets from the $Z \to \tau\tau$ they are caused by initial state radiation and therefore more central. Apart from the requirement of a large $\Delta\eta_{jj}$, the invariant di-jet mass m_{jj} is large for real VBF events. If there is an additional jet between the two tagging ones, events loose their VBF tag and end up in the 1- or 0-jet categories.

The transverse momentum of the reconstructed Higgs boson is the sum of the transverse components of the components assigned to the hard scattering:

$$p_{\mathrm{T}}{}^{\tau\tau} = |\vec{p_{\mathrm{T}}}^{\,L} + \vec{p_{\mathrm{T}}}^{\,L'} + \vec{\not{E}}_{\mathrm{T}}| \tag{2.45}$$

A large transverse momentum of the Higgs boson improves mass resolution to achieve better separation against the $Z \to \tau\tau$ background and for the reduction of the QCD contribution.

The 0- and 1-jet categories are also split up in terms of the transverse momentum of the hadronically decay tau in channels where is one, the muon momentum in $e\mu$ and the leading lepton in ee and $\mu\mu$. This is motivated by the fact that because of the higher mass of the Higgs boson $m_{\mathrm{H}} > m_{\mathrm{Z}}$ the leptons also get a higher boost.

The full categorization is summarized in Figure 2.28

Background estimation

To minimize uncertainties, the $H \to \tau\tau$ analysis uses data-driven background estimation techniques wherever possible.

- The $Z \to \tau\tau$ background is in most channels the largest. The modeling of this background is done in a semi-data-driven way making use of the lepton universality. In each data-taking period, from a loose $Z \to \mu\mu$ selection was done. The reconstructed particle-flow muons have been removed, hence leaving behind the Z boson recoil and the pile-up. The *embedding* then means, that two simulated tau leptons from a $Z \to \tau\tau$ decay are put instead in the event. After that, all steps of the event reconstruction based on particle flow candidates are done, like the jet clustering, lepton isolation and $\not{E}_{\mathrm{T}}$ reconstruction. The normalization is done using the inclusive $Z \to \mu\mu$ event yield, including differences in the event reconstruction and acceptances. The embedding leads to a dramatic improvement in all uncertainties not concerning the hard scatter like the jet energy scale, $\not{E}_{\mathrm{T}}$ and luminosity measurement.

- The $Z \to ll$ yield is normalized by subtracting all known backgrounds from data in each category independently. The W + Jets contribution is normalized in a high-$\mu\tau_h$ sideband region, while using the shape information from simulation. To allow this also in the VBF categories, the isolation requirement is loosened here.

- The $t\bar{t}$ background is normalized by events with at least one b-tagged jet, while keeping the shape from simulation.

- The QCD multijet production has a relatively high cross section. The QCD jets can as well fake hadronic tau decays as well as muons and electrons. The isolation criteria removes many of them, but there is still a contribution left. In the high-statistics categories, the QCD background is estimated by

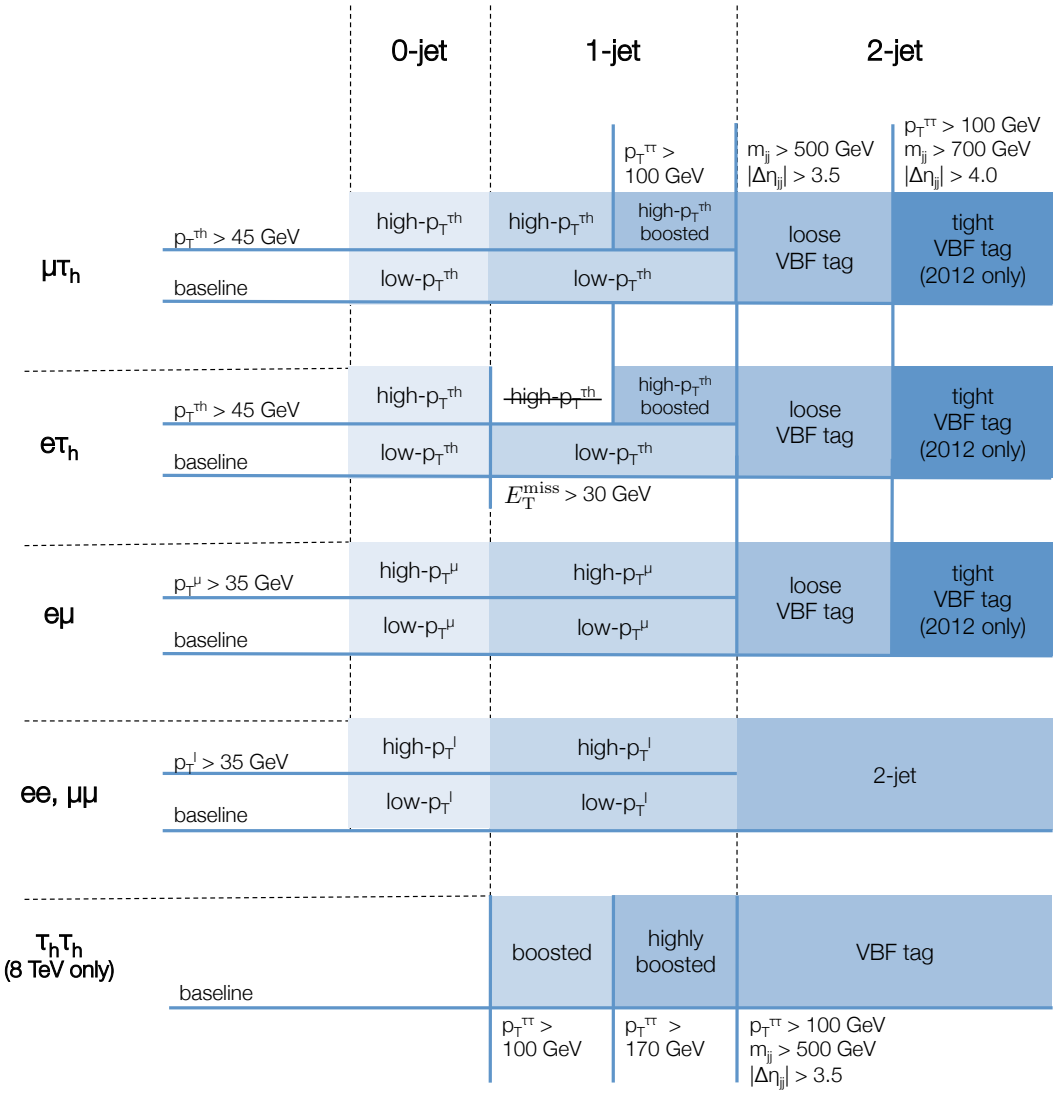


Figure 2.28: Event categories in the ggH and qqH to $\tau\tau$ searches.

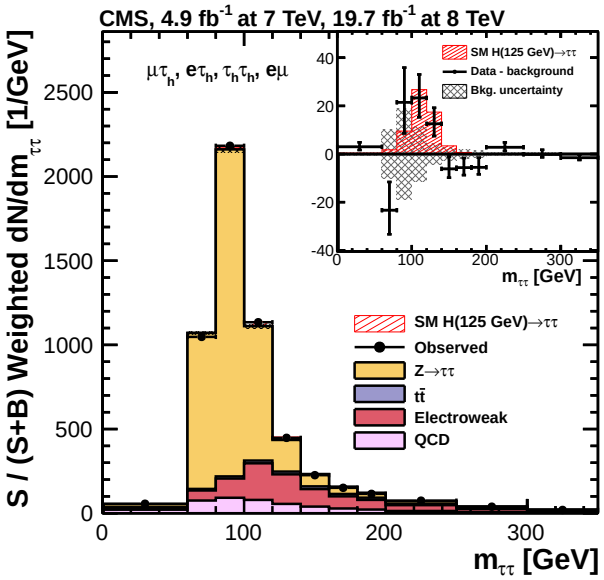


Figure 2.29: $s/(s+b)$ weighted distribution of $m_{\tau\tau}^{SVFit}$ in the $\mu\tau_h$, $e\tau_h$, $e\mu$ and $\tau_h\tau_h$ channels. The background normalization is taken from the global fit, the signal shows the Standard Model expectation. The inset shows the measured data minus the background expectation with the shaded red region as the Standard Model Higgs boson expectation for $m_{\mathrm{H}} = 125\,\mathrm{GeV}/c^2$.

selecting only leptons having the same charge, instead of the usual opposite-sign requirement. Additionally, the isolation criteria are inverted. From this selection, the DY, $\mathrm{t\bar{t}}$ and W + Jets contributions are subtracted. A correction scale factor of 1.06 is applied, measured on a pure QCD multijet sample. The shape is also used for the VBF and 1-jet high-p_{T} categories.

Results

The results are obtained calculating q_μ (see equation2.36) on all final states at the same time, using the final discriminating variables explained above. An $s/(s+b)$ weighted mass distribution can be found in figure 2.29. The overall expectation for the Standard Model Higgs boson is 3.7σ while 3.2σ have been observed. The best-fit value for the signal strength modifier is $\hat{\mu} = 0.78 \pm 0.79$ at $m_{\mathrm{H}} = 125\,\mathrm{GeV}/c^2$. Other masses are shown in figure 2.30.

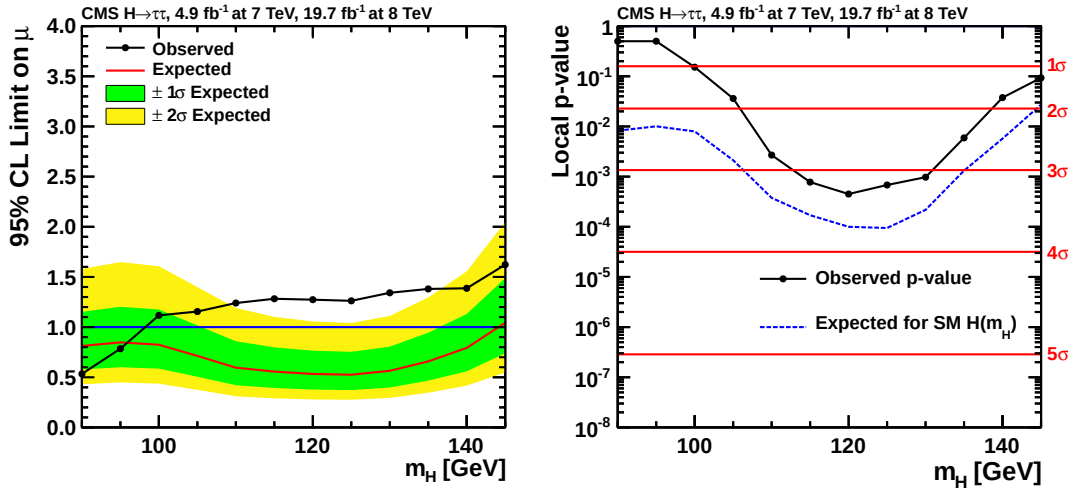


Figure 2.30: Final results of the Run I $H \to \tau\tau$ analysis. Both plots consider $H \to WW$ a background process. Left: The combined expected and observed 95% upper limit on the signal strength parameter μ. The observation shows clearly a deviation from the background-only hypothesis. The broad mass range is caused by the big mass resolution in the $H \to \tau\tau$ channel. Right: Local p-value and significance in multiples of standard deviations σ.

Chapter 3

Reconstruction of the Missing Transverse Energy

3.1 Introduction

The missing energy in the transverse plane $\not{E}_{\mathrm{T}}$ is one of the most complex observables measured at the CMS experiment [52, 53]. The $\vec{\not{E}}_{\mathrm{T}}$ is the negative sum of the momentum vectors of all measured and reconstructed objects $(e, \mu, \tau, \text{jets}, \gamma)$, projected to the transverse detector plane. See section 2.3 for the reconstruction of objects in the CMS experiment and the particle properties.

This quantity is usually given in polar coordinates with the magnitude $\not{E}_{\mathrm{T}}$ and polar angle $\phi_{\not{E}_{\mathrm{T}}}$.

The $\not{E}_{\mathrm{T}}$ serves many purposes. At the LHC, protons collide with negligible initial transverse momentum. Momentum conservation in the transverse plane requires that this is also the case after the collision. If it is not, there can be several reasons to measure a non-negligible $\not{E}_{\mathrm{T}}$. The main one is the production of neutrinos that are invisible to the detector. In addition, several other effects can result in $\vec{\not{E}}_{\mathrm{T}}$:

- Hypothetical new particles in physics beyond the Standard Model: Many searches performed by the CMS collaboration target a multitude of theories that go far beyond the Standard Model. These are for instance searches for black holes, dark matter, supersymmetry or extra dimensions. Dark matter searches, as an example, look for known particles decaying into a weakly interacting massive particle (WIMP), leaving a high $\not{E}_{\mathrm{T}}$.

- Detector misalignment: Even a slight misalignment of the detector with respect to the nominal interaction point causes the $\phi_{\not{E}_{\mathrm{T}}}$to be measurably non-uniform

- Detector performance: The largest contribution to the $\not{E}_{\mathrm{T}}$ apart from objects that due to its nature can not be identified, is the detector resolution. $\not{E}_{\mathrm{T}}$ can also result from subsequent interactions or from particle misidentification and detector malfunction. The latter is the reason why $\not{E}_{\mathrm{T}}$ plays a crucial role in

data quality management. Since all parts of the detector are involved in the calculation, the $\not\!\!E_{\mathrm{T}}$ is sensitive to inefficiencies in any part of the detector.

- Detector coverage: The forward region above $|\eta| > 5.0$ is not covered by the detector, hence an imbalance of particles escaping the detector in this region can cause $\not\!\!E_{\mathrm{T}}$.

3.2 Calculation

The raw $\not\!\!E_{\mathrm{T}}$ ($\not\!\!E_{\mathrm{T}}^{\mathrm{raw}}$) is defined as the negative sum in the transverse momentum of all detected and reconstructed particles, the PF candidates.

$$\not\!\!E_{\mathrm{T}}^{\mathrm{raw}} = -\sum_{\text{PF Cand.}} p_{\mathrm{T}} \tag{3.1}$$

The $\not\!\!E_{\mathrm{T}}^{\mathrm{raw}}$has been calculated using the particle flow candidates and does therefore not contain the calibration of the calorimeters. This calibration is performed on reconstructed jets in form of the jet energy corrections, see section 2.3.3. The jet energy corrections are propagated to the $\not\!\!E_{\mathrm{T}}$ by subtracting the transverse momenta of the uncorrected jets from the event and adding transverse momenta of the corresponding corrected jets.

$$\not\!\!E_{\mathrm{T}}^{\mathrm{PF}} = \not\!\!E_{\mathrm{T}}^{\mathrm{raw}} - \sum_{\text{jets}} (\vec{p}_{\mathrm{T}}^{\,\mathrm{corr}} - \vec{p}_{\mathrm{T}}^{\,\mathrm{uncorr}}) \tag{3.2}$$

For this correction, only jets with $p_{\mathrm{T}} > 15\,\mathrm{GeV}/c$ which deposit less than 90% of their energy in the ECAL are used. In the following the nomenclature is that $\not\!\!E_{\mathrm{T}}^{\mathrm{PF}}$ is the $\not\!\!E_{\mathrm{T}}^{\mathrm{raw}}$with jet energy corrections applied[1].

3.3 Benchmarking

To benchmark the performance of the $\not\!\!E_{\mathrm{T}}$, a setup is chosen where events without genuine $\not\!\!E_{\mathrm{T}}$ dominate, and for which the final decay products can be measured precisely. For that purpose, the decay $Z \to \mu\mu$ is ideal. Due to the absence of neutrinos, the measured $\not\!\!E_{\mathrm{T}}$ is dominated by detector effects.

The events used in the following are simulated $Z \to \mu\mu$ events. The two oppositely charged muons need to pass a relative muon isolation of 0.12. The invariant mass

[1] In the plots shown in this thesis, $\not\!\!E_{\mathrm{T}}$ definitions with jet energy corrections applied are marked with the label *T1*, standing for *type-1* corrected. This is a technical term in CMS, meaning that the jet energy corrections have been applied. The $\not\!\!E_{\mathrm{T}}^{\mathrm{MVA}}$, later to be introduced, is always based on the type-1 corrected $\not\!\!E_{\mathrm{T}}^{\mathrm{PF}}$ and needs no additional jet energy correction.

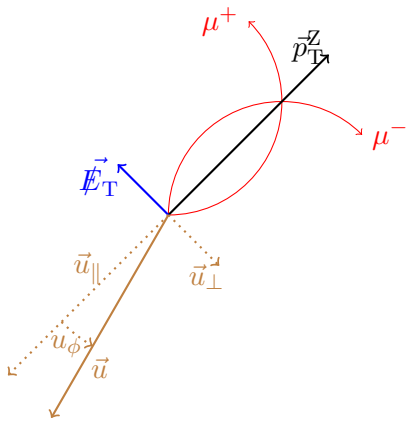


Figure 3.1: Sketch of the recoil $\vec{u}$ in an event with its projection on the di-muon momentum, resulting in $u_{\perp}$ and $u_{\parallel}$. The muons are just included for illustrative reasons, this definition is valid for any process.

of the di-muon system is required to be between 80 GeV/c and 100 GeV/c. Events containing another lepton passing the corresponding isolation criteria are neglected.

The muons in this decay are well understood and can be measured precisely with a momentum relative resolution of less than $\approx 1\%$ (see Figure 2.12). The missing transverse momentum comprises the negative sum of all PF candidates. If the muons are added back, the recoil of the Z/γ^* boson is gained. This recoil consists of initial state radiation and may also contain remnants of the proton. The recoil $\vec{u}$ is defined as

$$\vec{u} = \vec{\not{E}}_{\mathrm{T}} - p_{\mathrm{T}}^{\mathrm{Z}} \tag{3.3}$$

A graphical representation is shown in Figure 3.1. The recoil is a two-dimensional vector and usually expressed in the two components parallel and a longitudinal to the momentum of the di-muon system $u_{\parallel}$ and $u_{\perp}$, respectively.

The studied $Z \to \mu\mu$ process is free of genuine $\not{E}_{\mathrm{T}}$. Thus, we expect the recoil and the di-muon system transverse momentum to be exactly balanced. Therefore, the response $-u_{\parallel}/p_{\mathrm{T}}^{\mathrm{Z}}$ should be one, on average – assuming a well-calibrated detector reconstructing all final state particles. The response is calculated as a function of $p_{\mathrm{T}}^{\mathrm{Z}}$, to estimate its dependence on the scale of the recoil, and as a function of the number of reconstructed primary vertices (n_{PV}) probing the pile-up dependence. The $\not{E}_{\mathrm{T}}$ resolution is measured in the two components $u_{\parallel}$ and $u_{\perp}$. For the $u_{\perp}$ resolution measurement, a Gaussian function is fitted to the distribution to the $u_{\perp}$ in intervals of $p_{\mathrm{T}}^{\mathrm{Z}}$ and n_{PV}. The standard deviation of the Gaussian distribution is interpreted as the resolution. The same procedure is applied on the $u_{\parallel}$ component.

3.4 Alternative definitions of the Missing Transverse Energy

The $\not{E}_{\mathrm{T}}^{\mathrm{PF}}$ is the most holistic $\not{E}_{\mathrm{T}}$ definition. It does not differentiate between the PF candidates in any way – even though more properties than the momentum and energy are known. This section describes further possible definitions of $\not{E}_{\mathrm{T}}$, followed by a discussion of their differences and advantages.

The PF candidates can be classified according to their physical properties. One such property is the electric charge. If a PF candidate has a charge, a corresponding track can be reconstructed. This allows the assignment of an interaction vertex. In most analyses, we are most interested in the main primary vertex following the definition in equation (2.3). However, the majority of all measured PF candidates are not charged, therefore no track can be assigned to them. Clusters of particles are combined to jets (see section 2.3.3). Using the pileup jet id (see section 2.3.3), jets originating from pile-up interactions are identified. Another class of particles are the unclustered neutrals. They usually originate from additional vertices, carry little transverse momentum and have therefore only a small impact on the $\not{E}_{\mathrm{T}}$.

The five classes of PF candidates therefore are:

- Charged particles originating from the main primary vertex (Charged PV)
- Charged particles not originating from the main primary vertex (Charged PU)
- Neutral particles, clustered in jets identified as non-pile-up (Neutral PV)
- Neutral particles, clustered in jets identified as pile-up (Neutral PU)
- Neutral, unclustered particles (Neutral Unclustered)

An additional class of particles arises after the event cleaning step performed by the PUPPI algorithm. It was originally designed to clean the event from particles coming from pile-up interactions and therefore suppress the pile-up dependence of jets. It gives each PF candidate a weight, reflecting the probability of it to originate from pile-up.

The six different classes of PF candidates are combined into six $\not{E}_{\mathrm{T}}$ definitions, each reflecting a different aspect of bias caused by pile-up.

- The $\not{E}_{\mathrm{T}}^{\mathrm{PF}}$ as described above
- **Track** $\not{E}_{\mathrm{T}}$: The idea of the Track $\not{E}_{\mathrm{T}}$ is to consider only charged particles, which usually carry around 50 % of the transverse momentum in an event. In contrast to neutral particles, they can be unambiguously assigned to the leading primary vertex. The measurement of the track also gives them the best energy resolution of all PF candidates.

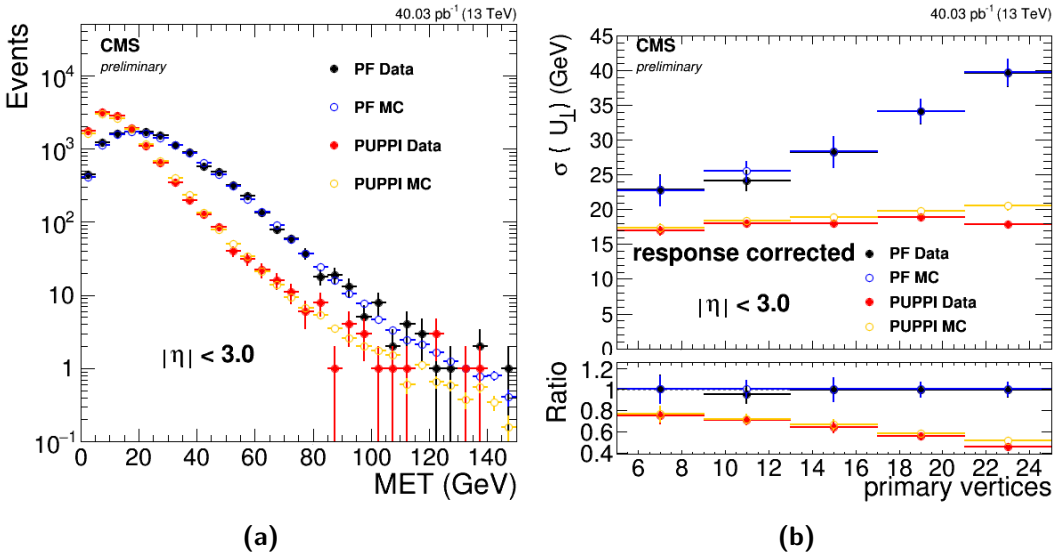


Figure 3.2: Performance of the PUPPI $\not\!\!E_{\mathrm{T}}$ on an integrated luminosity of $40\,\mathrm{pb}^{-1}$ of the first LHC Run II, recorded with 50 ns bunch spacing. The forward region of the ECAL was deactivated in this data taking period, so the $\not\!\!E_{\mathrm{T}}$ was calculated using only PF candidates in the region of $|\eta| < 3$. The plots shown are based on a loose di-muon selection, using the single muon trigger. (a) $\not\!\!E_{\mathrm{T}}$ distribution in logarithmic scale. Data and simulation agree reasonably well for this kind of early comparison. (b) Transverse resolution, measured on both simulation and data over the number of reconstructed primary vertices. The $\not\!\!E_{\mathrm{T}}^{\mathrm{PF}}$ resolution shows a strong dependence on pile-up, degrading by $17\,\mathrm{GeV}/c$ whereas the PUPPI $\not\!\!E_{\mathrm{T}}$ remains nearly unaffected, leading to a resolution gain of almost 50 % at high pile-up. The resolution is corrected for the response, which is 80 % to 90 % for PUPPI $\not\!\!E_{\mathrm{T}}$.

- No-Pileup $\not\!\!E_{\mathrm{T}}$ (**No PU $\not\!\!E_{\mathrm{T}}$**): This definition adds the clustered neutral candidates assigned to the hard scattering to the Track $\not\!\!E_{\mathrm{T}}$.
- **PU Corrected** $\not\!\!E_{\mathrm{T}}$: This definition adds the unclustered neutrals to the No PU $\not\!\!E_{\mathrm{T}}$.
- **PU** $\not\!\!E_{\mathrm{T}}$: This is the complementary definition to the PU Corrected $\not\!\!E_{\mathrm{T}}$, only consisting of PF candidates that either have a track and originating from a pile-up vertex or that are identified as part of a pile-up jet.
- PUPPI $\not\!\!E_{\mathrm{T}}$: PUPPI $\not\!\!E_{\mathrm{T}}$ uses the weighted PF candidates to calculate the $\not\!\!E_{\mathrm{T}}$ in equation (3.1). Special jet energy corrections derived for PUPPI are applied. PUPPI $\not\!\!E_{\mathrm{T}}$ was first presented on the BOOST 2015 conference, see Figure 3.2[54].

An overview on the combination of PF candidate collections and the resulting $\not\!\!E_{\mathrm{T}}$

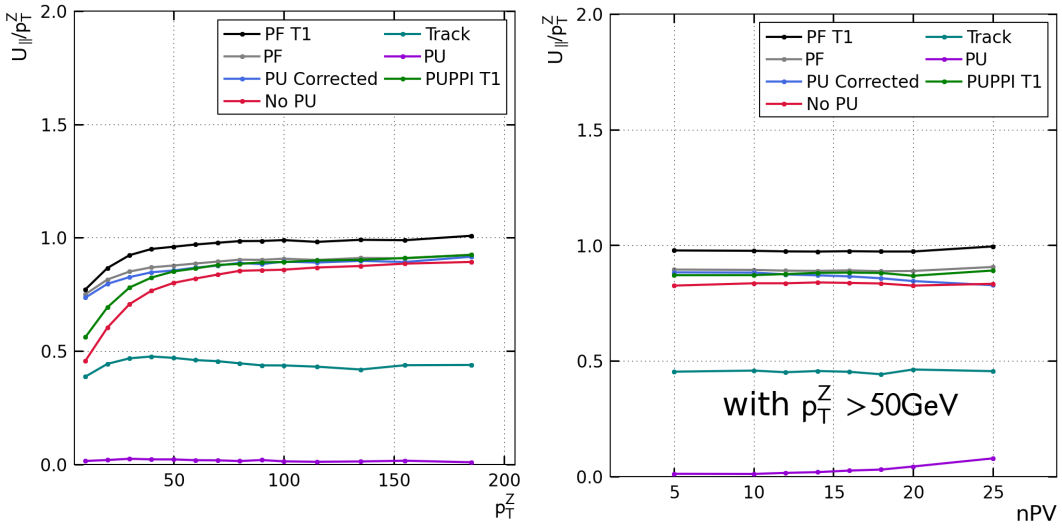


Figure 3.3: Response of the different $\not{E}_T$ definitions. Only the $\not{E}_T^{PF}$ reaches unity response. The uncorrected $\not{E}_T^{raw}$ is included for comparison reasons. Other $\not{E}_T$ definitions have a lower response, the more PF candidates get excluded. The response of the PU $\not{E}_T$ close to 0 means there really is no correlation to the hard scattering. The Track $\not{E}_T$ response is around 50% which is expected as the momentum is carried in average half by charged and half by uncharged particles. Even though the PUPPI $\not{E}_T$ does not reach unity response, which means the PUPPI event cleaning removes more particles than wanted, which causes an imbalance against the recoiling di-muon system.

definitions can be found in table 3.1. The six $\not{E}_T$ definitions all have their advantages and disadvantages. One disadvantage is, apart from the $\not{E}_T^{PF}$, that the response of all definitions is below 1.0. This means they do not include the whole recoil. See figure 3.3 showing the recoil of the different $\not{E}_T$ definitions. The alternative $\not{E}_T$ definitions on the other hand have a better resolution than the $\not{E}_T^{PF}$ and smaller dependence on the pile-up, as demonstrated in figure 3.4. Comparing the resolution of the $\not{E}_T$ definitions on five primary vertices to 25 primary vertices, a degradation in the $\not{E}_T^{PF}$ of 70 % is observed, while the no-pile-up $\not{E}_T$ only decreases by 45 %. PUPPI $\not{E}_T$ has a very good resolution in the perpendicular recoil component, but degrades faster than the other $\not{E}_T$ definitions with higher p_T^Z.

3.5 MVA Missing Energy in the Transverse Plane

The $\not{E}_T$ response and resolution measurements presented above did not identify a $\not{E}_T$ definition that is optimal in all cases.

To exploit the advantages and mitigate the disadvantages of the individual $\not{E}_T$

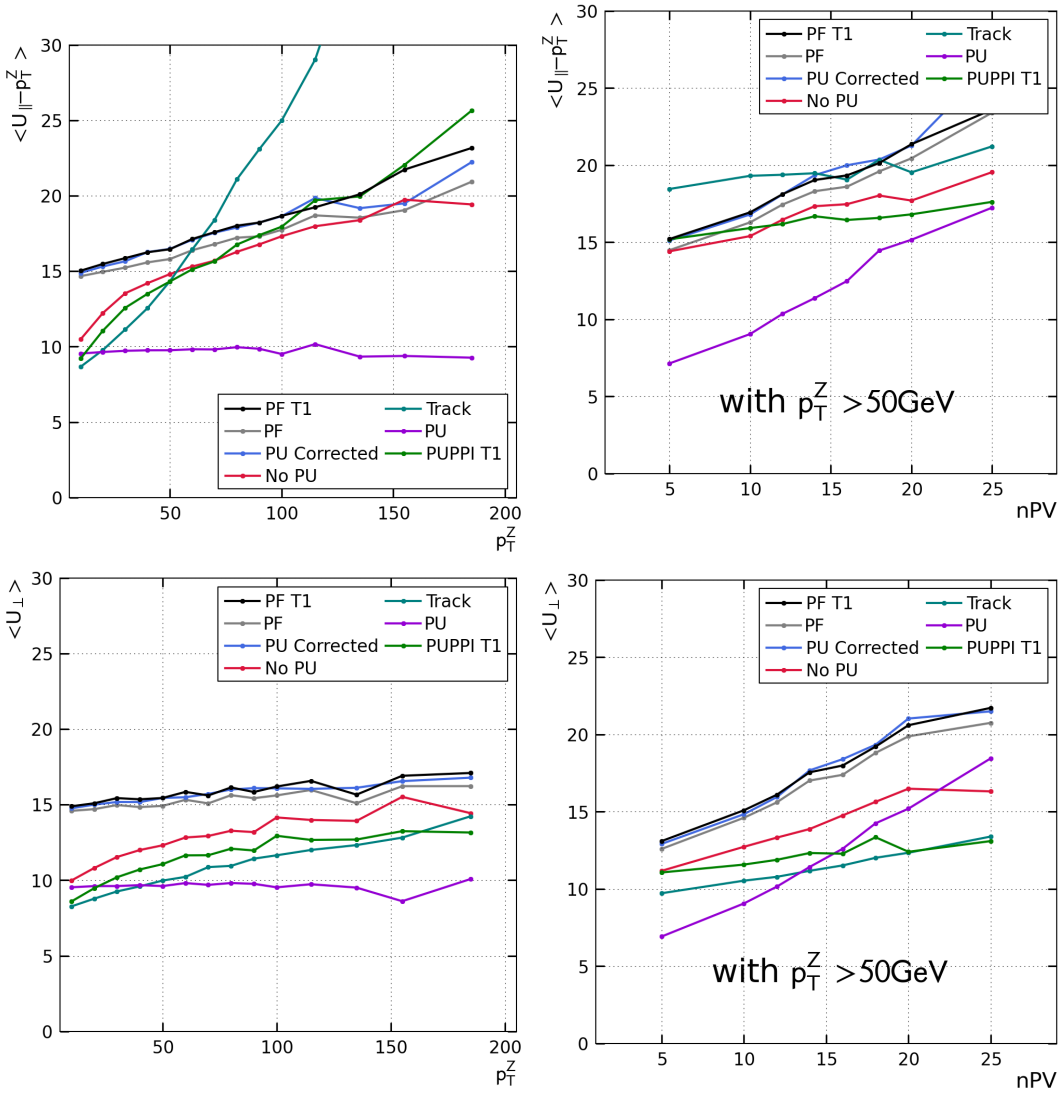


Figure 3.4: Resolution of different $\not{E}_{\mathrm{T}}$ definitions, not corrected by their response. An increase of resolution by the n_{PV} implies a pile-up dependence of the $\not{E}_{\mathrm{T}}$ definition.

Table 3.1: Overview on composition of $\not\!\!E_{\mathrm{T}}$ definitions used for the $\not\!\!E_{\mathrm{T}}^{\mathrm{MVA}}$

	Charged PV	**Charged PU**	**Neutral PV**	**Neutral PU**	**Neutral Unclustered**
PF $\not\!\!E_{\mathrm{T}}$	✓	✓	✓	✓	✓
Track $\not\!\!E_{\mathrm{T}}$	✓				
NoPU $\not\!\!E_{\mathrm{T}}$	✓		✓		
PU Corrected $\not\!\!E_{\mathrm{T}}$	✓		✓		✓
PU $\not\!\!E_{\mathrm{T}}$		✓		✓	
PUPPI $\not\!\!E_{\mathrm{T}}$	(✓)		(✓)		

definitions in an optimal way, a technique called $\not\!\!E_{\mathrm{T}}^{\mathrm{MVA}}$has been developed. It uses a gradient boosted regression tree technique (see section 2.3.5) and is explained in the following. The multivariate training uses many input variables. These are the p_{T}, ϕ and $\sum E_{\mathrm{T}}$ of the six different $\not\!\!E_{\mathrm{T}}$ definitions plus

- The ϕ, η and p_{T} of the two jets with the highest transverse momenta
- The number of jets with $p_{\mathrm{T}} > 30\,\mathrm{GeV}/c$
- The number of reconstructed primary vertices (n_{PV})

The number of jets and n_{PV} carry information about the pile-up regime in one event.

The component of interest is everything that is not related to the hard scattering decay products. For this reason, in the training case the two leading leptons are removed from the event (see equation 4.4.2). The recoil $\vec{u}$ is coming from initial state radiation or the underlying event and is irrespective of the hard scattering decay products. To make a best guess for this recoil $\vec{u}$, it is calibrated using the known and reconstructed $p_{\mathrm{T}}^{\mathrm{Z}}$.

The recoil $\vec{u}$ is a two-dimensional vector, so a recoil correction also needs two components. It is best if they are uncorrelated so that the multivariate regression does not need to take the correlation into account. This already disqualifies calibrating $u_{\parallel}$ and $u_{\perp}$, because they are correlated to the magnitude of the transverse momentum. The method chosen here is to correct the recoil angle u_{ϕ} first and then $u_{\parallel}$.

Angular correction of u_{ϕ}

The regression target T_1 for the first step is

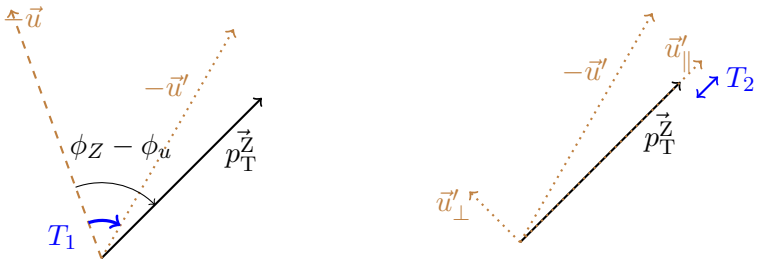


Figure 3.5: Sketch depicting the determination of the angular correction T_1 (left) and scalar correction on the recoil T_2 (right).

$$T_1 = \phi_{\mathrm{Z}} - (u_\phi + \pi) \text{ with } -\pi < T_1 \leq \pi \tag{3.4}$$

The regressed value is called u_ϕ^{MVA}. The addition of π to u_ϕ is helpful for the multivariate regression because it makes the target symmetrical around 0 and the regression algorithm works best on symmetrical input distributions.

Scale correction of $u_\parallel$

The second regression target depends on the angular-corrected recoil. To make it independent of the scale, the recoil is normalized to the recoiling boson momentum. Hence, the second training target T_2 is

$$T_2 = \frac{u_\parallel}{p_{\mathrm{T}}^{\mathrm{Z}}} \tag{3.5}$$

The evaluated value is called $u_\parallel^{\mathrm{MVA}}$, the resulting two-vector $\vec{u}^{\mathrm{MVA}}$.

Event-by-Event resolution estimation

As the dependence on n_{PV} shows, the $\vec{\not{E}}_{\mathrm{T}}$ resolution depends heavily on the event environment. Therefore, it is highly profitable for any fit-based technique using the $\not{E}_{\mathrm{T}}$ as input to get an event-by-event error estimation and not just an average over all events. This information is contained in the $\vec{\not{E}}_{\mathrm{T}}$ covariance matrix, which is defined as

$$\mathrm{V}(\vec{\not{E}}_{\mathrm{T}}) = E\left(\left(\vec{\not{E}}_{\mathrm{T}} - \vec{\not{E}}_{\mathrm{T}}^{\mathrm{True}}\right)\left(\vec{\not{E}}_{\mathrm{T}} - \vec{\not{E}}_{\mathrm{T}}^{\mathrm{True}}\right)^T\right) \tag{3.6}$$

with the true missing transverse momentum $\vec{\not{E}}_{\mathrm{T}}^{\mathrm{True}}$. The true value is only known in simulated data and is expected to be negligible in a $Z \to \mu\mu$ event selection in measured data.

The estimation is now done with two additional multivariate regressions. Since here we assume $u_{\parallel}$ and $u_{\perp}$ to be uncorrelated, the off-diagonal elements are zero. Hence, there are two diagonal elements to determine. This determination is based on the final $\not\!\!E_{\mathrm{T}}^{\mathrm{MVA}}$. On the $Z \to \mu\mu$ simulated dataset, $\vec{\not\!\!E}_{\mathrm{T}}^{\mathrm{True}}$ is known to be zero. For technical reasons, the square-root of the diagonal elements of the covariance matrix are the regression targets. The regression targets T_3 and T_4 are an estimation of the error on the parallel and perpendicular recoil components. Both are normalized to the magnitude of the regressed recoil u^{MVA}.

$$T_3 = \left| \sqrt{\frac{\pi}{2}} \frac{p_{\mathrm{T}}^{Z,\parallel} - u^{\mathrm{MVA}}}{u^{\mathrm{MVA}}} \right| \tag{3.7}$$

$$T_4 = \left| \sqrt{\frac{\pi}{2}} \frac{p_{\mathrm{T}}^{Z,\perp}}{u^{\mathrm{MVA}}} \right| \tag{3.8}$$

T_3 and T_4 depend on the projections of the boson momentum $\vec{p}_{\mathrm{T}}^{\,Z}$ on the recoil $\vec{u}^{\mathrm{MVA}}$ in the two components $p_{\mathrm{T}}^{Z,\parallel}$ and $p_{\mathrm{T}}^{Z,\perp}$. The factor $\sqrt{\frac{\pi}{2}}$ is used to ensure proper normalization later. The full covariance matrix in the coordinate system of $\vec{u}$ is

$$V(\vec{\not\!\!E}_{\mathrm{T}}) = \begin{pmatrix} (T_3)^2 \cdot u & 0 \\ 0 & (T_4)^2 \cdot u \end{pmatrix} \tag{3.9}$$

These two estimations account for the covariance of the recoil. The covariance due to the leptons is handled separately by adding the covariance matrix of the di-lepton system to the $\not\!\!E_{\mathrm{T}}^{\mathrm{MVA}}$covariance matrix as being used in the standard $\not\!\!E_{\mathrm{T}}^{\mathrm{PF}}$. This feature improves the $\not\!\!E_{\mathrm{T}}^{\mathrm{MVA}}$covariance matrix estimation significantly.

p_{T}^{Z} reweighting and mixing

The $\not\!\!E_{\mathrm{T}}^{\mathrm{MVA}}$algorithm aims for an outstanding $\not\!\!E_{\mathrm{T}}$ performance in any CMS analysis. Thus, its performance has to be stable up to high recoils of several hundred GeV/c. The inclusive $Z \to \mu\mu$ sample only contains events with recoils up to around 400 GeV/c. To extend this spectrum, special samples with special generator settings have been used in the training in addition to the inclusive $Z \to \mu\mu$ samples. These special samples require the scalar sum of all parton level particles transverse momenta to be above a certain threshold, called HT. Samples in HT bins of 200 GeV/c to 400 GeV/c, 400 GeV/c to 600 GeV/c and more than 600 GeV/c have been used. This results in enough training statistics to achieve a reasonable coverage of events up to 750 GeV. A strict selection is applied on events from the HT-binned samples, requiring them to have a reconstructed di-muon transverse momentum of at least 70 GeV/c, 90 GeV/c respectively 110 GeV/c. That way, boundary effects are minimized when

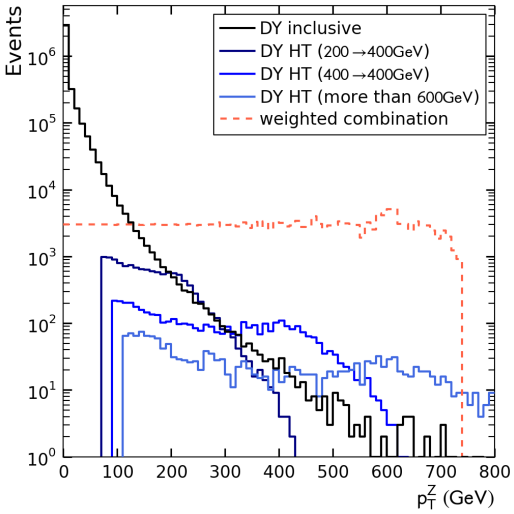


Figure 3.6: p_T distributions for the inclusive and the HT-binned samples. All events get binned equally populated and the events in each bin weighted with the bin width. This results in the weighted combination which gives a nearly flat distribution up to 750 GeV/c. In total, 886.023 events have been used for the training.

one process becomes dominant. Boundary effects lead to sudden jumps in the response and resolution, in case the BDT gets sensitive to the mixing of HT bins.

The mixing strategy is to treat all events equally, regardless of whether they are from the inclusive or special sample. The events are put into equally populated bins with each bin having 1000 entries. Each event is assigned a weight equal to the width of the bin. The last bin above 750 GeV/c is populated by less than 1000 events. Since no meaningful width can be assigned to it, the events in this last bin get dropped. The reweighting does not only allow the mixing, but also to reach a response of 1.0 due to the flat p_T^Z spectrum. Without the reweighting, the falling p_T^Z spectrum always biases the response towards lower values, since these events are more frequent. This effect is completely canceled out with an artificially flat p_T^Z spectrum. This is also the reason why this kind of training is often called the unity response training. The resulting p_T^Z spectrum with the inclusively simulated $Z \to \mu\mu$ sample and the three HT-binned samples is shown in figure 3.6.

A first evaluation is to apply the training on the same $DY \to \mu\mu$ MC event sample and compare the starting point, the $\not{E}_T^{PF}$, the PUPPI $\not{E}_T$ and the new $\not{E}_T^{MVA}$. $\not{E}_T^{PF}$ and PUPPI $\not{E}_T$ are only repetitions of figures 3.3 and 3.4. See figures 3.7 and 3.8 for the results gained on the training dataset.

An important performance indicator is the behavior of the $\not{E}_T$ under increasing pile-up. In figure 3.8, the function $a + b\sqrt{n_{PV}}$ has been fitted to the resolution with

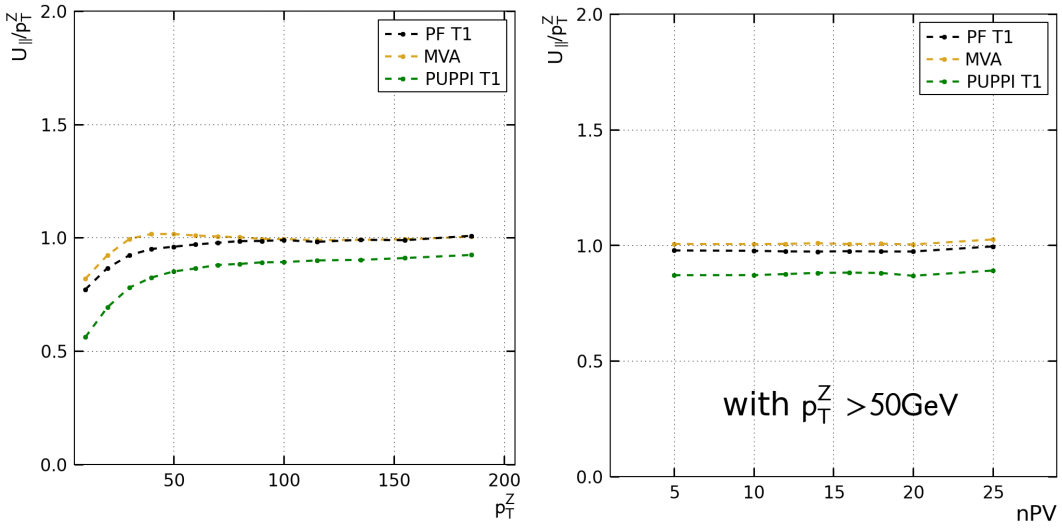


Figure 3.7: The $\not{E}_{\mathrm{T}}^{\mathrm{MVA}}$ response is close to unity above $p_{\mathrm{T}}^{\mathrm{Z}} > 30\,\mathrm{GeV}/c$. Thus, it covers an even broader range than the $\not{E}_{\mathrm{T}}^{\mathrm{PF}}$. Above $50\,\mathrm{GeV}/c$ the response is independent on the number of reconstructed primary vertices in all $\not{E}_{\mathrm{T}}$ definitions.

the fit results shown in table 3.2. The $\not{E}_{\mathrm{T}}^{\mathrm{PF}}$ has a smaller intercept, meaning it has the better resolution in low-pile-up conditions. With rising pile-up however, the $\not{E}_{\mathrm{T}}^{\mathrm{MVA}}$ resolution degrades only half as quick as the $\not{E}_{\mathrm{T}}^{\mathrm{PF}}$ does.

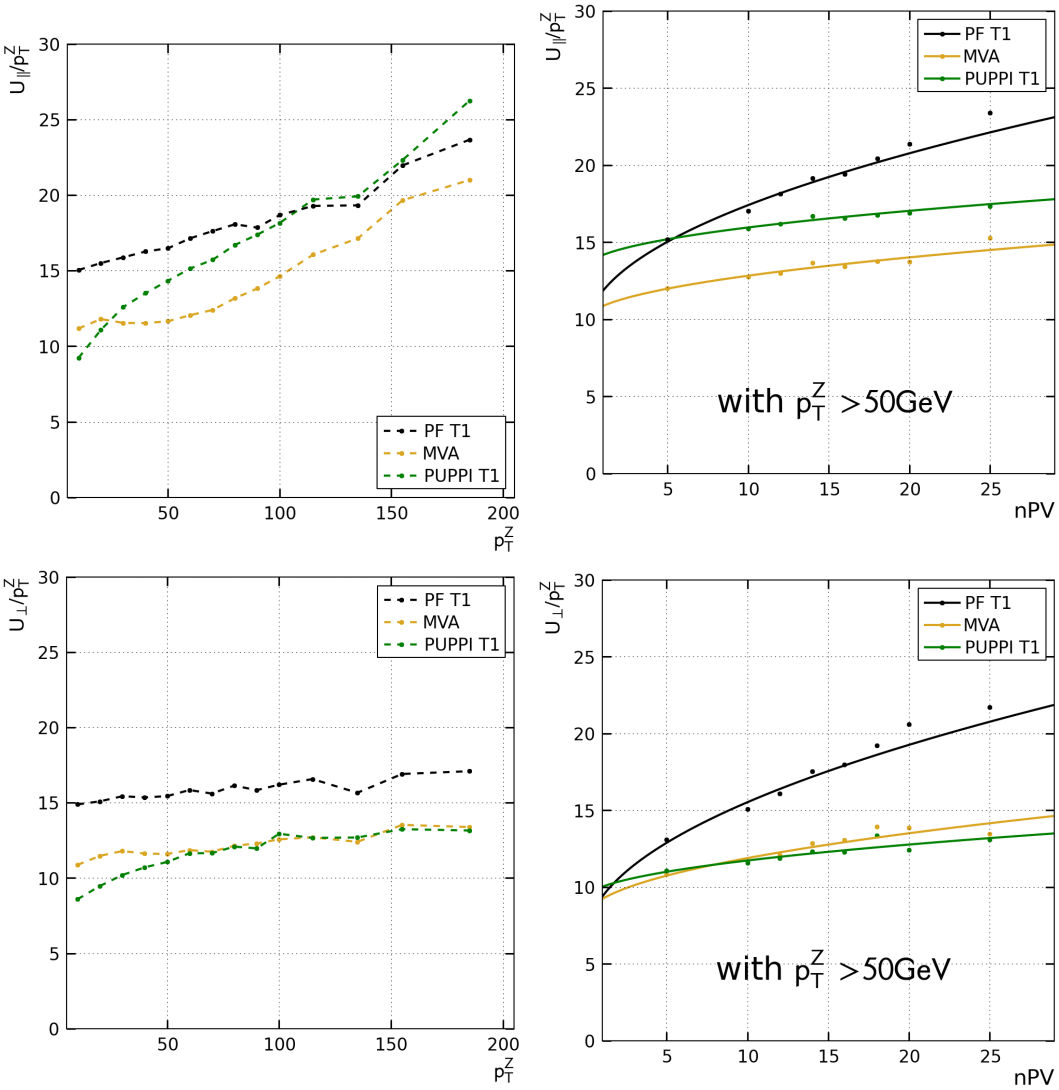
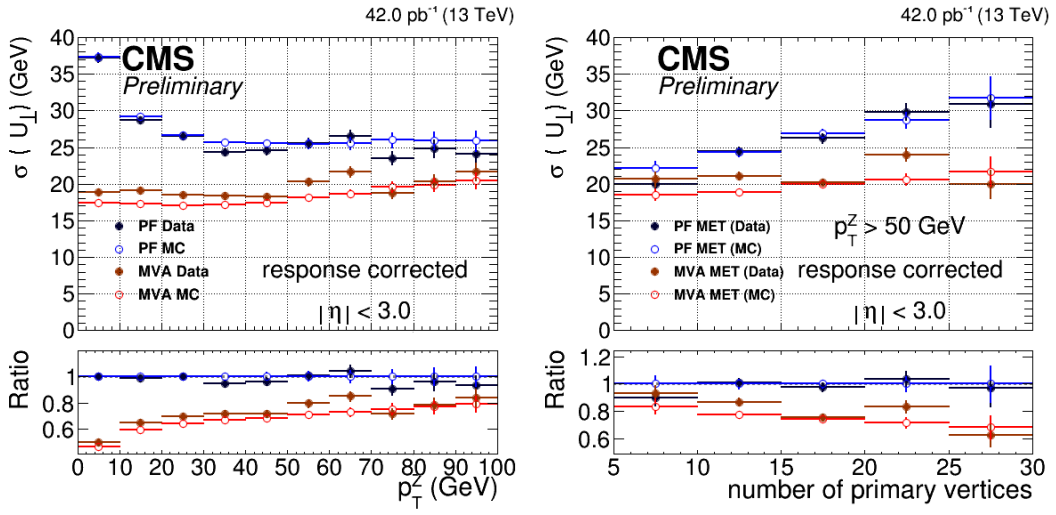


Figure 3.8: Resolution of different $\not{E}_T$ definitions, not corrected by their response. In the p_T^Z dependent resolution measurement, the $\not{E}_T^{MVA}$ has a gain of 20 to 30% in terms of resolution compared to PF $\not{E}_T$. In the n_{PV} dependent plots, the function $a + b\sqrt{n_{PV}}$ has been fitted. See the results in table 3.2.

		intercept a	slope b
$u_{\parallel}$	PF $\not\!\!E_{\mathrm{T}}$	9.3 GeV/c	2.6 GeV/c
	$\not\!\!E_{\mathrm{T}}^{\mathrm{MVA}}$	10.0 GeV/c	0.9 GeV/c
	PUPPI $\not\!\!E_{\mathrm{T}}$	13.4 GeV/c	0.8 GeV/c
$u_{\perp}$	PF $\not\!\!E_{\mathrm{T}}$	6.5 GeV/c	2.8 GeV/c
	$\not\!\!E_{\mathrm{T}}^{\mathrm{MVA}}$	8.0 GeV/c	1.2 GeV/c
	PUPPI $\not\!\!E_{\mathrm{T}}$	9.2 GeV/c	0.8 GeV/c

Table 3.2: Fit results to measure the dependence on the pile-up environment for the different $\not\!\!E_{\mathrm{T}}$ definitions. The $\not\!\!E_{\mathrm{T}}^{\mathrm{PF}}$ is affected by pile-up more than twice as much as PUPPI $\not\!\!E_{\mathrm{T}}$ and $\not\!\!E_{\mathrm{T}}^{\mathrm{MVA}}$.



(a) **(b)**

Figure 3.9: Response-corrected resolution for $\not{E}_{\mathrm{T}}^{\mathrm{PF}}$ and $\not{E}_{\mathrm{T}}^{\mathrm{MVA}}$. This early version of the $\not{E}_{\mathrm{T}}^{\mathrm{MVA}}$ did not use PUPPI $\not{E}_{\mathrm{T}}$ as input. Despite this, it performs similar to the PUPPI $\not{E}_{\mathrm{T}}$ with a resolution gain of 50 % compared to $\not{E}_{\mathrm{T}}^{\mathrm{PF}}$. The advantage is that the response closer to unity compared to PUPPI $\not{E}_{\mathrm{T}}$. The resolution gain for higher recoils becomes smaller, since $\not{E}_{\mathrm{T}}$-altering pile-up effects play a relatively smaller role. The study was performed using a di-muon selection, with the leading muon having $p_{\mathrm{T}} > 25\,\mathrm{GeV}/c$, the trailing one $20\,\mathrm{GeV}/c$ with the single muon trigger being fired. [55]

3.6 $\not{E}_{\mathrm{T}}$ performance

The training for Run II was first presented at the LHCP 2015 conference in St. Petersburg with the performance shown in Figure 3.9. Already for the early data analysis of an integrated luminosity of $42\,\mathrm{pb}^{-1}$, a resolution improvement of 20 % - 50 % has been measured.

In the following, the $\not{E}_{\mathrm{T}}^{\mathrm{MVA}}$ performance is measured on a different processes, having both genuine $\not{E}_{\mathrm{T}}$ from neutrinos as well as on the already presented $Z \to \mu\mu$ decay, which has no neutrinos in the final state. The integrated luminosity is $12.9\,\mathrm{fb}^{-1}$ for the Run II 2016 B - D datasets[2].

[2]The luminosity of $12.9\,\mathrm{fb}^{-1}$ with an uncertainty of 6.2 % has been the recommendation until mid of March 2017. At that point the recommendation was updated to $12.6\,\mathrm{fb}^{-1}$ with 2.5 % uncertainty. The $H \to \tau\tau$ analysis in the next chapter follows this updated recommendation. There is no impact of this change in the evaluation of the $\not{E}_{\mathrm{T}}$ since the normalization does not enter the resolution and response measurements.

3.6.1 Performance in a process without genuine $\not{E}_T$: $Z \to \mu\mu$

This measurement was done on the $Z \to \mu\mu$ selection which is also planned to be used as a control region in the $H \to \tau\tau$ analysis. It uses the single muon trigger and pairs of opposite-sign, isolated muons. The background contains mostly $Z \to \mu\mu$ events, but also includes W + Jets and $t\bar{t}$ events.

The $\not{E}_T^{PF}$ resolution shown in figure 3.10 has been confirmed in an independent measurement[53]. The improvement achieved by using the $\not{E}_T^{MVA}$is of 4-5 GeV/c for a recoil of 50 GeV/c. The more favorable behavior of the $\not{E}_T^{MVA}$compared to the $\not{E}_T^{PF}$ over increasing pile-up from five to 25 reconstructed primary vertices with a degradation of 5-6 GeV/c for $\not{E}_T^{MVA}$and around 9 GeV/c for $\not{E}_T^{PF}$ is confirmed here on 2016 data.

3.6.2 Performance on a process with one neutrino

For this measurement, an event selection according to the analysis chapter in the $H \to \tau\tau \to \mu\tau_h$ channel has been chosen (see chapter 4). The W + Jets process is a background for the $H \to \tau\tau$ search. It can be efficiently suppressed by reconstructing the transverse mass m_T, defined as the sum of the $\not{E}_T$ and a reconstructed lepton transverse momentum vector. This rejection especially profits from a good $\not{E}_T$ resolution. The comparison of the boson decay products and the recoil is in this case unhelpful for the determination of the resolution, since there are neutrinos in the final state. So the two possible checks are, first, how well the W + Jets background is modeled in data and second a comparison of the reconstructed $\not{E}_T$ with the generator $\not{E}_T^{True}$(see figure 3.11).

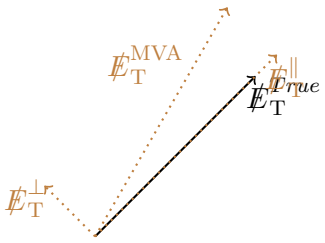


Figure 3.11: Definition of parallel and perpendicular components of the $\not{E}_T^{True}$.

The m_T^{μ} distribution based on the $\not{E}_T^{MVA}$agrees nicely with the observed data, see Figure 3.12.

The resolution improvement, see Figure 3.13, depends on the number of reconstructed primary vertices. While the $\not{E}_T^{PF}$ resolution doubles from 5 to 25 reconstructed vertices, the $\not{E}_T^{MVA}$increases by about 30%. This is not yet very important for the 2016 run but will become more and more crucial when the instantaneous luminosity further increases.

3.6.3 Performance on processes with several neutrinos

The reconstruction of the $\not{E}_T$ in the decay of two tau leptons is of particular importance for this thesis. The reconstruction of the full di-tau system requires the

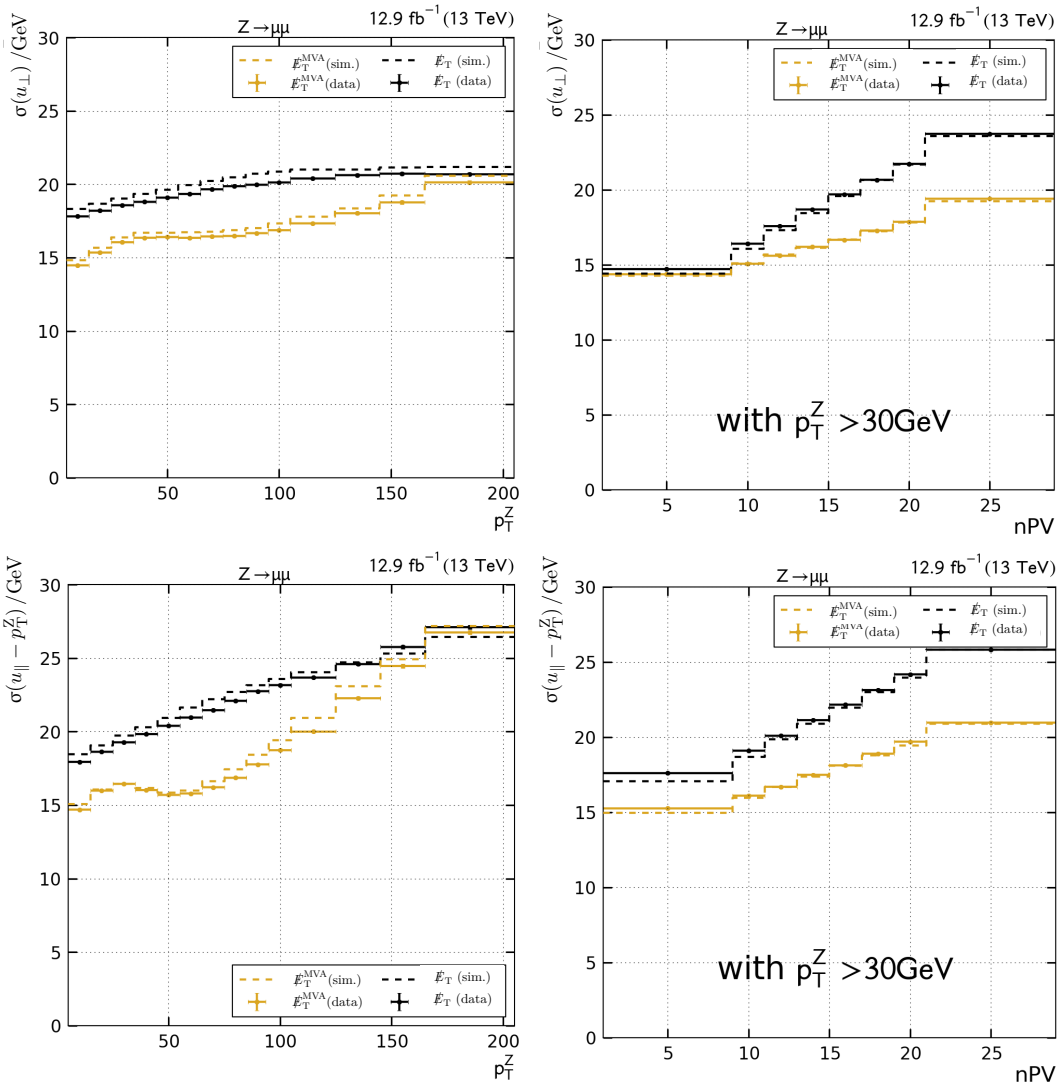


Figure 3.10: Resolution of $\not{E}_{\mathrm{T}}^{\mathrm{PF}}$ and $\not{E}_{\mathrm{T}}^{\mathrm{MVA}}$ for the 2016 dataset with an integrated luminosity of 12.9 fb^{-1}. The multivariate $\not{E}_{\mathrm{T}}$ corrects resolution-degrading pile-up effects. These effects become relatively less important with highly boosted Z/γ^* bosons. Hence, the resolution gain is highest with boosts around 50 GeV/c, where unity response is reached for both $\not{E}_{\mathrm{T}}$ definitions and pile-up effects remain important. At a boson p_{T} above 200 GeV/c, both $\not{E}_{\mathrm{T}}$ definitions are comparable. The smaller slope in $\not{E}_{\mathrm{T}}^{\mathrm{MVA}}$ compared to $\not{E}_{\mathrm{T}}^{\mathrm{PF}}$ is also present here in the n_{PV} distribution on data, showing an increasing advantage of $\not{E}_{\mathrm{T}}^{\mathrm{MVA}}$ with rising pile-up.

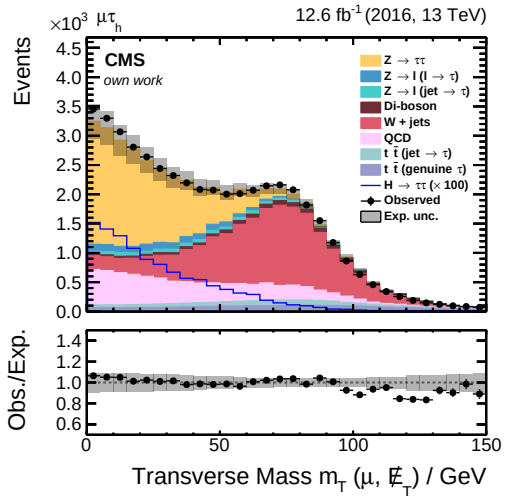


Figure 3.12: m_{T}^{μ} distribution using $\not{E}_{\mathrm{T}}^{\mathrm{MVA}}$ in the $\mu\tau_h$ final state of the $H \to \tau\tau$ analysis. The high- m_{T}^{μ} region is dominated by W + Jets.

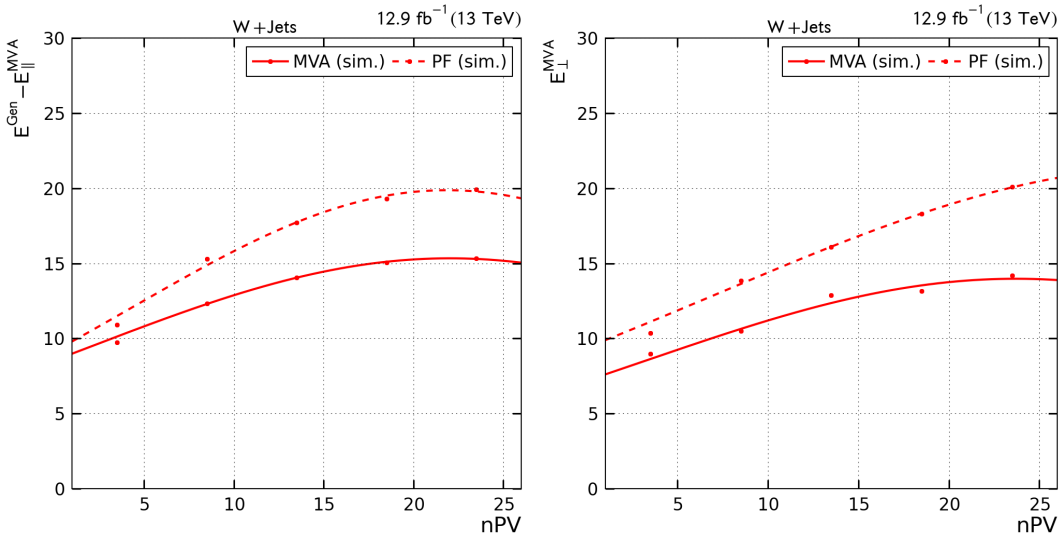


Figure 3.13: Resolution of different $\not{E}_{\mathrm{T}}$ definitions, not corrected by their response. An increase of resolution by the n_{PV} implies a pile-up dependence of the $\not{E}_{\mathrm{T}}$ definition. As already seen in the $\mu\mu$ channel, the resolution in $\not{E}_{\mathrm{T}}^{\mathrm{MVA}}$ degrades less with more activity in the event compared to the $\not{E}_{\mathrm{T}}^{\mathrm{PF}}$.

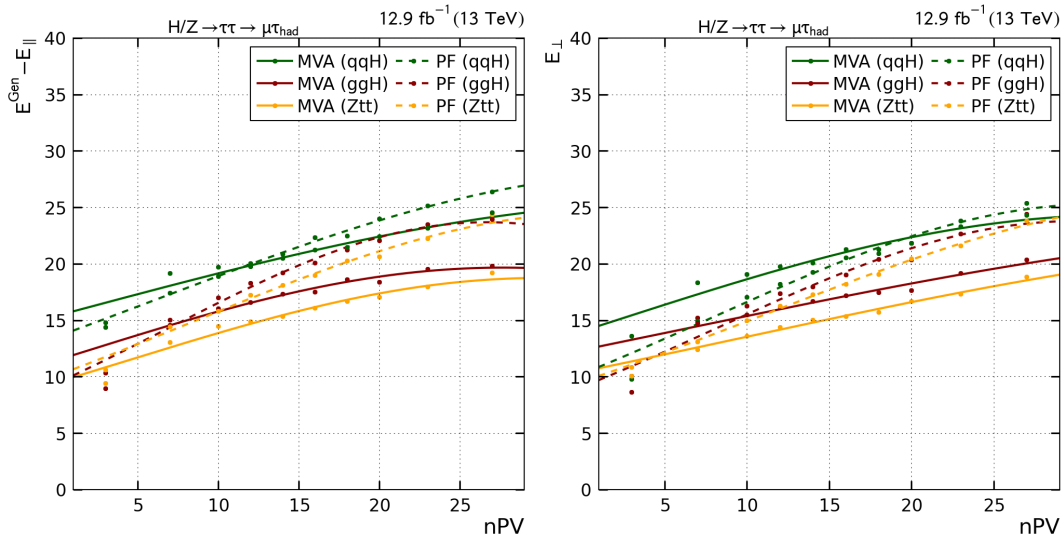


Figure 3.14: Resolution of the $\not{E}_{\mathrm{T}}^{\mathrm{PF}}$ and $\not{E}_{\mathrm{T}}^{\mathrm{MVA}}$ from the VBF Higgs boson production mode ($qq \to H$), gluon-gluon Fusion ($gg \to H$) and the Drell-Yan process. The VBF Higgs boson production is accompanied by two jets, which increase the $\not{E}_{\mathrm{T}}$ resolution in both components. $\not{E}_{\mathrm{T}}^{\mathrm{MVA}}$ has a smaller slope for all production processes, in the perpendicular as well as in the longitudinal components. Since the intercept for $\not{E}_{\mathrm{T}}^{\mathrm{MVA}}$ is also generally higher, the interesting point is the intersection. For first dataset of 2016 with an integrated luminosity of 12.9 fb^{-1}, a mean number of 15 reconstructed primary vertices is observed. So the crucial point is, if the intersection is above or below 15 reconstructed primary vertices, motivating the $\not{E}_{\mathrm{T}}$ to be used for the analysis. Only for the perpendicular component in $qq \to H$ this is the case, while all other intersections lie below. Therefore $\not{E}_{\mathrm{T}}^{\mathrm{MVA}}$ is expected to perform better than $\not{E}_{\mathrm{T}}^{\mathrm{PF}}$.

precise determination of the $\not{E}_{\mathrm{T}}$, as well as an estimate of the covariance matrix of the $\not{E}_{\mathrm{T}}$ on an event-by-event basis. During 2016, the covariance matrix for $\not{E}_{\mathrm{T}}^{\mathrm{PF}}$ became available allowing the comparison of performances. The reconstructed di-tau mass is the best discriminating variable between the two decays of the Z and the Higgs bosons to two tau. One of its crucial ingredients is the $\not{E}_{\mathrm{T}}$.

A comparison to measured data is not possible since there is no dedicated sideband to reach a high purity of $H \to \tau\tau$ signal events. Therefore, we only measure the performance on simulated data using the same metric as for W + Jets, using the two components $\not{E}_{\mathrm{T}}^{\parallel}$ and $\not{E}_{\mathrm{T}}^{\perp}$. The results can be seen in 3.14.

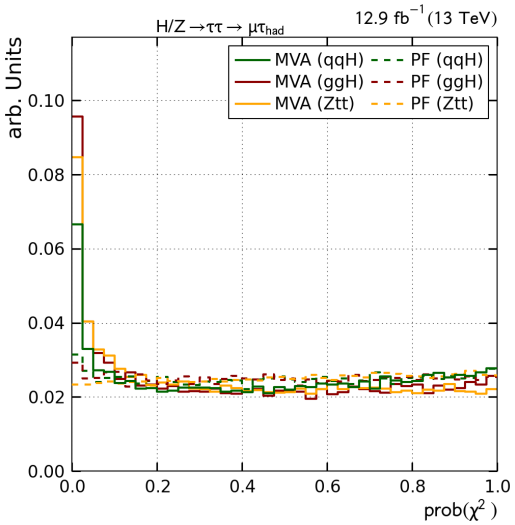


Figure 3.15: The prob (χ^2)distribution for the $\not{E}_{\mathrm{T}}^{\mathrm{MVA}}$and $\not{E}_{\mathrm{T}}^{\mathrm{PF}}$ distributions shows that there are still some problematic events in the $\not{E}_{\mathrm{T}}^{\mathrm{MVA}}$. But the broad range is modeled well in both definitions.

Measuring the covariance matrix estimation

The $\not{E}_{\mathrm{T}}$ covariance matrix V is probed by a χ^2 approach:

$$\chi^2 = \left(\vec{\not{E}}_{\mathrm{T}} - \vec{\not{E}}_{\mathrm{T}}^{\mathrm{True}}\right)^T \mathrm{V}^{-1} \left(\vec{\not{E}}_{\mathrm{T}} - \vec{\not{E}}_{\mathrm{T}}^{\mathrm{True}}\right) \tag{3.10}$$

which means that a good estimation of the covariance matrix is equivalent to a χ^2 distribution, i.e. randomly distributed. A good way to estimate the quality is to take the prob (χ^2). This should be flat, see Figure 3.15. For most events, this is the case. There is a step at low probabilities, that is higher for $\not{E}_{\mathrm{T}}^{\mathrm{MVA}}$as for $\not{E}_{\mathrm{T}}^{\mathrm{PF}}$.

3.7 Summary

It has been demonstrated that the recently trained $\not{E}_{\mathrm{T}}^{\mathrm{MVA}}$performs superior in nearly all scenarios compared to the commonly used $\not{E}_{\mathrm{T}}^{\mathrm{PF}}$. The full potential of the pile-up removing capabilities of the $\not{E}_{\mathrm{T}}^{\mathrm{MVA}}$will become evident once the instantaneous luminosity increases and therefore more pile-up collisions are expected [17].

The new $\not{E}_{\mathrm{T}}^{\mathrm{MVA}}$has been used in the 2015 and 2016 MSSM $H \to \tau\tau$ analyses [56, 57].

Chapter 4

Establishing the Higgs Boson Signal in the di-tau Final State

4.1 Analysis overview

The presented $H \to \tau\tau$ analysis strategy is in many aspects similar to the Run I analysis. It shares the cut-based event selection approach with the signal extraction from the fully reconstructed di-tau mass $m_{\tau\tau}^{SVFit}$.

The decay channels with at least one hadronic tau in the final state are analyzed, meaning that only about one out of nine $H \to \tau\tau$ decays is not in a final state covered by this analysis, see the fraction marked in red in figure 4.1. The fully hadronic channel is called $\tau_h\tau_h$, the semi-leptonic ones $e\tau_h$ and $\mu\tau_h$.

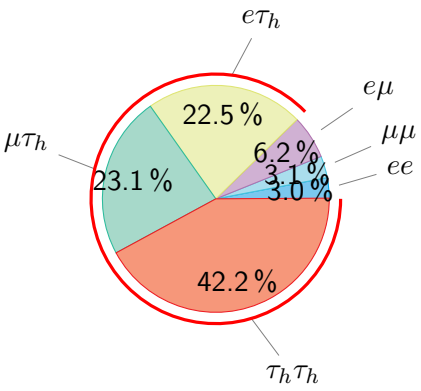


Figure 4.1: The branching ratios of di-tau systems

The channels $\mu\mu$ and ee are not included in the analysis, since their contribution to the overall significance is expected to be low for two reasons. One is the overall small number of expected Higgs boson events due to the small branching ratio of tau leptons to electrons or muons $BR(H \to \tau\tau \to ll) = 12.3\,\%$. Further, the irreducible $Z \to ll$ background is overwhelmingly high compared to the expected number of Higgs boson events even more than in the hadronic channels. The treatment of two dominant backgrounds, $Z \to \tau\tau$ and $Z \to ll$ makes special signal extraction methods necessary as e.g. shown in [41]. The $e\mu$ channel requires background estimation methods very different from the ones with hadronic taus, like a momentum-dependent ratio of the di-tau pairs with opposite-sign to the ones with same-sign. The background composition is also different to the hadronic channels, with $t\bar{t}$ plus jets being a dominant background depending on the event selection

which makes it rather challenging to control the channel. Since the branching ratio is also small with 6.2 %, it has not been analyzed in the context of considered for this thesis.

The presented $H \to \tau\tau$ analysis first reconstructs all possible di-tau pairs. These di-tau pairs have to fulfill several conditions that are introduced in the following. The conditions are there to suppress backgrounds while keeping the signal acceptance high. A set of corrections is applied, to improve the agreement between the simulated and observed events. After the selection, the events are *categorized* with the classification serving two purposes. One is the differentiation between the production processes $gg \to H$ and $qq \to H$, the other one is the improvement of the mass resolution by introducing boosted event categories. The statistical inference is performed on the fully reconstructed di-tau mass as final discriminator in 15 categories using the dataset with an integrated luminosity of $12.6\,\mathrm{fb}^{-1}$ which had been used and presented at the ICHEP 2016 conference.

4.2 Event reconstruction

This section quickly summarizes triggers used for the analysis, which physics objects are used for the analysis and how the Higgs boson is reconstructed from a selection of di-tau pairs from these objects. For a detailed description of the CMS experiment see section 2.2 and for the object reconstruction section 2.3.

4.2.1 Triggers

The three decay channels are based on three independent triggers. The CMS experiment has a sophisticated system capable of fully reconstructing physics objects at trigger level. These objects are called *online* objects, those used later in the analysis are called *offline* objects.

In the $\mu\tau_h$ channel, events firing the single-muon trigger with the transverse momentum requirement $p_\mathrm{T}^\mu > 22\,\mathrm{GeV}/c$ are selected. The transverse momentum requirement in the $e\tau_h$ channel using the single electron trigger where the electron has to have at least $p_\mathrm{T}^\mathrm{e} > 25\,\mathrm{GeV}/c$ is a bit higher. Additionally one is restricted to $|\eta^e| < 2.1$. Muons are more easy to identify and the misidentification rate is lower than the one of electrons, allowing a looser transverse momentum requirement for the single muon trigger. The $\tau_h\tau_h$ channel uses a di-tau trigger that requires two online hadronic taus in the event both having at least $p_\mathrm{T}^{\tau_h} > 35\,\mathrm{GeV}/c$ and being in the fiducial region $|\eta^{\tau_h}| < 2.1$. The higher transverse momentum requirements compared to the single-muon and single-electron trigger are the way to cope with the fact that many jets from QCD multijet production get misidentified as hadronically decaying taus.

Table 4.1: Lepton selection requirements for the constituents of di-tau pairs.

channel	trigger requirement	Lepton selection
$\mu\tau_h$	$p_{\mathrm{T}}^{\mu} > 22\,\mathrm{GeV}/c$	$p_{\mathrm{T}}^{\mu} > 23\,\mathrm{GeV}/c \& \lvert\eta^{\mu}\rvert < 2.4$ $p_{\mathrm{T}}^{\tau} > 20\,\mathrm{GeV}/c \& \lvert\eta^{\tau}\rvert < 2.3$
$e\tau_h$	$p_{\mathrm{T}}^{\mathrm{e}} > 25\,\mathrm{GeV}/c$ $\lvert\eta^{e}\rvert < 2.1$	$p_{\mathrm{T}}^{\mathrm{e}} > 26\,\mathrm{GeV}/c \& \lvert\eta^{e}\rvert < 2.1$ $p_{\mathrm{T}}^{\tau} > 20\,\mathrm{GeV}/c \& \lvert\eta^{\tau}\rvert < 2.3$
$\tau_h\tau_h$	$p_{\mathrm{T}}^{\tau_{1,2}} > 35\,\mathrm{GeV}/c$ $\lvert\eta^{\tau_{1,2}}\rvert < 2.1$	$p_{\mathrm{T}}^{\tau_1} > 40\,\mathrm{GeV}/c \& \lvert\eta^{\tau_1}\rvert < 2.1$ $p_{\mathrm{T}}^{\tau_2} > 40\,\mathrm{GeV}/c \& \lvert\eta^{\tau_2}\rvert < 2.1$

Table 4.1 summarizes the trigger requirements of the online objects and the requirements of the offline objects. The offline objects have to meet a higher transverse momentum requirement of 1 GeV/c ($\mu\tau_h$ and $e\tau_h$) respectively 5 GeV/c than the online objects requirements.

4.2.2 Vertices

All reconstructed leptons are required to have their origin in the leading primary vertex defined in equation 2.3. The fit reconstructing the primary vertex is done including the reconstructed taus, knowing this might bias the vertex position towards the secondary vertex of the tau lepton.

4.2.3 Electrons

Electrons passing a multivariate electron identification algorithm are used on the 90 % signal efficiency working point.

The precision of the electron energy scale measurement depends on the detector region they have been reconstructed in. In the barrel region ($|\eta_e| < 1.46$) a precision of 1 % has been measured, in the endcap region ($|\eta_e| > 1.46$) it is 2.5 %.

4.2.4 Muons

The muon identification algorithm uses two ways to identify an object as a valid muon:

1. If the tracks from the inside-out and outside-in muon reconstruction match well enough

2. The muon track has a sufficiently high fit quality and no kink.

Muons are allowed to have a distance of 0.045 mm from the primary vertex in the transverse direction and 0.2 mm along the beam axis.

The uncertainty on the momentum measurement for muons with a transverse momentum up to 100 GeV/c is well below 1 % (see Figure 2.12), and therefore considered negligible compared to the tau and electron energy scales.

4.2.5 Hadronically decaying Taus

Hadronically decaying taus are identified using the hadron-plus-strips algorithm, see section 2.3.4.

The MVA-Based isolation is rejecting about 99.2 % to 99.5 % of quark and gluon jets misidentified as hadronic taus while having an efficiency of 35 %.

There is no extra correction applied on the hadronic tau energy scale. A conservative uncertainty of 3 % on the hadronic tau energy scale is assumed, independent of the tau decay mode. The tau energy scale is treated as correlated among the final states.

4.2.6 Jets

Anti-k_T jets with a cone opening angle of 0.4 and a transverse momentum of $p_{\mathrm{T}} > 30\,\mathrm{GeV}/c$ are considered for the calculation of di-jet variables and to determine the number of jets, later used for the categorization.

All used jets in this analysis have been calibrated with the latest jet energy corrections in the usual CMS way, see section 2.3.3. The calibration modifies the transverse momenta of the reconstructed jets to the true energy on the parton level and is typically in the order of 10 %. The uncertainty on the determination of the jet energy scale enters the analysis as a shape uncertainty. This uncertainty is especially important since the categorization of events, to be introduced in the next section, largely depends on jet-related variables. The uncertainties on the jet energy determination are jet-p_{T} and η dependent and are in the range of 1 % to 10 %. The jet energy scale determination is independent from the final state and therefore treated as correlated among all channels and categories.

4.2.7 $\not\!\!E_{\mathrm{T}}$

The $\not\!\!E_{\mathrm{T}}^{\mathrm{MVA}}$ is used as described in the previous chapter with the combinatoric approach, assuring that the lepton selection used for the $\not\!\!E_{\mathrm{T}}^{\mathrm{MVA}}$ calculation exactly matches the final lepton selection of the analysis.

4.2.8 Di-tau pair building

The di-tau pair building algorithm is sophisticated in terms that it is able to cope with all kinds of ambiguity. First, the online leptons get matched to the offline leptons, meaning all leptons of the same flavors as the triggering ones get dropped

Table 4.2: Per channel: the total number of reconstructed di-tau pairs in the first 12.6 fb^{-1} of the CMS 2016 dataset, how many Higgs Boson events are expected and how many of them are expected to be reconstructed. The last column shows the overall signal significance.

channel	**valid $\tau\tau$ pairs**	**$H \to \tau\tau$ produced**	**$H \to \tau\tau$ reconstructed**	**selection efficiency**	$\frac{\mathbf{s}}{\sqrt{\mathbf{s+b}}}$
$\mu\tau_h$	289,399,228	4,053	1,623	40.0%	0.011
$e\tau_h$	439,854,362	4,125	898	21.8%	0.028
$\tau_h\tau_h$	169,504,648	15,066	783	5.2%	0.036

from consideration if they do not match the triggering lepton in a tight cone of $\Delta R < 0.2$.

In case different combinations of di-lepton pairs can be created, the following algorithm is applied. The term *leading* always refers to the object with the highest transverse momentum.

1. Prefer the pair with the most isolated lepton ($e\tau_h$ and $\mu\tau_h$) or leading hadronic tau ($\tau_h\tau_h$)

2. If the previous condition does not lead to a clear decision, take the pair with the highest electron p_T ($e\tau_h$), muon p_T ($\mu\tau_h$) or leading hadronic tau p_T ($\tau_h\tau_h$)

3. If still ambiguous, take the pair with the most isolated candidate hadronic tau ($e\tau_h$ and $\mu\tau_h$) or trailing hadronic tau ($\tau_h\tau_h$)

4. In case this fails, take the di-tau pair with the highest transverse momentum

By this procedure the momentum-sorting of the di-tau pair in the $\tau_h\tau_h$ channel is made sure automatically, meaning that $p_\text{T}^{\tau_1} > p_\text{T}^{\tau_2}$. It is further required that the two selected leptons have a distance in $\Delta R > 0.5$. This is necessary to make sure that a single particle is not interpreted twice, e.g. once as electron and once as hadronic tau.

The necessity for an advanced analysis strategy is obvious when looking at the numbers in table 4.2. Hundreds of millions of di-tau pairs get reconstructed while only a few thousand $H \to \tau\tau$ decays are expected, out of which many are expected to be below the trigger thresholds. The selection efficiency is calculated by comparing the total number of expected $H \to \tau\tau$ with the expected number of reconstructed $H \to \tau\tau$ events.

- $H \to \tau\tau$ **produced**: This is the number of all Higgs bosons produced times the branching ratio to two taus in the 12.6 fb^{-1} of integrated luminosity.

- $H \to \tau\tau$ **reconstructed**: This is the number of Higgs events for which a valid di-tau pair could be reconstructed, estimated from Monte Carlo simulation.

The efficiency is highly dependent on the trigger thresholds. It is highest for the $\mu\tau_h$ channel and lowest for the $\tau_h\tau_h$ channel with its high requirements on the tau transverse momenta. The signal significance is

$$s/\sqrt{s+b} \tag{4.1}$$

with s the number of Higgs boson events and the total number of recorded events $s + b$. The significance is below 0.01 for this first selection of di-tau pairs. It is obvious, that in order to get sensitive to the existence of the $H \to \tau\tau$ decay, advanced data analysis methods have to be used. The following section introduces the different kinds of backgrounds and how they are suppressed.

4.3 Simulated and recorded data for the $H \to \tau\tau$ search

Figure 4.2 lists the production cross sections of the different processes at the LHC for center-of-mass energies 7, 8 and 13 TeV. Most of them are backgrounds to the $H \to \tau\tau$ search.

This section presents in detail the various background processes, their estimation and suppression techniques as well as the associated uncertainties.

4.3.1 Recorded data

The dataset used in this analysis is the first $12.6\,\mathrm{fb}^{-1}$ taken in the year 2016. It contains data certified by the CMS data quality management group, meaning that it is ensured that all subdetectors have been working properly. The preliminary recommendation for the uncertainty for the 2016 data taking period is 2.5 % with a total integrated luminosity of $12.6\,\mathrm{fb}^{-1}$[59][1].

The number of recorded events per period can be found in the appendix in table A.1.

4.3.2 $H \to \tau\tau$

The mass of the SM Higgs boson is well known (see section 2.4), so a $H \to \tau\tau$ search of a mass range like in Run I is not necessary any more. Therefore, only simulated Higgs Boson events with the mass $m_\mathrm{H} = 125\,\mathrm{GeV}/c^2$ are used for this study. The processes $gg \to H$, $qq \to H$, VH and $qq \to ZH$ (see section 1.4.1) are used as signal samples.

[1] This reference documents the luminosity measurement procedure for the 2015 run that has not changed since.

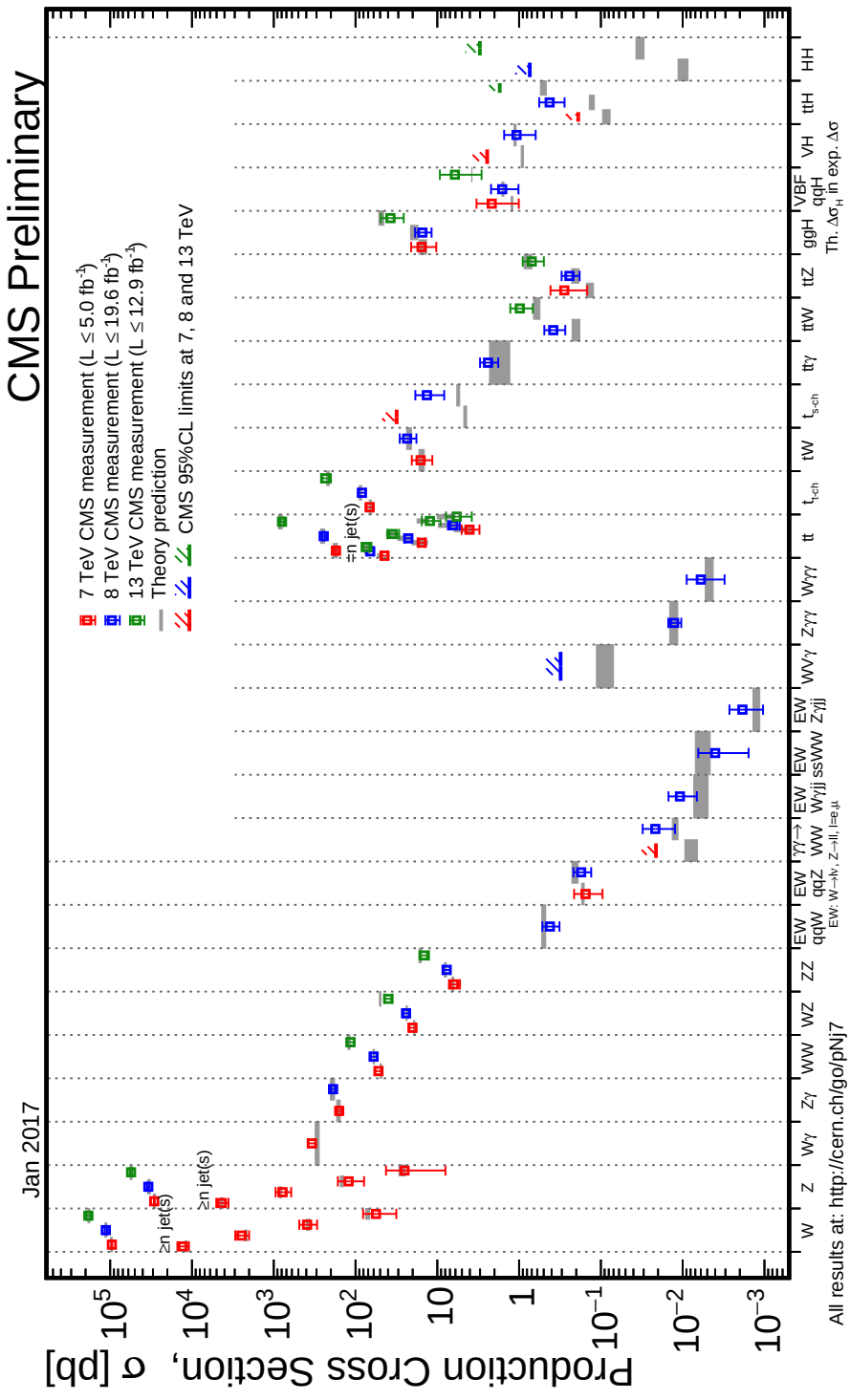


Figure 4.2: Cross sections of important production processes at the LHC with their prediction from the Standard Model and the measurements in Run I. The minimum-bias cross section for 13 TeV is $O(10^{11}\text{pb})$ [58]. Many of these processes are backgrounds to this analysis. The Z Boson production is the dominant, irreducible background. The W production in association with jets is the largest background caused by mis-identification of jets as hadronically decaying taus. Its cross section is orders of magnitude higher than the one of the Higgs boson that is listed on the right side of this plot.

The event generator POWHEG 2.0[60–62] has been used for the simulation interfaced to PYTHIA 8.1[63] to take into account the effects of parton showering. The $gg \to H$ inclusive production cross section has been calculate to N3NLO precision in QCD an next-to-leading order in elektroweak theory, the VBF production cross section at NNLO in QCD and NLO elektroweak theory.

The migration of events between categories has been studied by different simulations of the α_s, the QCD refactorization- and normalization scale and different parton density functions. In this approach, the acceptance of each category is studied. A limit of this model is that uncertainties with a common source like the uncertainty on α_s might be correlated among different production processes, but are treated as uncorrelated. This approach has been chosen for simplicity and is a conservative approximation.

QCD scale

The nominal QCD scale used for the simulation of Higgs boson events is $\mu = \mu_R = \mu_F = m_H = 125\,\mathrm{GeV}/c^2$. The difference in acceptance has been studied by the variation of the renormalization scale μ_R and refactorization-scale μ_F to 1/2 and 2 resulting in the acceptances $A_{\mu_R\mu_F}$ while leaving out the extreme combinations (see figure 4.3). The uncertainty has been estimated by

$$\frac{1}{2} \cdot \frac{A_{max} - A_{min}}{A_{\mu_R=1,\ \mu_F=1}} \tag{4.2}$$

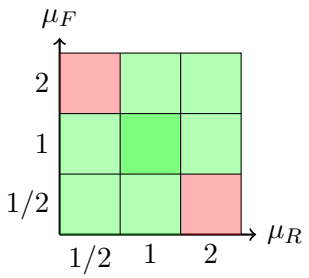


Figure 4.3: QCD scales variation scheme. The red parts are left out. The center is the nominal value, the light green ones are the variations.

The uncertainty is treated as correlated among different jet multiplicities. See figure 4.4 for the individual categories. The QCD scale uncertainty is especially high in the $gg \to H$ production process in association with one or two jets. The uncertainty on the inclusive Higgs boson production cross section is caused by the variation of μ_R and μ_F is taken from [64]. It is 3.9 % for $gg \to H$ production and 0.4 % for $qq \to H$ production. It is correlated among all channels and final states.

Uncertainty on α_s

The central value of α_s is $\alpha_s = 0.118 \pm 0.0011$ [65]. The effect on the categorization is estimated by the maximum difference of the up- and down-variations of α_s relative to the central value. The uncertainties can be found in figure 4.5. Since they have a common origin, they are treated as correlated. The uncertainty depends on the category and is at most 1.9 %.

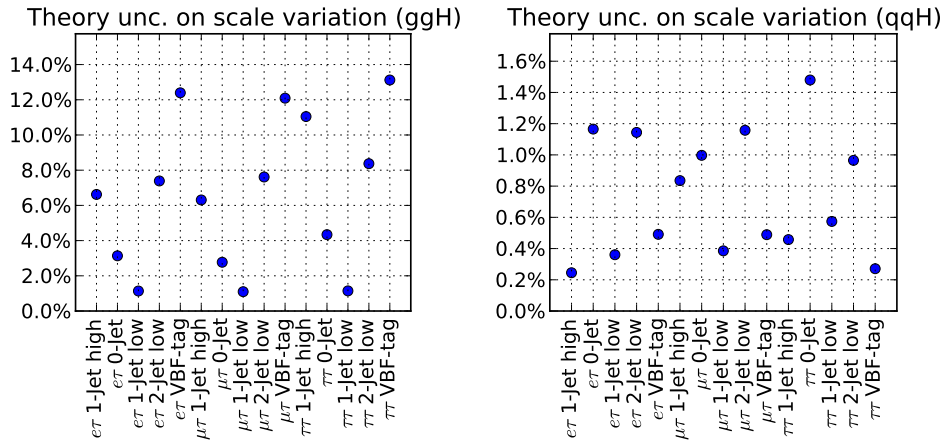


Figure 4.4: The uncertainties on the migration of Higgs Boson events between categories, caused by the variation of the QCD renormalization and refactorization scales for $gg \to H$ (left) and $qq \to H$ (right). The terminology concerning the categorization will be introduced later in section 4.6. For now it is only important to know that the *high* categories are defined via requirements on the boost of the reconstructed Higgs Boson system and via a requirement on the $p_{\mathrm{T}}^{\tau_h}$. The VBF-tag requires two reconstructed jets with a large gap in pseudorapidity in between. The QCD scale uncertainty for $qq \to H$ is small, while for $gg \to H$ production in association with one boosted jet or event with two jets it can be up to 13 %.

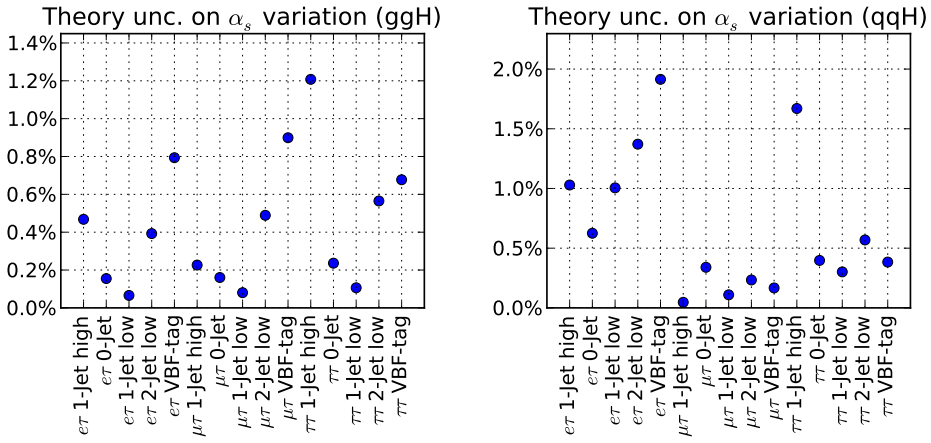


Figure 4.5: The uncertainties on the migration of Higgs Boson events between categories, caused by the uncertainty on α_s for $gg \to H$ (left) and $qq \to H$ (right).

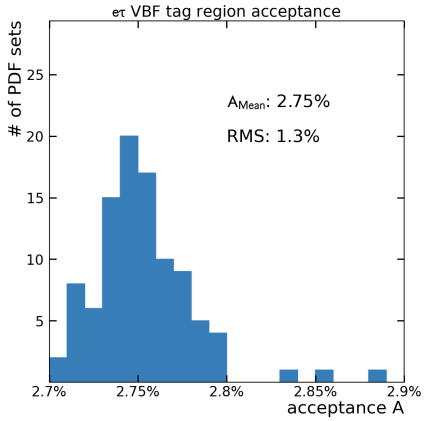


Figure 4.6: An example distribution of acceptances under the variation of PDF sets. Shown is the VBF tagged category in the $e\tau_h$ decay channel.

PDF uncertainty

For the simulation of the $H \to \tau\tau$ signal parton density functions have been chosen from the PDF4LHC15 [66]. For the VBF Higgs boson production additionally NNPDF3.0 [67] has been used for the electroweak corrections and photon parton density function.

The uncertainty on the parton density function is estimated by effective event weights of 100 PDF replicas. The standard deviation of the resulting acceptance variation is taken as PDF uncertainty. An example distribution of acceptances can be found in figure 4.6.

The uncertainty on the inclusive Higgs boson production cross section is caused by the uncertainty on α_s and the choice of the PDF set is taken from [64]. It is 3.2 % for $gg \to H$ production and 2.1 % for $qq \to H$ production. It is correlated among all channels and final states. See figure 4.7 for the resulting uncertainties.

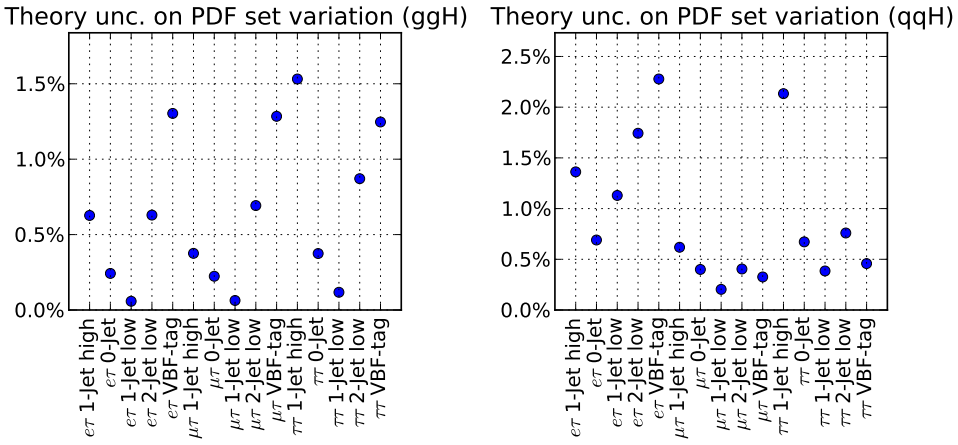


Figure 4.7: The uncertainty on the choice of the parton distribution function in the $gg \to H$ production mode (right) and $qq \to H$ production mode (left). The uncertainty is at most 2.3 %.

4.3.3 Drell-Yan Z background

About 70 % of all Z bosons decay directly into hadrons, but they also have a probability of 3.3 % each to decay into pairs of electrons, muons or taus. The remaining ones decay *invisibly*, which is the term for the decay into neutrinos that can not be detected in collider experiments. The main production process for the Z boson in hadron-hadron scattering is the *Drell-Yan* process. At leading order, a quark and an antiquark annihilate and form a pair of same-flavor, opposite charge leptons via exchange of a (virtual) Z boson or photon. One next-to-leading order process is $qg \to qZ \to ql^+l^-$, see Figure 4.8. Drell-Yan events in association with a jet can be boosted and have a higher reconstruction efficiency.

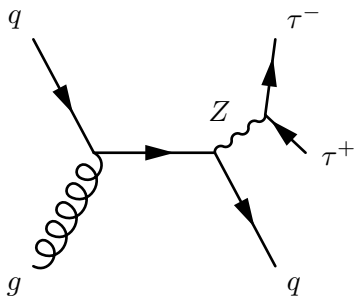


Figure 4.8: Drell-Yan Z production with an additional quark jet in the final state

The cross section for Z+Jets production drops by approximately one order of magnitude for each additional jet, see figure 4.2. The reason is that the Z boson production in association with jets is a higher-order effect. To ensure sufficiently large number of events from the simulation with at least two jets, this analysis uses a combination of special simulated samples focusing on the $Z + N_{jets}$ processes with $N_{jets} = 1, 2, 3, 4$. These samples get weighted relatively to each other to keep the inclusive shape and normalization but with improved statistics of simulated events for the higher number of jets.

Drell-Yan events have been simulated using the MADGRAPH 5[68] Monte Carlo generator. The parton shower and hadronization is done with PYTHIA 8 [69] in the CUETP8M1 [70] tune. The theory uncertainty on inelastic Drell-Yan Z boson production is 4 %, caused by the uncertainties on α_s, the QCD scale and the choice of the parton density function.

The largest irreducible background for the $H \to \tau\tau$ analysis is the $Z \to \tau\tau$ decay. The main difference between two taus in the final state being produced from a Higgs boson to the ones produced by the Z boson is the invariant di-tau mass. Since tau decays are always accompanied with the occurrence of undetectable neutrinos, the di-tau mass can only be estimated. The mass reconstruction algorithm used in this analysis SVFIT (*S*econdary *V*ertex fit) has been introduced in section 2.4.6.

The leptonic decay of the Z boson is denoted as $Z \to ll$, where $l = e, \mu$. This is a non-negligible background in case only one of the leptons properly reconstructed. The τ_h can result from

- the other l mis-identified as a hadronically decaying tau, mainly in the single hadron ($h^{\pm}$) mode and resulting in a usual di-tau signature ($Z \to l(l \to \tau_h)$)
- a jet from initial state radiation or pile-up that has been mis-identified as a

hadronically decaying tau ($Z \to l(\text{jet} \to \tau_h)$)

Both sources of background, $Z \to l(l \to \tau_h)$ and $Z \to l(\text{jet} \to \tau_h)$ are treated independently since they are linked to different uncertainties.

The $Z \to l(\text{jet} \to \tau_h)$ background can be suppressed by around 90% with a veto of additional leptons in the $\mu\tau_h$ and $e\tau_h$ channel. This veto is applied if apart from the selected di-lepton pair leptons, another electron ($e\tau_h$ channel) or muon ($\mu\tau_h$ channel) is present in the event, having a relative isolation $R^L < 0.3$ and a transverse momentum of $p_\mathrm{T} > 15\,\mathrm{GeV}/c$. The uncertainty on the rate of jets being misidentified as hadronic taus has been estimated in the context of the MSSM analysis[57], being 0.2 % per $1\,\mathrm{GeV}/c$ with a maximum of 40 %.

The $Z \to l(l \to \tau_h)$ background caused by electrons and muons mis-identified as hadronic taus, is suppressed by a special anti-electron and anti-muon discriminator (see section 2.3.4 for details). The background caused by electrons misidentified as hadronic taus is suppressed that way by 99.2 % in the $e\tau_h$ channel. The corresponding suppression efficiency of muons in the $\mu\tau_h$ channel is 96 %.

4.3.4 W + Jets

W + Jets events have been simulated with the same settings as the Drell-Yan event simulation. In the $\tau_h\tau_h$ channel, only 1 out of 40 events are expected to be from W + Jets production. It is therefore a subdominant background and its shape and yield are taken from simulation. The uncertainty on the inclusive yield has been estimated from the choice of the QCD scale, the α_s uncertainty and the choice of the PDF set is 4 % [71].

In the $\mu\tau_h$ and $e\tau_h$ channels 5 % to 10 % of the di-tau pairs in the final selection originate from W + Jets, making it a more important background. Its shape is taken from simulation, while the yield is controlled and eventually corrected as described in the following.

A variable reconstructing the W boson mass is introduced called m_T. It is defined as transverse mass component of the combined leptonic tau and $\not\!\!E_\mathrm{T}$ system:

$$m_\mathrm{T} \equiv \sqrt{2 p_\mathrm{T}{}^l \not\!\!E_\mathrm{T} (1 - \cos(\Delta\phi))} \tag{4.3}$$

with the difference in the azimuthal angle $\Delta\phi$ between the muon ϕ and the $\phi_{\not\!\!E_\mathrm{T}}$ in the transverse plain in the $\mu\tau_h$ channel and correspondingly the electron ϕ in the $e\tau_h$ channel. For muons or electrons originating from the resonant W production the m_T variable is higher than for muons or electrons from tau decays. The corresponding distributions of m_T can be found in figure A.2.

The m_T variable is used for the suppression of the W+Jets background by rejecting events with $m_\mathrm{T} > 40\,\mathrm{GeV}/c^2$. The W + Jets rejection shows superior performance by using the $\not\!\!E_\mathrm{T}^\mathrm{MVA}$, see figure 4.10 and table 4.3. The $gg \to H$ signal efficiency in

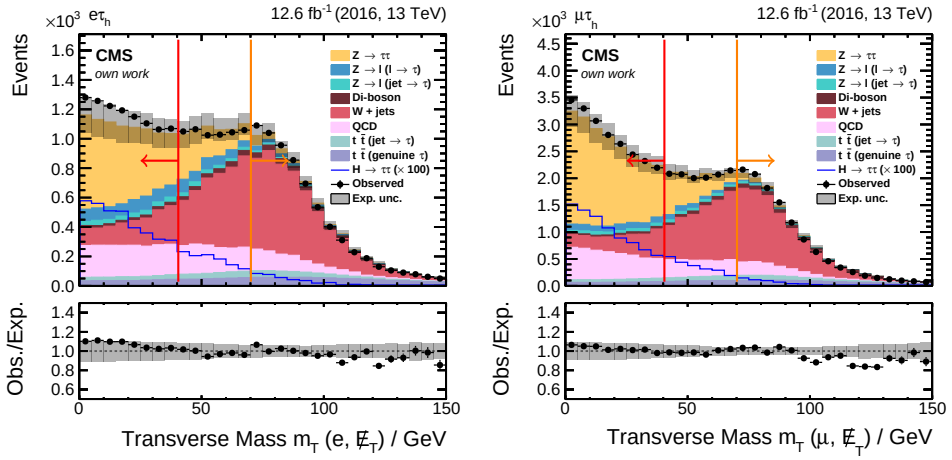


Figure 4.9: Distribution of m_{T} in the $e\tau_h$ channel (left) and $\mu\tau_h$ channel (right). The transverse mass of the lepton-$\not\!\!E_{\mathrm{T}}$ system m_{T} has its highest value around $80\,\mathrm{GeV}/c^2$ for the $W+\mathrm{Jets}$ process. The error bands represent the total systematic and statistic uncertainties. The red lines mark the signal regions, the orange lines the control regions used for the determination of the normalization.

	$gg \to H$ efficiency		W rejection	
	$\not\!\!E_{\mathrm{T}}^{\mathrm{MVA}}$	$\not\!\!E_{\mathrm{T}}^{\mathrm{PF}}$	$\not\!\!E_{\mathrm{T}}^{\mathrm{MVA}}$	$\not\!\!E_{\mathrm{T}}^{\mathrm{PF}}$
$\mu\tau_h$	0.791	0.767	0.813	0.795
$e\tau_h$	0.747	0.729	0.804	0.815

Table 4.3: Comparison of the $gg \to H$ signal efficiency and $W+\mathrm{Jets}$ rejection for the chosen m_{T} cut value of $m_{\mathrm{T}} < 40\,\mathrm{GeV}/c^2$.

the $\mu\tau_h$ channel is 3.1 % higher while rejecting 2.3 % more $W+\mathrm{Jets}$ events using $\not\!\!E_{\mathrm{T}}^{\mathrm{MVA}}$compared to $\not\!\!E_{\mathrm{T}}^{\mathrm{PF}}$. The $e\tau_h$ channel has a signal efficiency 2.5 % higher using $\not\!\!E_{\mathrm{T}}^{\mathrm{MVA}}$. The $W+\mathrm{Jets}$ rejection is 1.3 % smaller. Still, this means that a better signal to background ratio can be achieved by using the $\not\!\!E_{\mathrm{T}}^{\mathrm{MVA}}$for the calculation of the $\mu\tau_h$ variable.

The yield estimation is done in the high-m_{T} control region with $m_{\mathrm{T}} > 70\,\mathrm{GeV}/c^2$ that is dominated by $W+\mathrm{Jets}$ events. In events with opposite-sign di-tau pairs, the $Z \to \tau\tau$, $Z \to ll$, $\mathrm{t\bar{t}}$ and Di-Boson backgrounds are subtracted, leaving the estimate for the combined $W+\mathrm{Jets}$ and QCD yield, denoted as N^{OS}_{W+QCD}. The challenge now is that there is no QCD estimation and there is no adequate number of simulated QCD events that can correspondingly be subtracted.

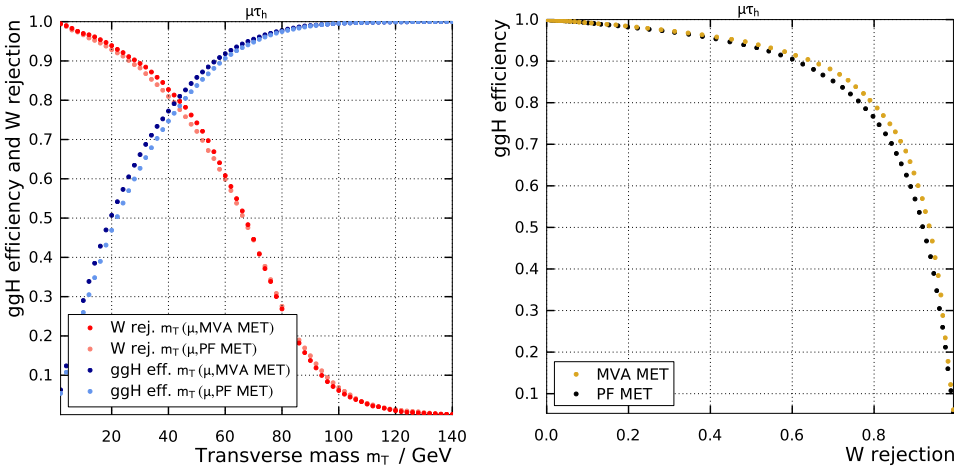


Figure 4.10: Evaluation of the effect of the choice of $\not{E}_{\mathrm{T}}$ definition on the transverse mass m_{T}. The left plot shows in blue the efficiency if $gg \to H$ events depending on the m_{T} cut value once for the m_{T} using $\not{E}_{\mathrm{T}}^{\mathrm{MVA}}$ and once for m_{T} using $\not{E}_{\mathrm{T}}^{\mathrm{PF}}$. The red points represent the W + Jets rejection under the two $\not{E}_{\mathrm{T}}$ options. The right plot shows the W rejection over the $gg \to H$ efficiency. The m_{T} definition using the improved $\not{E}_{\mathrm{T}}$ resolution from the $\not{E}_{\mathrm{T}}^{\mathrm{MVA}}$ has a superior efficiency and higher rejection than the m_{T} definition using the $\not{E}_{\mathrm{T}}^{\mathrm{PF}}$.

The number N^{SS}_{W+QCD} is determined for events with same-sign di-tau pairs in the same manner as N^{OS}_{W+QCD}. These N are combined with the different same-sign to opposite-sign di-tau pair extrapolation factors for W + Jets and QCD. For W + Jets, the factor $F_W^{SS\to OS}$ is taken from simulation. See section 4.3.7 for the determination of the factor $F_{QCD}^{SS\to OS}$. Typical values are 1.0 to 1.2 (Figure 4.14).

The full system of equations this is:

$$
\begin{aligned}
N^{OS}_{W+QCD} &= N^{OS}_{W} + N^{OS}_{QCD} \\
N^{SS}_{W+QCD} &= N^{SS}_{W} + N^{SS}_{QCD} \\
F_W^{SS\to OS} &= \frac{N^{OS}_{W}}{N^{SS}_{W}} \quad \text{(from simulation)} \\
F_{QCD}^{SS\to OS} &= \frac{N^{OS}_{QCD}}{N^{SS}_{QCD}} \quad \text{(from separate estimation)}
\end{aligned}
\tag{4.4}
$$

This system is solved to get the number N^{OS}_{W} in the high-m_{T} region. See figure 4.11 for the $m_{\tau\tau}^{SVFit}$ distributions in the high-m_{T} region. The number of W + Jets events with opposite sign in the high-m_{T} region N^{OS}_{W} is extrapolated with a high- to low-m_{T} extrapolation factor taken from simulation $F_W^{m_{\mathrm{T}}^{\mathrm{high}}\to m_{\mathrm{T}}^{\mathrm{low}}}$ to the signal region

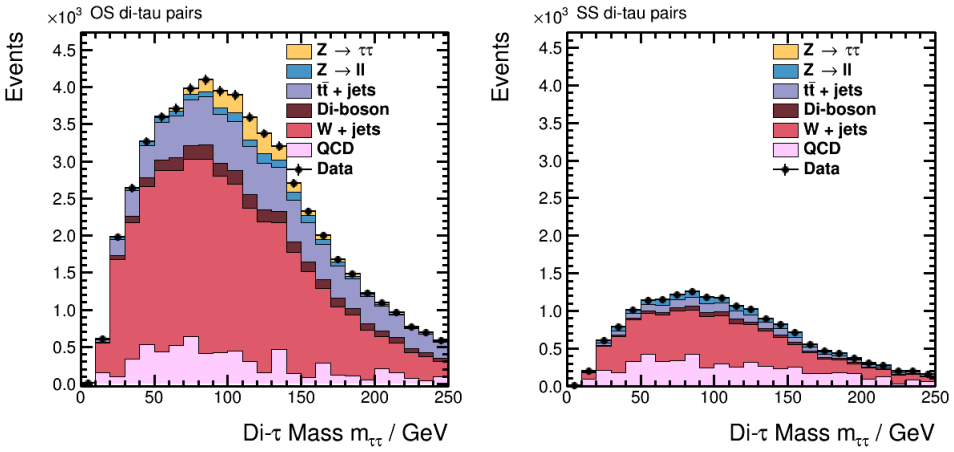


Figure 4.11: Distribution of the fully reconstructed di-tau mass $m_{\tau\tau}^{SVFit}$ in the high-m_{T} region ($m_{\mathrm{T}} > 70\,\mathrm{GeV}/c^2$). *Left:* opposite-sign pairs, *right* same-sign pairs. The yield of QCD events is approximately the same while $W+\mathrm{Jets}$ events are three times more likely to have opposite charge in their reconstructed di-tau pair. The $W+\mathrm{Jets}$ normalization in this figure is taken from simulation, the remaining events are interpreted as QCD. Therefore, the measured data points match the background estimation by construction.

$$N_W^{m_{\mathrm{T}}^{\mathrm{low}}} = N_W^{OS} \cdot F_W^{m_{\mathrm{T}}^{\mathrm{high}} \to m_{\mathrm{T}}^{\mathrm{low}}} \tag{4.5}$$

which is the estimate for the $W+\mathrm{Jets}$ normalization in the inclusive signal region.

As explained later, the analysis is performed in categories that each contain a subset of the events of a decay channel, in the most extreme case only 0.13 % of all events. The $W+\mathrm{Jets}$ estimation is done depending on the number of jets with $p_{\mathrm{T}} > 30\,\mathrm{GeV}/c$ that can be zero, one or two or more. This number is also called *jet multiplicity*. A scale factor from the selection of events within a jet multiplicity $F_W^{\text{jet mult.}\to\text{final sel.}}$, determined on simulated events, extrapolating to the final categorization is applied.

$$N_W^{\mathrm{final}} = N_W^{m_{\mathrm{T}}^{\mathrm{low}}} \cdot F_W^{\text{jet mult.}\to\text{final sel.}} \tag{4.6}$$

4.3.5 $\mathrm{t\bar{t}}$

The production of $\mathrm{t\bar{t}}$ pairs happens via the gluon fusion process $gg \to \mathrm{t\bar{t}}$ or by quark annihilation $q\bar{q} \to \mathrm{t\bar{t}}$, see Figure 4.12. The $\mathrm{t\bar{t}}$ pair production has been simulated using POWHEG[60] interfaced to Pythia 8 [69] for the parton shower and hadronization. Nearly all top quarks decay into a W boson and a b-jet. This gives an event signature very close to the one of $H \to \tau\tau$ decays. The $\mathrm{t\bar{t}}$ pair events tend to have high jet

multiplicity. The shape and normalization of the $t\bar{t}$ background in all channels is taken from simulation.

The cross section has been determined to $\sigma_{t\bar{t}} = 831.76^{+19.77}_{-29.20}(\text{scale}) \pm 35.06(\text{PDF} + \alpha_s)\text{pb}$ [72] as calculated with the Top++2.0 program to next-to-next-to-leading order in perturbative QCD, including soft-gluon resummation (see [73] and references therein), and assuming a top-quark mass $m_{top} = 172.5\,\text{GeV}/c^2$. The first uncertainty comes from the independent variation of the factorization and renormalization scales, μ_F and μ_R, while the second one is associated to variations in the PDF and α_s, following the PDF4LHC prescription with the MSTW2008 68 % CL NNLO, CT10 NNLO and NNPDF2.3 5f FFN PDF sets (see [66] and references therein). As combined $t\bar{t}$ production cross section uncertainty, 6 % is used.

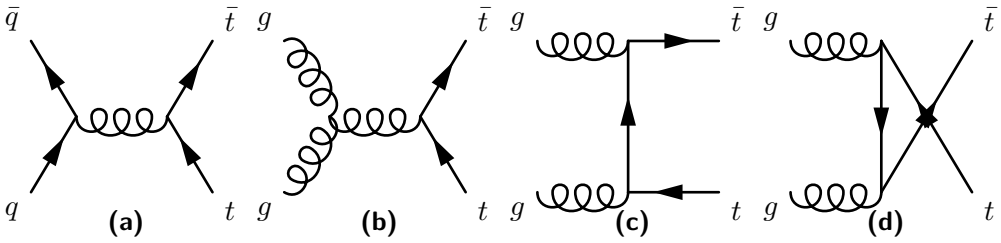


Figure 4.12: Feynman diagrams in leading order that contribute to top quark pair production at the LHC.[74]

4.3.6 Di-Boson

The Di-Boson production consisting of two W, two Z or WZ contributes both by real taus decaying hadronically as well as jets misidentified as hadronic taus. The Di-Boson production cross section is the smallest of all considered backgrounds. It has been simulated using the Madgraph 5 software package at next-to-leading order accuracy [75, 76] with the FxFx merging scheme.

The total production cross section uncertainty resulting from the variation of μ_R and μ_F, α_s and the PDF sets on the Di-Boson production is 10 % [71].

4.3.7 QCD Multijet

The QCD multijet background summarizes all QCD induced processes initiated by light quarks or gluons that lead to final states with large jet multiplicities.

The simulation of the QCD contribution in the $H \to \tau\tau$ analysis is challenging. The QCD production cross section is orders of magnitude higher than the one of all other background processes. The yield in the analysis however is small, due to the high efficiency of to reject this background by the lepton and τ_h isolation requirements. The usual simulation approach would require a huge amount of QCD

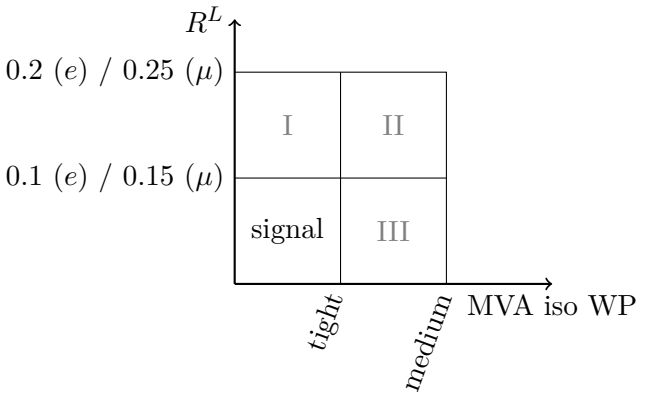


Figure 4.13: Three opposite-sign to same-sign extrapolation factors are calculated in the region with an anti-isolated lepton but isolated tau (region I), lepton and tau anti-isolated (II) and lepton isolated but tau anti-isolated (III). These factors are used to extrapolate to the signal region where both the lepton and tau are isolated.

events for a proper background estimation. The choice therefore is to derive the QCD contribution from data in sideband regions.

In the $\mu\tau_h$ and $e\tau_h$ channels, the QCD shape is estimated from recorded events with same-sign di-tau pairs. From data, the $Z \to \tau\tau$, $Z \to ll$, Di-Boson, $\mathrm{t\bar{t}}$ and W + Jets contributions are subtracted leading to a QCD shape estimation. The normalization is calculated by a scale factor $F_{QCD}^{SS \to OS}$, which is introduced in the following. See figure 4.15 for the $m_{\tau\tau}^{SVFit}$ distribution in the $\mu\tau_h$ channel with same-sign di-tau pairs.

The assumption that one can use this approach to estimate QCD as background for $H \to \tau\tau$ by subtracting all other backgrounds from observed data. This is only valid when the region is free of $H \to \tau\tau$ signal. For that reason, the scale factor $F_{QCD}^{SS \to OS}$ is derived in sideband regions with inverted isolation criteria as demonstrated in figure 4.13. The same-sign to opposite-sign di-tau pair extrapolation factors are derived in regions where either the lepton, the hadronic tau or both are anti-isolated. Anti-isolated means a relative $\Delta\beta$ corrected isolation $0.1 < R^L < 0.2$ for electrons, $0.15 < R^L < 0.25$ for muons and for hadronic taus having a MVA score lower than 0.85. From these three factors the extrapolation to the signal region is obtained by

$$F_{signal}^{OS/SS} = \frac{F_{III}^{OS/SS} \cdot F_{II}^{OS/SS}}{F_{I}^{OS/SS}} \tag{4.7}$$

See figure 4.14 for the final scale factors and their uncertainties.

The QCD multijet background is suppressed by the isolation criteria using a working point of $R^L < 0.15$ for muons and $R^L < 0.1$ for electrons in the $\mu\tau_h$ and $e\tau_h$ channel (see formula 2.38 and 2.15). The MVA tau isolation, introduced in section 2.3.4 is applied on the tight working point. The isolation variable distributions can be found in figure 4.16.

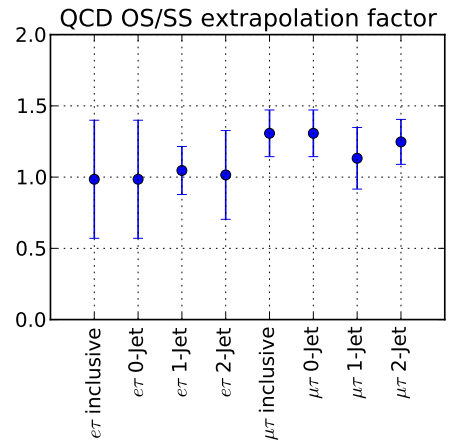


Figure 4.14: The scale factors $F^{OS/SS}$. The uncertainties shown are the statistical uncertainties plus an inclusive uncertainty of 10 % on the individual components that have been subtracted.

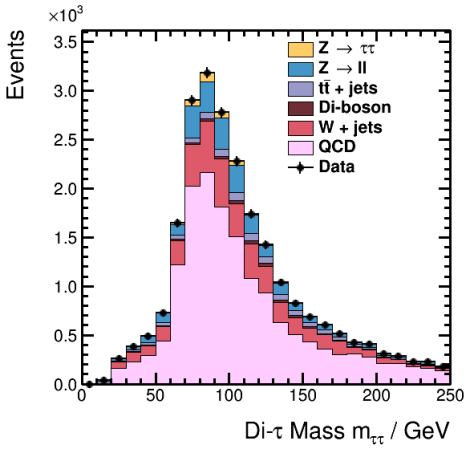


Figure 4.15: QCD background estimation, here in the $\mu\tau_h$ channel. In same-sign di-tau pairs all known backgrounds are subtracted. The remaining events are interpreted as QCD events.

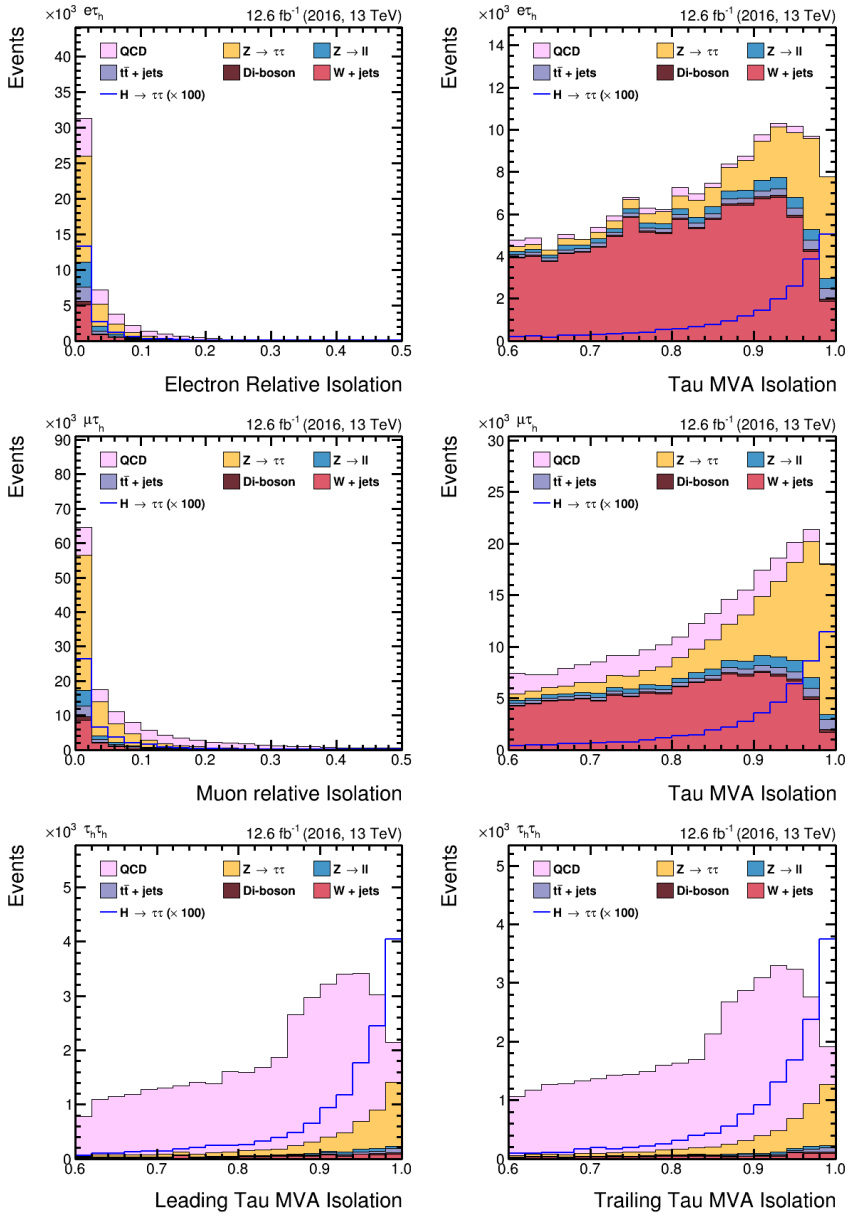


Figure 4.16: Lepton isolation variables for the $e\tau_h$ (top), $\mu\tau_h$ (middle) and $\tau_h\tau_h$ (bottom) channels. The electron and muon isolation are the relative isolations R^L, the hadronic tau isolation is the BDT score where the tight working point corresponds to a score of 0.85. Events not passing the isolation requirement of the corresponding pair lepton are not included in these distributions, which means that e.g. in the R^L distribution for electrons in the $e\tau_h$ channel, the hadronic tau MVA isolation requirement is already applied.

4.3.8 A general remark on simulated datasets

A simulated event has three, basically independent ingredients. One is the above-mentioned hard scattering processes, defining the ingredients in the event the analysis is interested in. The second is the **underlying event** consisting of the remnants of the proton that did not take part in the hard scattering process. The third part, the pile-up, consists of additional soft proton-proton collisions. These additional collisions are simulated independently from the simulation of the hard scattering process. Both are superimposed before the full event reconstruction is run. The third is the **pile-up** that can have two reasons. One are additional proton-proton collisions. The other is the **out-of-time** pile-up which is the effect that within the short period of 25 ns two proceeding collisions happen. This can cause a residual signal in some sub-detectors. The full event simulation always contains all three ingredients together, making the simulated events as realistic as possible.

A special kind of challenge in the analysis is imposed by limited statistic of simulated events. Especially in signal categories with a very tight selection, $m_{\tau\tau}^{SVFit}$ histograms may be sparsely populated, leading to large fluctuations in neighboring bins. This gives a natural limit in how tight one can define the categories. The corresponding uncertainty is incorporated by the **bin-by-bin** uncertainties. The bin-by-bin uncertainties assign each individual process in each bin the poissonian error of the weighted bin content as uncertainty.

4.4 Correction Factors for the Simulation

Simulating processes always comes with assumptions, that might reflect reality only to a certain degree. The following section presents the corrections that have been applied on simulated data in order to improve their description of measured data.

4.4.1 Pile-up reweighting

The Monte Carlo event generation typically starts before data taking, allowing analysts to be prepared before data taking. This comes with the downside that one is forced to assume some kind of pile-up scenario, because the $\mathcal{L}_I$ and n_b in formula 2.6 change basically from collision to collision. This is handled by a Poisson distribution of pile-up interactions as close as possible to the expectation of the run period. The residual differences are corrected using a re-weighting procedure. The effect on the number of reconstructed primary vertices can be found in figure 4.17. The agreement between simulation and data is improved. This is important since many variables that are used in the analysis naturally depend on the pile-up, e.g. the number of reconstructed jets or the lepton isolation.

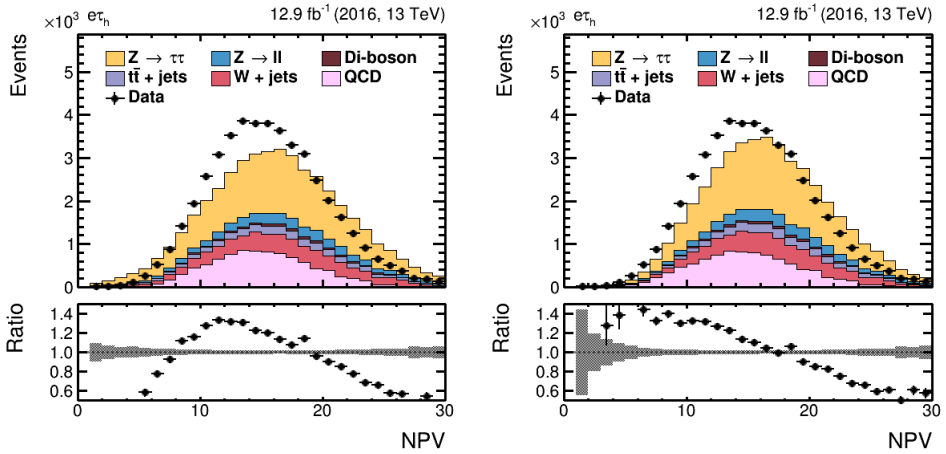


Figure 4.17: The number of primary vertices before *(left)* and after *(right)* the pile-up reweighting in the $e\tau_h$ decay channel. The agreement is slightly improved for the lower number of primary vertices.

4.4.2 Recoil Correction

The hadronic recoil, defined as in equation (), is not well modeled in the simulation of the Drell-Yan, W + Jets and Higgs boson events. This is corrected by the so-called **recoil corrections** that are applied on the recoil of the hard scatter objects comparable to the $\not{E}_{\mathrm{T}}^{\mathrm{MVA}}$. The $\not{E}_{\mathrm{T}}^{\mathrm{MVA}}$procedure improves the resolution while the recoil corrections improve the agreement between simulation and data.

In a $\mu\mu$ selection, one observes a disagreement of the simulation with the observation in regions that are dominated by the $Z \to \mu\mu$ process having no genuine $\not{E}_{\mathrm{T}}$ while the modeling in the high-$\not{E}_{\mathrm{T}}$ region is good. The simulation is calibrated by correcting the differences to data of the two recoil components $u_{\parallel}$ and $u_{\perp}$, defined like in figure 3.1 as the parallel and perpendicular projections of the recoil on the di-muon transverse momentum vector.

The used method assumes a Gaussian distribution of the parallel and perpendicular recoil components. The resolutions $\sigma(u_{\parallel})$ and $\sigma(u_{\perp})$ are determined on both observed and simulated data, additionally for $u_{\parallel}$ the mean $\left\langle u_{\parallel} \right\rangle$ is determined. The perpendicular component $u_{\perp}$ is assumed to be symmetrical and with the mean at 0.0.

The correction factors hence are

$$u'_{\parallel} = \left\langle u_{\parallel} \right\rangle_{data} + \left(u_{\parallel} - \left\langle u_{\parallel} \right\rangle_{MC} \right) \frac{\sigma\left(u_{\parallel}\right)_{data}}{\sigma\left(u_{\parallel}\right)_{MC}} \tag{4.8}$$

$$u'_{\perp} = u_{\perp} \frac{\sigma\left(u_{\perp}\right)_{data}}{\sigma\left(u_{\perp}\right)_{MC}} \tag{4.9}$$

They are calculated and propagated to a corrected $\not{E}_{\mathrm{T}}^{\mathrm{MVA}}$.

The corrections have been derived dependent on the occurrence of no, one or two or more jets with a transverse momentum of $p_{\mathrm{T}} > 30\,\mathrm{GeV}/c$ and dependent on the magnitude of the recoil, where the borders are 0, 10, 20, 30, 50 GeV/c and more than 50 GeV/c.

The systematic uncertainty on the $\not{E}_{\mathrm{T}}^{\mathrm{MVA}}$response and resolution corrections are up to 3 % for Drell-Yan, W+Jets and Higgs boson production and 3 to 11 % for $\mathrm{t\bar{t}}$. The uncertainties on the response and resolution are incorporated by systematic shifts of the $\not{E}_{\mathrm{T}}$ response and resolution and re-calculation of the di-tau mass $m_{\tau\tau}^{SVFit}$. Due to differences in the $\not{E}_{\mathrm{T}}^{\mathrm{MVA}}$performance depending on the lepton selection, these uncertainties are considered to be uncorrelated among jet multiplicities and decay channels.

4.4.3 Lepton identification and isolation scale factors

Lepton identification and isolation scale factors are depend on the lepton p_{T} and $|\eta|$, assigning each event individually a weight. This weight compensates differences on the lepton identification and isolation efficiency in simulated and data.

The scale factors have been derived by a tag-and-probe method using $Z \to ee$ and $Z \to \mu\mu$ events in the context of the 2016 MSSM Analysis [57]. The first lepton is used as a tag, required to fulfill the usual kinematic, identification and isolation requirements. In data, it additionally has to be matched to the triggering lepton within a cone of $\Delta R < 0.5$. The probe leptons only have to pass a set of kinematic cuts. A selection targeting Z boson decays is performed, requiring the di-lepton pair to have the same flavor, opposite charge, being well separated with a distance of at least $\Delta R > 0.5$ and having an invariant mass of $m_{ll} > 50\,\mathrm{GeV}/c^2$. The efficiencies have been extracted by a simultaneous fit of an exponential function to the background and two asymmetric Gaussian to model the signal in a mass range of $75\,\mathrm{GeV}/c^2 < m_{ll} < 105\,\mathrm{GeV}/c^2$. The uncertainty on both the electron and muon identification and isolation efficiency has been estimated to be 2 %. See figure 4.19 for the impact of the identification and isolation scale factors.

4.4.4 Tau identification efficiency scale factors

To determine the efficiency of the τ reconstruction, a selection to aim for $Z \to \tau\tau$ events with one τ decaying to a muon and the other decaying hadronically was performed. This was done by selecting events with one muon and exactly one oppositely charged τ_h candidate. Depending on the efficiency of the working point in

question, a scale factor of 0.9 ± 0.1 was extracted by a maximum likelihood fit to the data. The scale factor is independent of the τ decay mode. The improvement of the application of the scale factor can be followed in comparison of figure 4.18a and 4.18b.

The uncertainty of 10 % is split up in two parts. The first is a correlated uncertainty among all final states of 8 % for the semileptonic and 16 % for the fully hadronic channel. The uncorrelated part is 4 % for $\mu\tau_h$ and $e\tau_h$ channels and 10 % for the $\tau_h\tau_h$ channel, which includes a 3.5 % uncertainty for each leg related to the trigger efficiency.

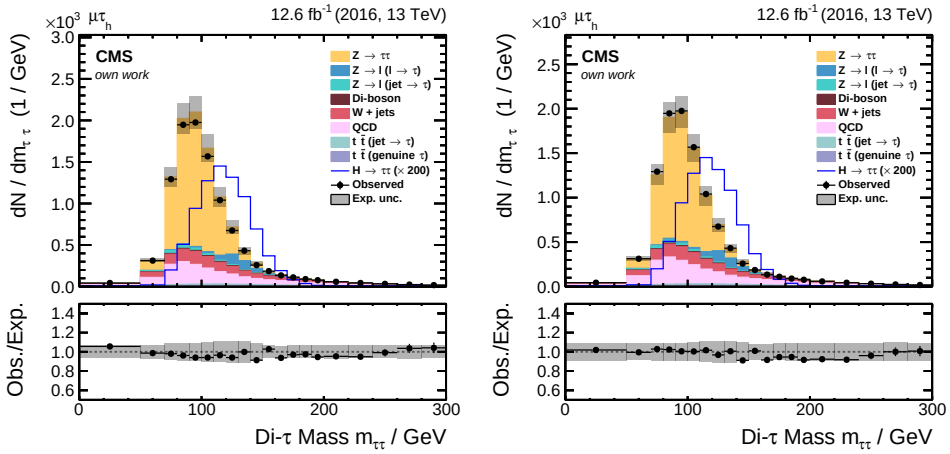


Figure 4.18: The $m_{\tau\tau}^{SVFit}$ distribution in the $\mu\tau_h$ channel without (left) and with (right) the hadronic tau identification scale factor applied. The agreement of simulated and observed data in the mass region of the Z resonance is improved.

4.4.5 Trigger efficiencies

The simulated events lack the trigger information, which corresponds to an acceptance of 100 %. The trigger efficiencies on data are also derived using the tag-and-probe technique, following the same approach as for the lepton identification and isolation measurement.

In the semi-leptonic channels the probes are electrons and muons from triggers with loser requirements on the trigger objects. The probe can either pass or fail, depending on the trigger decision of the final high level trigger. The $\tau_h\tau_h$ trigger efficiency is determined by comparing the efficiencies of a single muon trigger and a $\mu + \tau$ cross-trigger. That way, the efficiency of an individual tau leg has been measured. See figure 4.19 for the impact of the trigger efficiency scale factors.

4.4.6 Anti-lepton discriminator scale factors

The $Z \to l(l \to \tau_h)$ component of the $Z \to ll$ background is reduced by an anti-electron respectively an anti-muon veto on the τ_h candidate. The description in data is largely improved by the application of two corresponding scale factors. The factor depends on the final state, the pseudorapidity of the τ_h and ranges from 1.4 to 2.6. Especially where $Z \to ll$ is the dominant background, the data description is largely improved by the anti-lepton discriminator scale factors. See table 4.4 for the applied scale factors.

Table 4.4: Scale factors correcting the differences in rate between simulation and observation for electrons (left) and muons (right) misidentified as hadronically decaying taus.

$Z \to e(e \to \tau_h)$				$Z \to \mu(\mu \to \tau_h)$			
$\lvert\eta\rvert$ range			scale factor	$\lvert\eta\rvert$ range			scale factor
0.00	-	1.46	$1.51^{+0.07}_{-0.12}$	0.0	-	0.4	1.5 ± 0.1
1.46	-	1.58	-	0.4	-	0.8	1.4 ± 0.1
1.58	-	2.30	$2.00^{+0.35}_{-0.35}$	0.8	-	1.2	1.2 ± 0.1
				1.2	-	1.7	2.6 ± 0.9
				1.7	-	2.3	2.1 ± 0.9

4.4.7 Energy scale of leptons misidentified as hadronic taus

If electrons or muons misidentified as τ_h are not rejected from the anti-electron and anti-muon discriminator, it means that they have been very badly reconstructed. This is, e.g., the case for muons without a hit in the muon system.

The energy of misidentified leptons is calibrated using the Z resonance. In events without a reconstructed jet with $p_\mathrm{T} > 30\,\mathrm{GeV}/c$, $Z \to ll$ is the dominant background around the nominal Z mass. Here, a disagreement between the expectation and observation of up to 40 % is visible, see Figure 4.20.

To calibrate the energy of misidentified leptons, a constant function has been fitted to the ratio of observed events vs. expected events. The fit result is not used, but its χ^2 value. The transverse momentum of the lepton p_T^l is varied in steps of 0.5 % between 97 % and 15 % and the visible mass is recalculated:

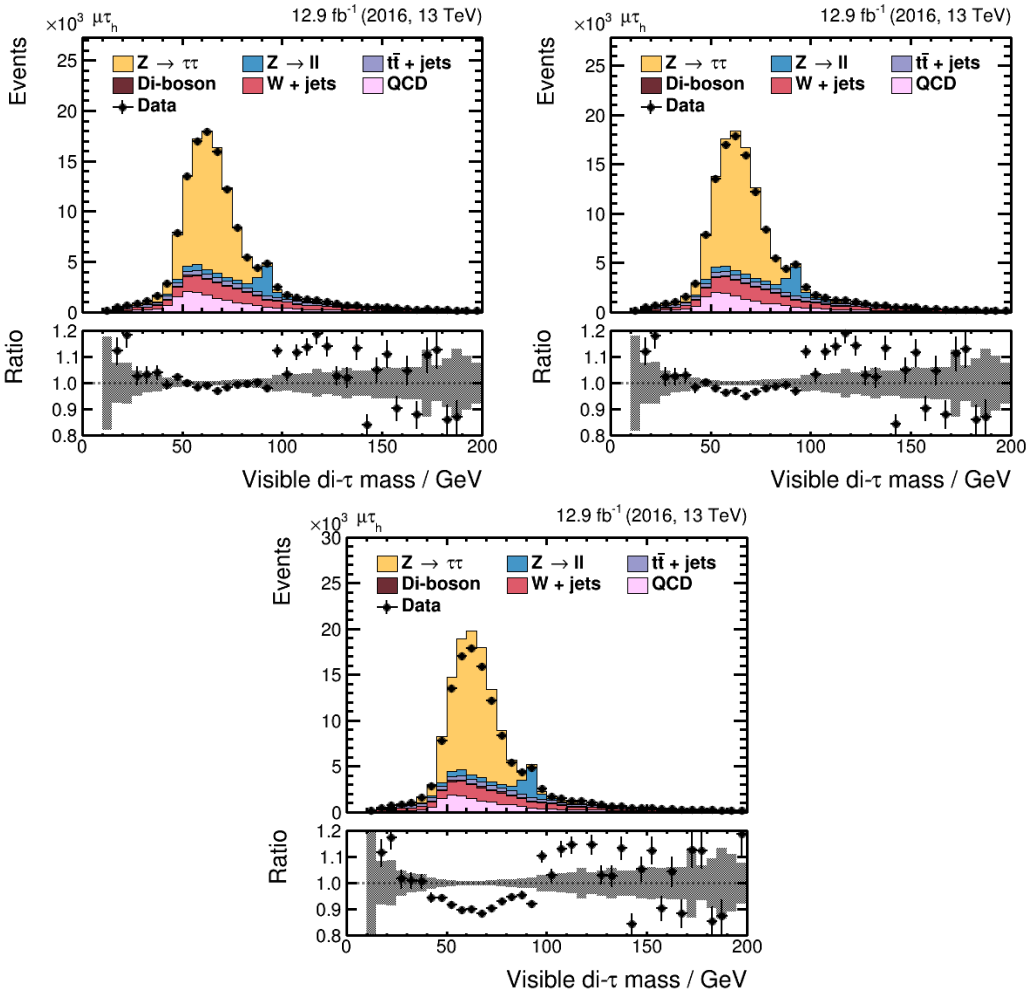


Figure 4.19: Comparison of the inclusive reconstructed di-tau mass in the $\mu\tau_h$ final state with the error bands only reflecting statistical uncertainties. In the top left plot, all corrections are applied. The top right plot is without the identification and isolation scale factors applied. This modifies the acceptance slightly, leading to an overestimation of the backgrounds. The bottom plot is without the trigger scale factors applied on simulation, leading to a mismodelling of around 10 %.

	$\lvert\eta\rvert < 0.9$	$0.9 < \lvert\eta\rvert < 1.46$	$\lvert\eta\rvert > 1.46$
$e\tau_h$	1.07	1.09	1.0
$\mu\tau_h$	1.01	1.01	1.0

Table 4.5: Minima of the fitted parabola in bins of $|\eta|$, corresponding to the final scale factor. The correction is 7 % to 9 % for electrons and at most 1 % for muons. In the forward region, no correction is necessary.

$$
\begin{aligned}
(m_{\tau\tau}^{vis})^2 = \Bigg(& \sqrt{m_l^2 + \left(p_{\mathrm{T}}^{\mathrm{l}} * \cos\phi_1\right)^2 + \left(p_{\mathrm{T}}^{\mathrm{l}} * \sin\phi_1\right)^2 + \left(\frac{p_{\mathrm{T}}^{\mathrm{l}}}{\tan(2 * \arctan e^{-|\eta_l|})}\right)^2} \\
& + \sqrt{m_{\tau_h}^2 + \left(p_{\mathrm{T}}^{\tau_h} * \cos\phi_{\tau_h}\right)^2 + \left(p_{\mathrm{T}}^{\tau_h} * \sin\phi_{\tau_h}\right)^2 + \left(\frac{p_{\mathrm{T}}^{\tau_h}}{\tan(2 * \arctan e^{-|\eta_{\tau_h}|})}\right)^2}\Bigg)^2 \\
& - \left(p_{\mathrm{T}}^{\mathrm{l}} * \cos\phi_l + p_{\mathrm{T}}^{\tau_h} * \cos\phi_{\tau_h}\right)^2 \\
& - \left(p_{\mathrm{T}}^{\mathrm{l}} * \sin\phi_l + p_{\mathrm{T}}^{\tau_h} * \sin\phi_{\tau_h}\right)^2 \\
& - \left(\frac{p_{\mathrm{T}}^{\mathrm{l}}}{\tan(2 * \arctan e^{-|\eta_l|})} + \frac{p_{\mathrm{T}}^{\tau_h}}{\tan(2 * \arctan e^{-|\eta_{\tau_h}|})}\right)^2
\end{aligned}
\tag{4.10}
$$

A parabola is fitted to the χ^2 values. The minimum of this parabola has by definition the best agreement between data and simulation and is therefore interpreted as the required correction value. See figure 4.20 for the visible di-tau mass before and after the correction is applied. An uncertainty of 100 % of the energy scale factory is assumed.

Since the reconstruction- and identification conditions largely vary depending on the detector region, the measurement has been done separately in the central ($|\eta|$ < 0.9), transition ($0.9 < |\eta| < 1.46$) and forward region ($|\eta| > 1.46$) of the detector. The results are summarized in table 4.5. The electrons misidentified as hadronically decaying taus in the transition region between the central and the forward detector get a correction factor of 1.09. The resulting, corrected di-tau mass distribution shows a greatly improved agreement between data and simulation.

This approach of correcting the energy scale of leptons misidentified as taus by the variation of the $p_{\mathrm{T}}^{\tau_h}$ has been picked up by the CMS group at the University Wisconsin-Madison for the Standard Model $H \to \tau\tau$ analysis. The factors have been re-derived on the full LHC 2016 dataset and are compatible to the results shown above.

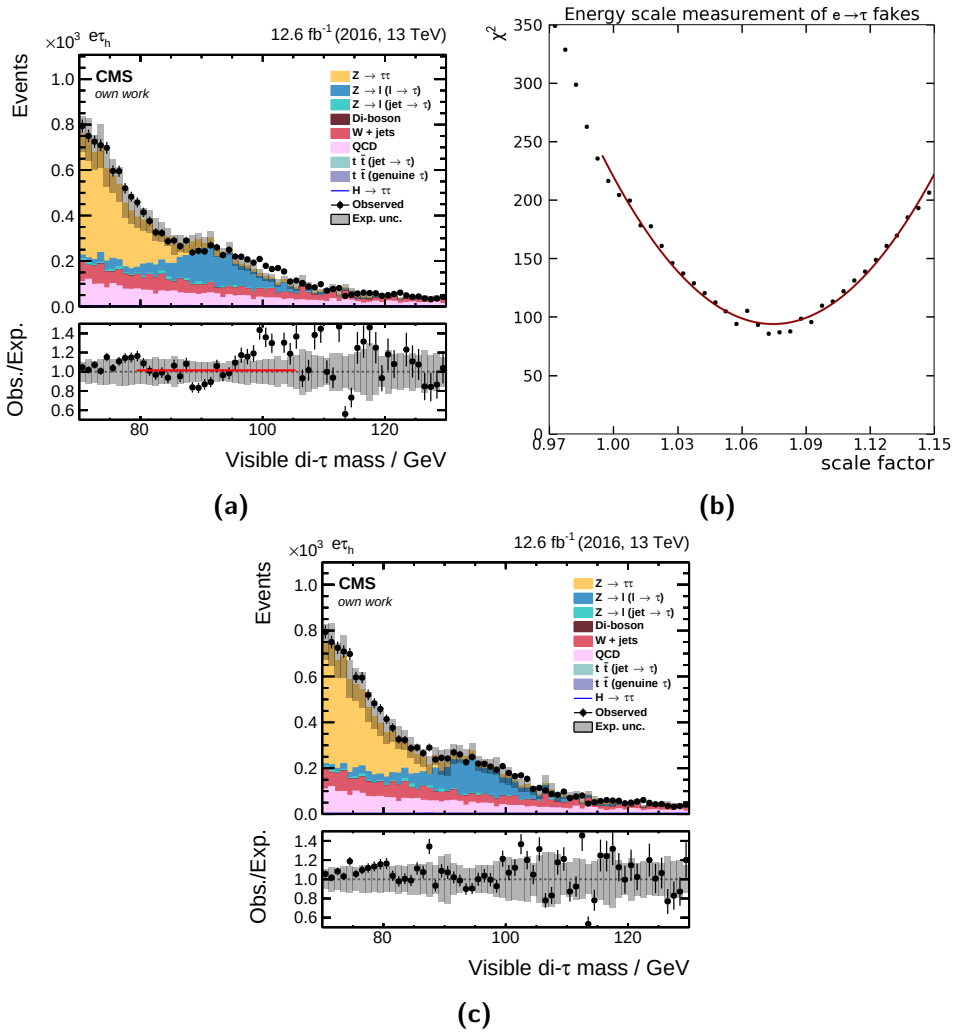


Figure 4.20: (a) The visible di-tau mass distribution in the $e\tau_h$ channel in events without jets with a transverse momentum of 30 GeV/c before the correction. (b) The χ^2 value distribution with the fitted parabola and a minimum at 1.07. (c) The $m_{\tau\tau}^{vis}$ distribution after the scale factor is applied.

4.4.8 Top p_{T} reweighting

A kinematic reweighting is applied on $\mathrm{t\bar{t}}$ events to improve the data/MC agreement. A data/simulation disagreement from the measured top quark transverse momentum to the observed one has been found [77], see Figure 4.21.

The resulting scale factor depending on the top transverse momentum is

$$SF(p_{\mathrm{T}}) = e^{0.0615 - 0.0005 \cdot p_{\mathrm{T}}/\mathrm{GeV}} \tag{4.11}$$

with an overall factor w

$$w = \sqrt{SF(t) \cdot SF(\bar{t})} \tag{4.12}$$

The systematic uncertainty corresponding to this reweighting is accounted for by a systematic shift where it is not applied and one where it is applied twice. It is correlated among all final states.

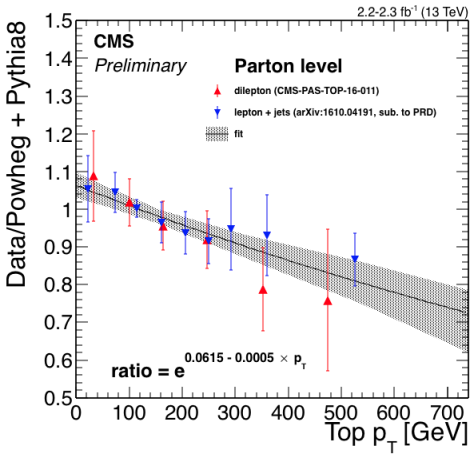


Figure 4.21: The ratio between 13 TeV data from Run II 2015 and the POWHEG + PYTHIA 8 leading order simulation for the differential top quark pair cross section as a function of the top p_{T}. The simulation underestimates the data. The difference was parametrized and used in this analysis to correct the top p_{T} spectrum [77].

4.4.9 Z Boson leading order reweighting

The Z boson **L**eading **O**rder reweighting is applied to all Drell-Yan events and is depending on the generator-level transverse momentum and mass of the di-lepton system. It has been derived on $Z \to \mu\mu$ decays and corrects the simulation. The correction factors tend to be higher with increasing di-lepton system boost. It has been verified that the kinematic reweighting estimates effects of higher orders. The uncertainty to this correction is assumed to be 100 % of the effect and incorporated in a shape uncertainty where for one systematic shift the correction is applied twice and for the other not at all. It is correlated among all channels and categories.

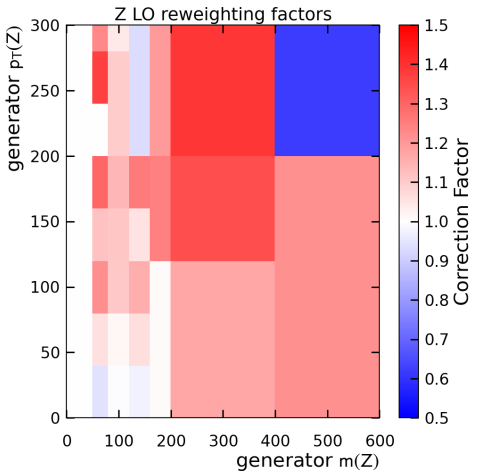


Figure 4.22: The Z leading-order correction factors, dependent on the generated boson mass and transverse momentum.

4.5 Discussion of the systematic uncertainties

The uncertainty model allows for technical reasons two kinds of uncertainties that are quickly explained in the following.

1. **Normalization** uncertainties only affect the normalization of one or several processes. A typical example is the uncertainty on the luminosity. It is applied on all backgrounds where the yield estimation is not extracted in a sideband region. It is correlated among all channels and final states.

2. **Shape** altering uncertainties are variations of an underlying variable based on which the full di-tau system reconstruction is run afterwards. There is always one systematic up- and one down-shift, corresponding to the variation of the modified variable. An example is the uncertainty on the τ_h energy scale of $\pm 3\,\%$.The di-tau mass distribution or $m_{\tau\tau}^{SVFit}$ *shape* potentially changes by the application of this shift, giving this type of uncertainty its name. The normalization can also correspondingly change.

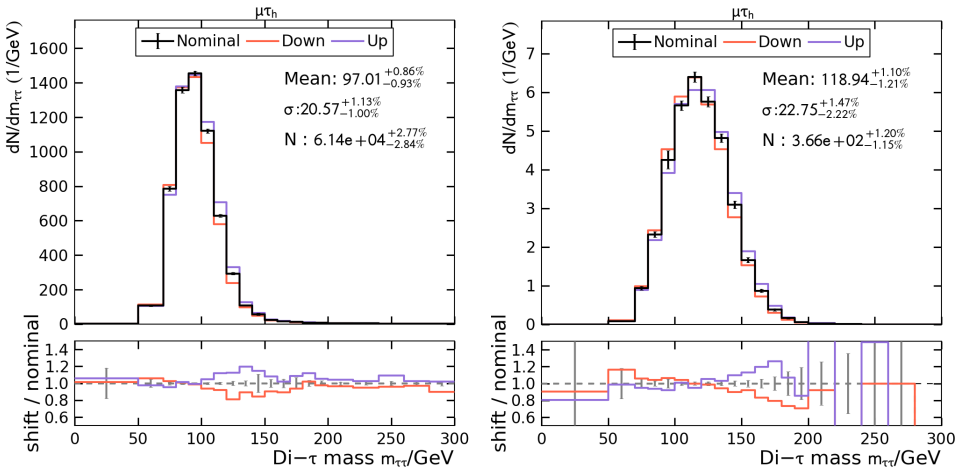
The impact of the uncertainties on the signal strength measurement varies a lot. Only a handful of uncertainties do really have a significant impact on the signal strength measurement.

See figure 4.23 for the comparison of the tau energy scale variation in the $\mu\tau_h$ channel on the $Z \to \tau\tau$ process and the $gg \to H$ process. The number of expected events at nominal mass values of the Z or Higgs boson changes only in the order of a few percent. The normalization change of $3\,\%$ in the $Z \to \tau\tau$ process happens mainly in the mass region above $m_{\tau\tau}^{SVFit} > 100\,\mathrm{GeV}/c^2$. This is exactly where the Higgs boson signal is expected and is the reason why the hadronic tau energy scale is one of the most important uncertainties for the Standard Model $H \to \tau\tau$ analysis. This illustrates that keeping signal-free sideband regions with reconstructed di-tau masses above and below the Higgs boson resonance makes sense to constrain the hadronic tau energy scale.

The uncertainty on the jet energy scale has the highest impact on the migration of events between categories since the number of jets above $p_\mathrm{T} > 30\,\mathrm{GeV}/c$ depends on it.

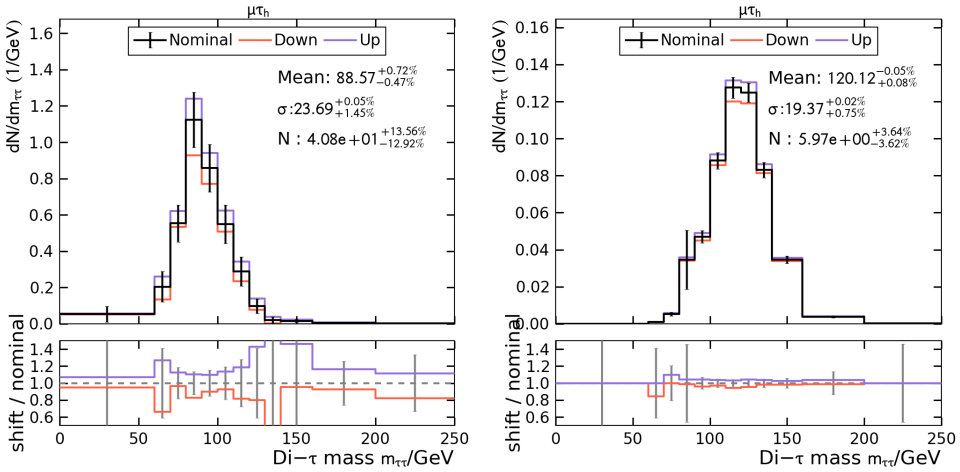
The impact of the jet energy scale variation in a special VBF signal region can be seen in figure 4.24. Comparable to the hadronic tau energy scale variation, the migration of events due to the jet energy scale variation changes the normalization of the $Z \to \tau\tau$ background significantly especially in the mass region of the Higgs boson.

Table 4.6 summarizes the normalization uncertainties that have been introduced in the previous sections and that have been taken from literature or the corresponding preliminary physics objects group recommendations.



(a) $Z \to \tau\tau$ **(b)** $gg \to H$

Figure 4.23: Comparison of the hadronic tau energy scale variation in the $\mu\tau_h$ channel on the a $Z \to \tau\tau$ process and the b $gg \to H$ process.



(a) $Z \to \tau\tau$ **(b)** $qq \to H$

Figure 4.24: The effect of the jet energy scale variation on the (a) $Z \to \tau\tau$ di-tau mass distribution and the (b) $gg \to H$ distribution in a event selection requiring at least two jets with $\Delta\eta_{jj} > 4.5$ and $p_{\mathrm{T}}^{\mathrm{H}} > 100\,\mathrm{GeV}/c$. While the number of expected $qq \to H$ events changes only by 3.6 %, the change in the number of expected $Z \to \tau\tau$ events in the di-tau mass region between $110\,\mathrm{GeV}/c^2$ and $130\,\mathrm{GeV}/c^2$ changes by 40 %. This means that a precise determination of the jet energy is crucial to distinguish $qq \to H$ signal from a upward fluctuation of the jet energy scale.

	$\tau_h\tau_h$	$\mu\tau_h$	$e\tau_h$
luminosity	2.5	2.5	2.5
$t\bar{t}$ cross section	6.0	6.0	6.0
$\mathrm{Di-Boson}$ cross section	10.0	10.0	10.0
$W+\mathrm{Jets}$ cross section	4.0		
Z/γ^* cross section	4.0	4.0	4.0
ggH PDF inclusive	3.2	3.2	3.2
VBF PDF inclusive	2.1	2.1	2.1
ggH QCD inclusive	3.9	3.9	3.9
VBF H QCD inclusive	0.4	0.4	0.4
e id efficiency			2.0
μ id efficiency		2.0	
τ id efficiency (correlated)	16.0	8.0	8.0
τ id efficiency (uncorrelated)	10.0	4.0	4.0

Table 4.6: Summary of the systematic uncertainties (in percent) on the production process and on the electron, muon and tau identification efficiencies.

The normalization uncertainties derived in the context of this thesis can be found in table 4.7.

4.6 Event Categorization

The categorization of events serves two purposes. One is to focus on phase-space regions where the signal is enriched compared to the backgrounds. The approach to categorize events instead of rejecting them has the advantage that control regions can still be used. These control regions allow constraining nuisance parameters and therefore effectively reduce the uncertainty in the signal regions. The other purpose of categorization is to improve the mass resolution of the discriminating variable, $m_{\tau\tau}^{SVFit}$. In boosted topologies, the relative resolution of the $\not{E}_{\mathrm{T}}$ and p_{T} of the reconstructed taus improve and therefore allows also a di-tau mass reconstruction with improved resolution. This is especially important to separate the Higgs Boson signal from the $Z \to \tau\tau$ background.

The basic categorization of events is in jet multiplicity, i.e. that there are events with zero, one or two or more jets with a transverse momentum of $p_{\mathrm{T}} > 30\,\mathrm{GeV}/c$. In the following, the motivation for this choice is explained and how the final definition

	$\tau_h\tau_h$	$\mu\tau_h$	$e\tau_h$
W estimation		2.8 - 7.8	3.6 - 9.7
QCD estimation	3.1 - 72.0	0.2 - 3.5	0.4 - 5.5
OS/SS extrapolation factor		16.4 - 21.6	16.8 - 41.4
VBF H α_s variation	0.5 - 1.8	0.1 - 0.4	0.3 - 1.6
ggH α_s variation	0.1 - 0.7	0.1 - 0.6	0.1 - 1.2
VBF H PDF variation	0.6 - 2.2	0.2 - 0.8	0.4 - 2.1
ggH PDF variation	0.1 - 1.2	0.1 - 0.9	0.1 - 1.2
VBF H scale	0.2 - 1.2	0.4 - 1.2	0.2 - 1.5
ggH scale	0.5 - 12.2	0.5 - 12.0	0.5 - 13.0

Table 4.7: Summary of the systematic uncertainties (in percent) derived in the context of this thesis. When ranges are given, the smallest and the highest of the corresponding uncertainty is given.

of the categories has been made. This determination is fully based on an **Asimov dataset**[2] , meaning the observation has been replaced by the prediction from the simulation. That way it is made sure that one does not optimize the analysis towards statistical fluctuations in the dataset.

4.6.1 0 jet category

In the semileptonic channels, the $Z \to l(l \to \tau_h)$ background is higher in the 0 jet category than in the ones with jets. The uncertainty on the misidentification-rate of electrons and muons being misidentified as hadronic taus is up to 34 % (see table 4.4) and the $Z \to l(l \to \tau_h)$ background has an expected di-tau mass distribution very close to the expected Higgs boson signal, the measurement of the Higgs boson signal is not very significant.

In the $\tau_h\tau_h$ channel, the QCD multijet background is overwhelmingly large compared to the expected SM Higgs boson signal.

The requirement of no reconstructed jets with transverse momentum above 30 GeV/c already limits the expectation of a boost of the di-tau system. A boosted di-tau system needs some kind of hadronic recoil - and the boost is an important requirement for a good mass resolution. So even though most $H \to \tau\tau$ events end up in the 0 jet category (see table A.4), they are indistinguishable from the large number of background events.

[2]The term *Asimov dataset* refers to the story *Franchise* by Isaac Asimov where a democratic election is replaced by an interview of the most representative voter

Therefore, no further categorization has been made for events with jet multiplicity Zero. This category is mainly important as control region to constrain the uncertainties correlated to the one- and two-jet categories.

4.6.2 1 jet categories

The 1 jet category allows best the measurement of the $gg \to H$ production process in association with a jet from initial state radiation.

The estimated Higgs Boson transverse momentum $p_{\mathrm{T}}^{\mathrm{H}}$, defined as

$$p_{\mathrm{T}}^{\mathrm{H}} = (\vec{p}_{\tau_1} + \vec{p}_{\tau_2} + \vec{\not{E}}_{\mathrm{T}})_{\mathrm{T}} \tag{4.13}$$

is used for background suppression, see figure 4.25a. While for transverse momenta of $10\,\mathrm{GeV}/c$ to $30\,\mathrm{GeV}/c$ only about one out of 200 events is expected to be a Higgs boson, above $150\,\mathrm{GeV}/c^2$, there is about one Higgs boson event per 50 background events.

An improved mass resolution in events with boosted hadronic taus allows a better separation of Higgs boson events from the $Z \to \tau\tau$ and W + Jets background, see figure 4.25b. The overlap of the $m_{\tau\tau}^{SVFit}$ distributions of the $Z \to \tau\tau$ and $H \to \tau\tau$ processes are greatly reduced with a higher boost of the hadronic tau. The W + Jets distribution gets so broad that its overlap with the $H \to \tau\tau$ signal also decreases. Only the $Z \to ll$ background has an overlay to the signal over the whole transverse momentum range. This is the reason that a precise measurement of the lepton to τ_h and jet to τ_h fake rate is important.

The final choice of the $p_{\mathrm{T}}^{\tau_h}$ and $p_{\mathrm{T}}^{\mathrm{H}}$ cut values has been derived using a two-dimensional scan, meaning that several hundred possible combinations of cut values have been evaluated. Each channel has been optimized on its own. The events with jet multiplicity 0 got their own category, as well as the events with 2 or more jets. The events with one reconstructed jet got split up into a *high* and a *low* category. The categorization scheme can be found in figure 4.26.

The range of the scan is limited by the possibility to cover a very tight selection with a sufficient number of simulated events. For low-populated bins on the one hand the bin-by-bin uncertainties rise, hence decreasing the expected significance. On the other hand, poorly populated background templates that are by definition the same in the Asimov dataset can lead to unnaturally high significances. The expected significances in the $p_{\mathrm{T}}^{\tau_h}$ range of 0 to $50\,\mathrm{GeV}/c$ and $p_{\mathrm{T}}^{\mathrm{H}}$ from 0 to $80\,\mathrm{GeV}/c$ can be found in figure 4.27.

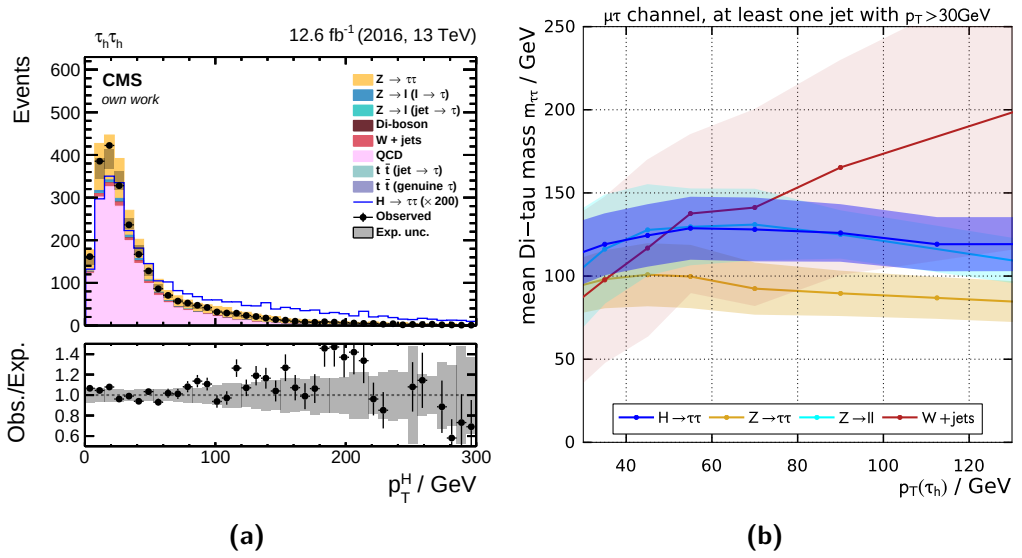


Figure 4.25: (a) Distribution of $p_{\mathrm{T}}^{\mathrm{H}}$ in the $\tau_h\tau_h$ channel. Higgs bosons tend to have a higher boost, leading so a signal enrichment in the high-$p_{\mathrm{T}}^{\mathrm{H}}$ region. Figure (b) shows the mean values of $m_{\tau\tau}^{SVFit}$ as a function of $p_{\mathrm{T}}^{\tau_h}$ in the $\mu\tau_h$ channel of events with at least one jet. The error bands represent the standard deviation of the $m_{\tau\tau}^{SVFit}$ distribution. The mass distributions do not follow a gaussian distribution. The method is anyway suitable to show the effect of the improved mass resolution over the $p_{\mathrm{T}}^{\tau_h}$. The overlap of the $m_{\tau\tau}^{SVFit}$ distributions of the $Z \to \tau\tau$ and $H \to \tau\tau$ processes are greatly reduced with a $p_{\mathrm{T}}^{\tau_h}$. The W + Jets distribution gets so broad that its overlap with the $H \to \tau\tau$ signal also decreases.

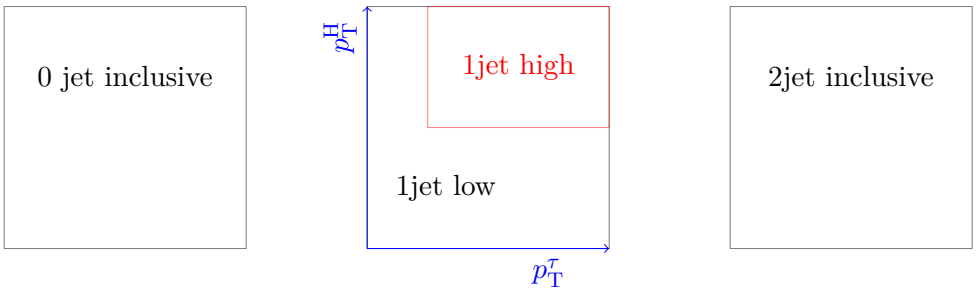
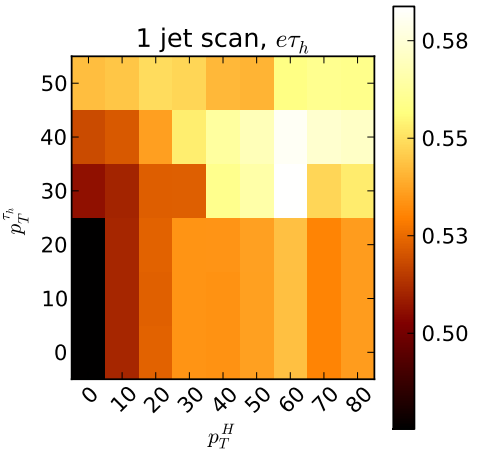
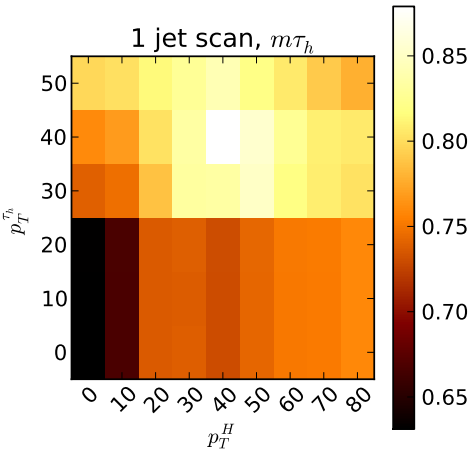


Figure 4.26: Categorization for the search of the optimal $p_{\mathrm{T}}^{\mathrm{H}}$ and p_{T}^{τ} cut values. Only events with exactly one reconstructed jet above $p_{\mathrm{T}} > 30\,\mathrm{GeV}/c$ are categorized differently depending on the Higgs- and hadronic tau transverse momentum.

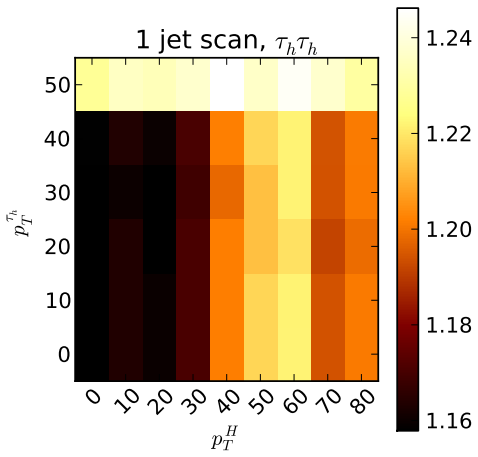
Figure 4.27: Significance scans in the 1-jet categories



(a) $e\tau_h$



(b) $\mu\tau_h$



(c) $\tau_h\tau_h$

The expected significance of a potential SM Higgs boson signal, coded in the color scales. Shown is the expected signal significances in the (a) $e\tau_h$ channel, (b) $\mu\tau_h$ channel and (c) $\tau_h\tau_h$ channels under the variation of the $p_{\mathrm{T}}^{\mathrm{H}}$ and $p_{\mathrm{T}}^{\tau_h}$ cuts. While in the semileptonic channels the optimum is already reached within the scan range, in the $\tau_h\tau_h$ channel it is on the border of $p_{\mathrm{T}}^{\tau_h} = 50\,\mathrm{GeV}/c$ above which the background templates start to be poorly populated.

4.6.3 2 jet category

The VBF Higgs Boson production mode is associated with two reconstructed jets in the final state. These jets can be used as additional tag of events to largely suppress the backgrounds. The $t\bar{t}$ background may cause a comparable topology where the reconstructed di-tau system is accompanied by two or more jets at leading order. The Di-Boson background with its many possible combinations of Z and W decays also is there, but thanks to its low cross section not dominant. The $Z \to l(l \to \tau_h)$ and $Z \to l(\text{jet} \to \tau_h)$ backgrounds are nearly negligible, since they are rarely accompanied by a second jet. The $W+\text{Jets}$ production cross section for a W in association with three jets (one misidentified as the τ_h, the other two as the tagging jets) is even still one order of magnitude higher than the VBF Higgs Boson production cross section, leaving $W+\text{Jets}$ as an important source of background. The $Z \to \tau\tau$ background also contributes with two real taus, one jet e.g. from initial state radiation and one pile-up jet.

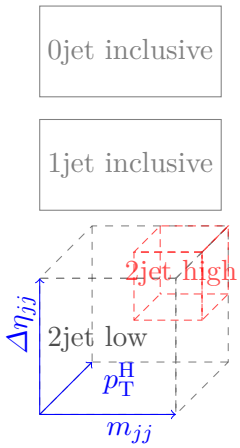
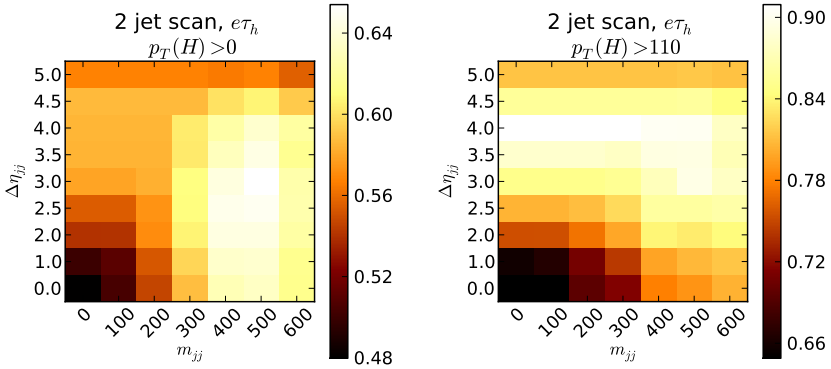
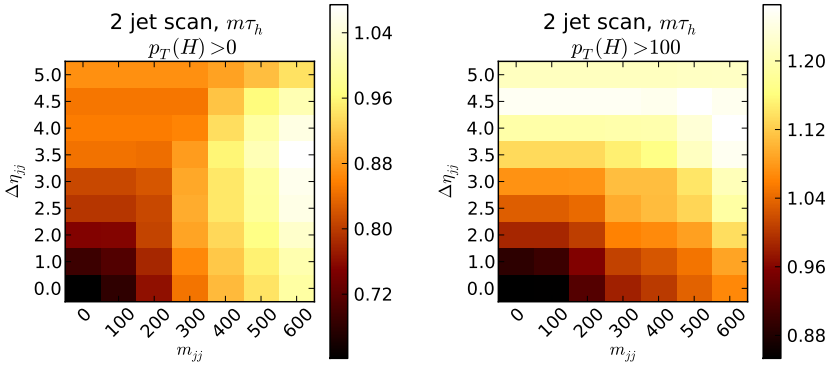


Figure 4.29: Categorization scheme for the determination of the m_{jj}, $\Delta\eta_{jj}$ and $p_{\mathrm{T}}^{\mathrm{H}}$ values leading to the highest signal significance.

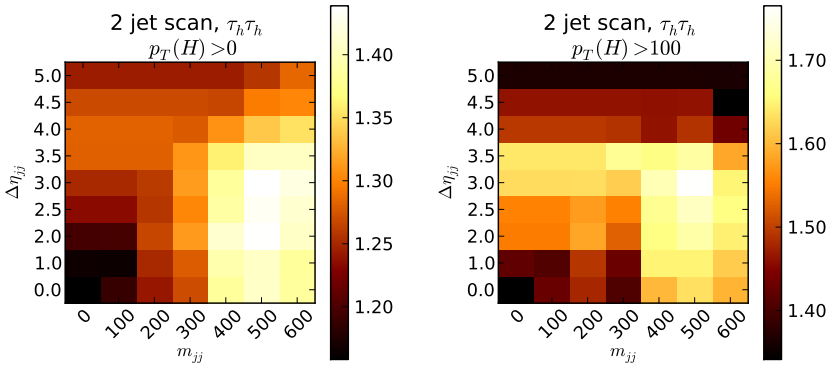
Not only the occurrence of jets but also the jet kinematics are exploited to form signal-enriched categories. The two variables of interest are the difference in pseudorapidity of the two jets $\Delta\eta_{jj}$ and the mass of the di-jet system m_{jj}. Requiring a boosted topology also improves mass resolution and significantly reduces the background expectation. A three-dimensional cut search has been performed in each channel, sorting the events in a *high* category when passing the $\Delta\eta_{jj}$, m_{jj} and $p_{\mathrm{T}}^{\mathrm{H}}$ cuts and a *low* category. The distributions of the expected sensitivity can be found in 4.30.



(a) $e\tau_h$ with no $p_{\mathrm{T}}^{\mathrm{H}}$ cut (left) and $p_{\mathrm{T}}^{\mathrm{H}} > 110\,\mathrm{GeV}/c$



(b) $\mu\tau_h$ with no $p_{\mathrm{T}}^{\mathrm{H}}$ cut (left) and $p_{\mathrm{T}}^{\mathrm{H}} > 100\,\mathrm{GeV}/c$



(c) $\tau_h\tau_h$ with no $p_{\mathrm{T}}^{\mathrm{H}}$ cut (left) and $p_{\mathrm{T}}^{\mathrm{H}} > 100\,\mathrm{GeV}/c$

Figure 4.30: The scan result distributions for no $p_{\mathrm{T}}^{\mathrm{H}}$ cut (left) and the $p_{\mathrm{T}}^{\mathrm{H}}$ values with the highest significances (right). For boosted topologies in the semileptonic channels, no cut on m_{jj} is necessary.

4.6.4 Final categorization

The final categorization is summarized in table 4.8. It consists of a 0 jet category, a 1jet low, 1jet high, 2jet low and 2jet high category in each of the three final states, thereby 15 categories in total. A distributions of expected and observed events can be found for one category in 4.32 and in the appendix in section A.4.2.

Table 4.8: Determined, optimized cut values to define the five event categories in each channel. The *high* categories are above all the cut values listed in the table, the *low* categories correspondingly under at least one of them.

		$e\tau_h$	$\mu\tau_h$	$\tau_h\tau_h$
0jet				
1 jet	$p_{\mathrm{T}}^{\mathrm{H}}$ (GeV/c)	30	40	40
	$p_{\mathrm{T}}^{\tau_h}$ (GeV/c)	60	40	60
2 jet	$\Delta\eta_{jj}$	4.5	4.5	3.0
	m_{jj} (GeV/c)	0	0	500
	$p_{\mathrm{T}}^{\mathrm{H}}$ (GeV/c)	110	110	100

4.7 Statistical inference

Everything shown up to now has was a *blinded* analysis following the CMS convention introduced in section 2.2. The penalty parameter in formula 2.5 has been chosen to be $\epsilon = 0$ and the bin-wise threshold on the significance to 0.2 above which the complete bin content in the observation and expectation has manually been set to 0. From this section on, all data get unblinded.

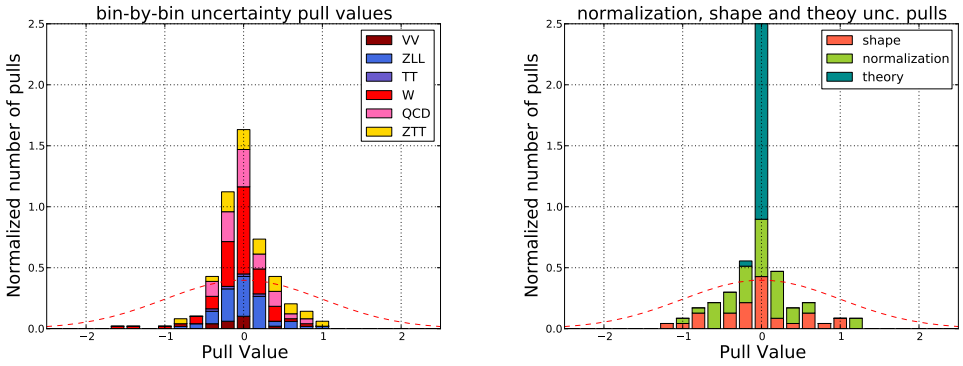
First, statistical pulls on the maximum-likelihood fit to the data on the introduced uncertainties are discussed. Then, the compatibility between the prediction and observation gets quantified using a saturated goodness-of-fit test. Finally, the signal strength and the corresponding signal significance is presented.

4.7.1 Pulls

All previously introduced uncertainties estimate on how precisely the corresponding quantities have been measured. In the form of nuisance parameters, they give the fit degrees of freedom by the variation of s_k and b_k in formula 2.24, while the fit

minimizes the negative log likelihood under the background-only and the signal plus background hypothesis. The nuisance parameters are identical to a $\pm 1\sigma$ variation.

The **pulls** are modifiers on the nuisance parameters that minimize negative log likelihood. Since the uncertainty errors are usually determined rather conservative, it is expected that the distribution of all pulls does not have a standard deviation of 1 but smaller. This is also reflected in the pull distribution shown in figure 4.31. The highest pull values, sorted by the signal plus background pull value can be found in table 4.9. The highest value is 1.23σ which is completely within the expectation of the pull distribution.



(a) bin-by-bin pulls

(b) other pulls

Figure 4.31: Visualization of the observed pull values. (a) The bin-by-bin uncertainties follow a Gaussian distribution, but with a standard deviation smaller than 1.0. There is no particular process that is pulled in an unexpected way. (b) The theory uncertainties, reflecting the uncertainties on the QCD renormalization/refactorization scale, α_s uncertainty and PDF set uncertainty are all close to 0, meaning they hard hardly constrained by the fit. The shape and normalization uncertainties are not constrained more than expected. This overall shows a rather conservative uncertainty model.

4.7.2 Goodness-of-fit test

The agreement between the observed data and prediction of the background model has been checked with a saturated model test [78]. This is a χ^2-like test, based on a likelihood ratio λ

$$\lambda = \prod_i exp\left(-\frac{(d_i - f_i)^2}{2\sigma_i^2}\right) \tag{4.14}$$

where $d_i \pm \sigma_i$ is the ith measured data point with the standard deviation σ_i and the model prediction f_i. One can convert this into the χ^2-like value χ'^2via

Nuisance parameter	s+b fit	b-only fit
$jet \to \tau$ misid. rate	1.23	1.25
Tau efficiency	1.14	0.83
$\not{E}_{\mathrm{T}}$ Recoil H/DY/W $e\tau$ 0-Jet	-1.14	-0.57
$\not{E}_{\mathrm{T}}$ Response H/DY/W $e\tau$ 1-Jet	-1.07	-1.06
$\not{E}_{\mathrm{T}}$ Response H/DY/W $\tau\tau$ 1-Jet	1.04	0.99
QCD estimation $\tau\tau$ 0-Jet	-1.0	-0.9
$\not{E}_{\mathrm{T}}$ Response H/DY/W $\tau\tau$ 0-Jet	0.91	0.78
$\not{E}_{\mathrm{T}}$ Response H/DY/W $\mu\tau$ 1-Jet	0.81	0.83
$\not{E}_{\mathrm{T}}$ Recoil H/DY/W $\mu\tau$ 0-Jet	-0.8	-0.75
$\not{E}_{\mathrm{T}}$ Response H/DY/W $\mu\tau$ 0-Jet	-0.8	-0.84
$t\bar{t}$ Cross-section	-0.77	-0.7
$\not{E}_{\mathrm{T}}$ Response H/DY/W $e\tau$ 2-Jet	-0.72	-0.71
Muon efficiency	-0.67	-0.68
Luminosity	-0.67	-0.29
Tau efficiency	0.66	0.43
$p_T(Z)$ NLO Reweighting	0.62	0.62
SS/OS factor $\mu\tau$ NoJets	-0.6	-0.59
Tau efficiency	0.58	0.69
$p_T(t)$ Reweighting	0.57	0.71
$\not{E}_{\mathrm{T}}$ Response H/DY/W $e\tau$ 0-Jet	0.54	-1.07
Tau efficiency	-0.53	-0.56
W + Jets estimation $e\tau$ 1-Jet high	-0.53	-0.43

Table 4.9: The nuisance parameters with pulls in the signal plus background fit above 0.5. A full list can be found in the appendix in section A.3

$$\chi'^2 = -2 \ln \lambda. \tag{4.15}$$

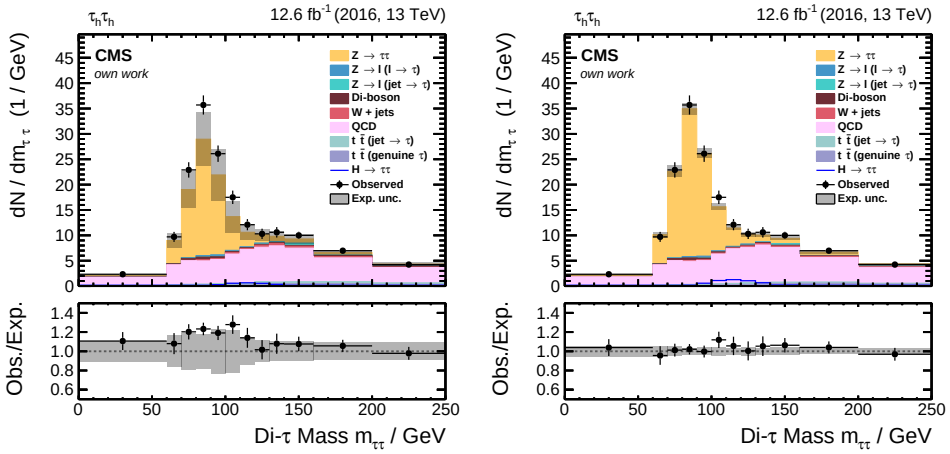
This likelihood ratio was calculated for both the observation as well as for 1000 toy experiments, that have been thrown randomly following the uncertainty model. By comparing the observed value with the expectation based on the toy experiments on can conclude on the probability p that the uncertainty model is describing the observed data. This is what has been done for individual categories, for combined channels and for the full combination. If the probability p that the data can be a result of the expectation is above 5 %, the data is considered as well understood.

Table 4.10 summarizes the results of the Goodness-of-fit tests. The lowest p-value is observed in the $\tau_h\tau_h$ 1 jet low category. An example for a good agreement is the $\tau_h\tau_h$ 2 jet low category (figure 4.32). The deviation between the observation and prediction is largely reduced by the fit, resulting in a p-value of 0.98.

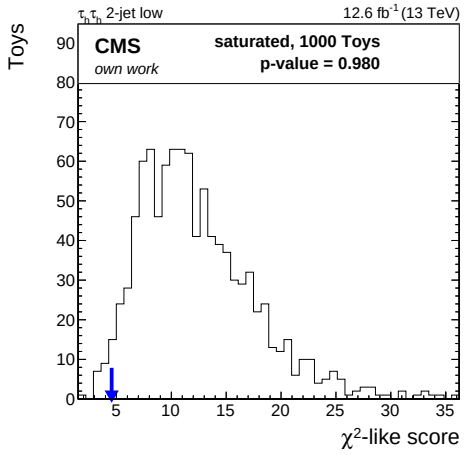
The combination of all channels and categories has a p-value of 0.16. From this one can conclude that the observed data and uncertainty model are well understood.

combined	channel	category	
		0 Jets :	0.38
		1 Jet low :	0.79
	$\mu\tau_h$: 0.33	1 Jet high :	0.82
		2 Jet low :	0.57
		2 Jet high :	0.62
		0 Jets :	0.60
		1 Jet low :	0.24
0.16	$\tau_h\tau_h$: 0.36	1 Jet high :	0.70
		2 Jet low :	0.98
		2 Jet high :	0.44
		0 Jets :	0.43
		1 Jet low :	0.98
	$e\tau_h$: 0.58	1 Jet high :	0.28
		2 Jet low :	0.64
		2 Jet high :	0.70

Table 4.10: Comparison of the p-values derived via the Goodness-of-fit test using a saturated model test. The *combined* column shows the full combination of all channels and categories. The *channel* column shows the p-values for the analysis run only on the individual channels, the *category* column correspondingly on the individual categories.



(a) prefit **(b)** s+b postfit

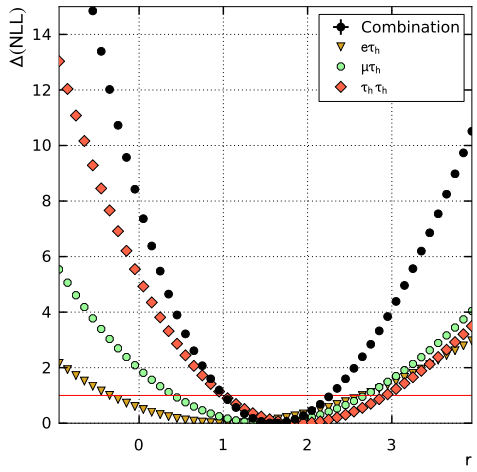


(c) Goodness-of-fit test with saturated model

Figure 4.32: $\tau_h\tau_h$ 2 jet low category comparison (a) prefit, (b) postfit and (c) the Likelihood-ratio distribution with the highest p-value of 0.98. The deviation between observed and simulated distributions is largely reduced, which is reflected by the p-value of 0.98.

4.7.3 Signal strength and signal significance

The extraction of the signal strength is performed by the global maximum likelihood fit over all categories and channels. It has been also done on the three individual channels to examine the compatibility between them. Figure 4.33 shows the $\Delta(NLL)$ scan over the signal strength modifier r with the minimum at $\hat{r} = 1.61$. The individual channels show compatible results while the combination largely improves the significance of the result. The observed excess of 61 % is only slightly above one standard deviation of 0.46 and therefore still compatible with the Standard Model signal strength prediction.



(a) Sampling of the $\Delta(NLL)$ values at different signal strength modifiers r

combination			channel			
			$\mu\tau_h$	1.50	±	0.82
1.61	±	0.46	$\tau_h\tau_h$	1.90	±	0.71
			$e\tau_h$	1.05	±	1.10

(b) Best-fit signal strengths $\hat{r}$ for the full combination and the individual channels

Figure 4.33: Scan of $\Delta(NLL)$ of different signal strengths r for the full combination and the individual channels (left) and the best-fit signal strengths $\hat{r}$ (right). While the best-fit values for $\hat{r}$ are close to each other, the width and asymmetry is different for the different distributions. While a higher signal strength modifier can be excluded from all channels, $e\tau_h$ still is compatible with $r = 0$ meaning the absence of the $H \to \tau\tau$ decay. The $\tau_h\tau_h$ channel however shows that this scenario is very unlikely.

Since the Higgs boson signal peak in the $m_{\tau\tau}^{SVFit}$ histograms is small even in the high categories and can not be seen by eye, a re-ordering of the bins helps in the visualization. To construct Figure 4.34, the bins of all 15 $m_{\tau\tau}^{SVFit}$ distributions have been sorted by their expected signal yield $s/(s+b)$. The yield represents the yield in these individual bins.

The significances for the observation of the Standard Model Higgs boson can be found in table 4.11. A significance for the existence of the $H \to \tau\tau$ decay of $2.77\,\sigma$ has been observed with an expectation of $2.51\,\sigma$. The most significant channel is

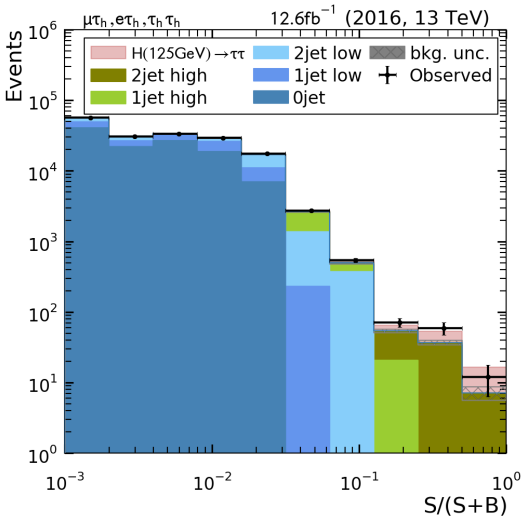


Figure 4.34: The final discriminator histograms re-ordered depending on the signal fraction $s/(s+b)$ per bin. The compatibility of the observation is higher with the existence of the $H \to \tau\tau$ decay than its non-existence. This plot includes all events of the analysis of all three channels. The leftmost bin is the overflow bin where also the events with no signal expectation at all are included.

the fully hadronic channel out of which the 2 jet high category with its expected significance of $1.55\,\sigma$ and an observed significance $1.81\,\sigma$ is the largest.

combined	channel	category	
		0 Jets :	0.24 (0.24)
		1 Jet low :	0.22 (0.19)
	$\mu\tau_h$: 1.33 (1.34)	1 Jet high :	0.44 (0.35)
		2 Jet low :	0.38 (0.37)
		2 Jet high :	0.97 (1.01)
		0 Jets :	0.36 (0.38)
		1 Jet low :	0.38 (0.34)
2.77 (2.51)	$\tau_h\tau_h$: 2.04 (1.90)	1 Jet high :	0.68 (0.56)
		2 Jet low :	0.69 (0.64)
		2 Jet high :	1.81 (1.55)
		0 Jets :	0.16 (0.18)
		1 Jet low :	0.22 (0.22)
	$e\tau_h$: 1.06 (0.97)	1 Jet high :	0.37 (0.33)
		2 Jet low :	0.32 (0.31)
		2 Jet high :	0.83 (0.80)

Table 4.11: Observed (expected) significance for the SM $H \to \tau\tau$ decay for the individual categories (right), the combination of all categories within one channel (middle) and the full combination of all channels and categories (left). The uncertainties do not add up quadratically since each of the numbers corresponds to a fit that considers the cross-correlations between the individual categories.

Chapter 5

Conclusion

With the LHC Run II a new era in high energy physics has been entered. While one of the most important achievements of Run I was the discovery of the existence of a particle compatible with the Standard Model expectation of the Higgs boson, the goals in Run II have become much more diverse. Precision measurements as well as searches for physics beyond the Standard Model are the new challenges, both having their very own requirements.

The total integrated luminosity of the 2016 Run II recorded with the CMS experiment is $40.8\,\mathrm{fb}^{-1}$. This exceeds all previous data taking periods and was only possible by the unexpectedly high performance of the LHC. How precious this 2016 dataset is should not be underrated. In the 2017 data taking period, trigger thresholds will again rise. Standard Model Higgs boson physics in Run II has already become low-energy physics just above trigger thresholds with signal acceptances in the percent level. This trend will continue with further increasing instantaneous luminosity.

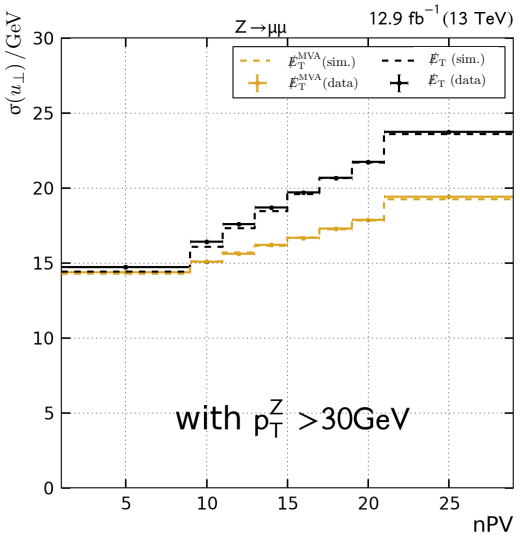


Figure 5.1: $\not\!\!E_\mathrm{T}$ resolution ($u_\perp$) as a function on the number of reconstructed primary vertices (nPV). The additional pile-up interactions degrade the $\not\!\!E_\mathrm{T}^\mathrm{PF}$ resolution while the $\not\!\!E_\mathrm{T}^\mathrm{MVA}$ can remove nearly half of this degradation effect.

With more pile-up interactions, the resolution of the $\not\!\!E_\mathrm{T}^\mathrm{PF}$ suffers. By the use of the $\not\!\!E_\mathrm{T}^\mathrm{MVA}$ presented in this thesis, the $\sqrt{nPV}$ dependent degradation in resolution with additional pile-up interactions could be reduced by 60 % (Figure 5.1). The $\not\!\!E_\mathrm{T}^\mathrm{MVA}$ reimplementation presented in this thesis has been written specifically to be future-proof and suits well in the

officially provided $\not\!\!E_{\mathrm{T}}$ tools. It comes with many improvements like the combinatoric approach to calculate the $\not\!\!E_{\mathrm{T}}^{\mathrm{MVA}}$ exactly for those leptons that are used later in the analysis and with a better description of the covariance matrix due to the inclusion of the hadronic tau recoil components.

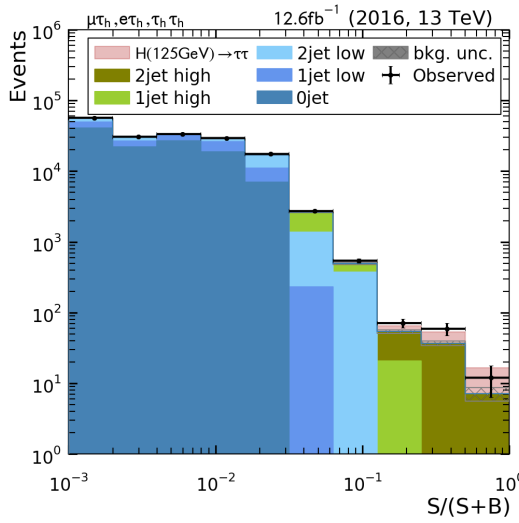


Figure 5.2: The final discriminator histograms re-ordered depending on the signal fraction $s/(s+b)$ per bin. The compatibility of the observation is higher with the existence of the $H \to \tau\tau$ decay than its non-existence.

The precise reconstruction of the $\not\!\!E_{\mathrm{T}}$ plays an important role in the sensitivity of the $H \to \tau\tau$ analysis since it is used for the W + Jets rejection as well as input for the full di-tau system reconstruction. Apart from the presented Standard Model $H \to \tau\tau$ analysis, also other analyses like the MSSM $H \to \tau\tau$ search profit from the superior $\not\!\!E_{\mathrm{T}}^{\mathrm{MVA}}$ resolution. Examples for synergies of developments done in the making of this thesis are the tool chain leading to the measurement of theory uncertainties on the Higgs Boson production cross section, which has been re-used for the $Z \to \tau\tau$ cross section measurement or the common KAPPA skims shared among several German CMS analysis groups.

The presented Standard Model $H \to \tau\tau$ analysis includes a complete set of background estimation methods out of which existing ones have been picked up and modified wherever needed. The applied existing and self-derived corrections have proved to well describe the dataset with an integrated luminosity of $12.6\,\mathrm{fb}^{-1}$. A lot of the results obtained in the context of this thesis are or will be part of official CMS publications, like the measurement of the energy scale of leptons misidentified as hadronic taus.

The first finalized CMS $H \to \tau\tau$ analysis based on Run II covering three decay channels has been presented in this thesis. The observed significance for the decay of the Standard Model Higgs boson to pairs of taus is $2.77\,\sigma$ ($2.51\,\sigma$ expected). The measured signal-strength modifier of $\hat{r} = 1.61 \pm 0.46$ is compatible with the expectation of the Standard Model and gives yet another hint on the completeness of this theory, where the fermion masses are a result of the Yukawa coupling to the Standard Model Higgs boson.

Appendix A

Appendix

A.1 Software

The analysis Software used to run the Standard Model $H \to \tau\tau$ analysis is described in the following. The analysis is split up into several, subsequent steps. Each step is associated with a special software and a data format that are explained in the following. The used computing resources are listed in the next section with an overview in Figure A.1.

A.1.1 Skimming

The CMS experiment delivers their simulated and measured data in the so-called MiniAOD (Mini Analysis Object Data) format. This format has been designed to contain enough the necessary information for most analyses with minimal overhead, while saving 90 % of file size compared to its predecessor AOD. For this analysis, the MiniAOD format was well sufficient.

The data can be read out using the **CMS S**oft**W**are framework. CMSSW is a highly modular and adaptable framework with a wide range of applications, running already at prompt reconstruction during data-taking and is also able to do a full physics analysis. Because the MiniAOD files still are way too large to handle them on common institute resources, a step called *skimming* is performed. It is done by sending computing jobs running CMSSW to a computing site within the WLCG (see section A.2.1) using the **C**MS **R**emote **A**nalysis **B**uilder (CRAB) [79] as skim job submission tool. The output files of these jobs, that have only the size of several Tera bytes, can then be stored on a local Tier-2 center or institute resources. The KIT $H \to \tau\tau$ and jet energy calibration group uses the data format of Kappa[1],(**K**arlsruhe **P**ackage for **P**hysics **A**nalysis). Kappa has a very lean data format based on root [80] and comes with a plugin for CMSSW to fill this data format. Since 2015, the author became the most active developer of Kappa.

[1] Kappa, https://github.com/KappaAnalysis/Kappa

A.1.2 Event-by-event data analysis

The Artus [81] framework is used for the event-by-event analysis of the Kappa output files. Its development started at the end of 2013 and is still ongoing. The idea behind the Artus framework is to share common analysis tasks of different analyses by a modular structure. The individual modules are configurable via configuration files in the JSON[82] (**J**ava**S**cript **O**bject **N**otation) file format. The different kinds of modules are the event provider, producers, filters and consumers.

The event provider reads in the Kappa files. It always opens exactly one event with all its associated physics objects and event metadata. Producers are there to perform specific tasks like the calculation of the di-tau mass. Filters can either store their decision in the output or cancel the further processing of the current event. The latter one significantly speeds up the analysis. The producers can store the event content in a configurable way, either as histograms or as root n-tuples.

Artus supports pipelining, meaning each event can be processed with different settings while it is read only once from disk. There are two kinds of pipelines. The first is the global pipeline, that every event has to go through. The events then are produced by local pipelines where also producers and filters are run, with different settings. That way e.g. a systematic variation of energy scales can be performed.

Since all events are treated as uncorrelated, the Artus analysis jobs can be trivially parallelized. The submission to a batch system is performed with the help of grid-control[2], a versatile batch job submission tool unifying the interface of several batch systems for optimal portability of configurations. The Artus output files are usually merged into one file per dataset, which fastens up the proceeding histogram-based analysis-step.

A.1.3 Histogram-based data analysis

This is the first step where all events converge into a single distribution. The tool developed to handle this last step is called Harry Plotter. It is a ROOT-based postprocessing-tool, covering a wide range of tasks due to its modularity. There are analysis modules that can perform fits, do calculations and even all background estimation methods are implemented as modules of Harry Plotter. Harry Plotter can either produce plots via Matplotlib[3] or a ROOT Plot interface. It can also export ROOT histograms that are further used for the statistical combination.

[2] grid-control, https://github.com/grid-control/grid-control
[3] matplotlib, http://matplotlib.org/

A.1.4 Statistical combination

The statistical combination has been performed by COMBINE[4] interfaced by the COMBINEHARVESTER[5] framework. A *datacard* containing the final discriminator histograms on both expectation and observation being produced in the last step is handed over to COMBINEHARVESTER. The statistical combination is performed here extracting the signal strength and signal significance.

A.2 Computing resources

The large bandwidth of requirements on the individual analysis steps make it absolutely necessary to have access to different computing resources fulfilling different requirements.

A.2.1 The Worldwide LHC computing grid - WLCG

The WLCG[6] is a collaboration of over 170 computing centers in 42 countries all around the earth. The computing centers are organized in tiers. While the Tier-0 and Tier-1 are mostly reserved for running jobs of the collaborations itself, users can their computing jobs to 160 Tier-2 sites, while preferably the job is run where the data is stored to minimize network traffic.

The full $H \to \tau\tau$ analysis skim consists of 9 datasets from the detector and 32 simulated datasets. Each dataset can contain several thousand individual files for each MiniAOD file exactly one job has been ran producing exactly one KAPPA output file. In total, the analysis is based on 50.000 input files with a file size of 48.74 TB, where the runtime to produce them is in the range of a few hours. The resulting KAPPA files are only 6 TB in size. They have been stored on the DESY that used to provide easy access via dCache[83].

A.2.2 NEMO and the NAF

The NEMO[7] cluster is intended for compute activities related to research in the fields **N**euroscience, **E**lementary Particle Physics and **M**icrosystems Engineering. It provides 756 worker nodes with each 20 having cores. The job submission is possible via HTCONDOR[8], a batch system developed at University of Wisconsin-Madison. A sophisticated system of virtual machines allows the execution of CMSSW and related analyses [84].

[4] COMBINE, HTTPS://GITHUB.COM/CMS-ANALYSIS/HIGGSANALYSIS-COMBINEDLIMIT
[5] COMBINEHARVESTER, HTTPS://GITHUB.COM/CMS-ANALYSIS/COMBINEHARVESTER
[6] WLCG website, http://wlcg.web.cern.ch
[7] NEMO, https://www.nemo.uni-freiburg.de/
[8] HTCondor, https://research.cs.wisc.edu

The **N**ational **A**nalysis **F**acility at DESY is set up in the framework of the Helmholtz Alliance 'Physics at the Terascale', providing computing resources to the LHC, ILC and BELLE communities. It is in operation since 2007 and has been re-designed in 2013/2014.

Both resources have been intensively utilized to produce this thesis. Especially the calculation of the reconstructed di-tau system is very time-consuming: A complete analysis run from scratch, run in parallel in over 12 thousand jobs takes in total 101 775 h or 11.6 a. The NEMO cluster reduced this to a real-time need for an analyst of less then a week. Since there is no local disk space available for KIT users at NEMO, the merging had to be done at the NAF batch system. This reduced the time for file merging from several days to about four hours.

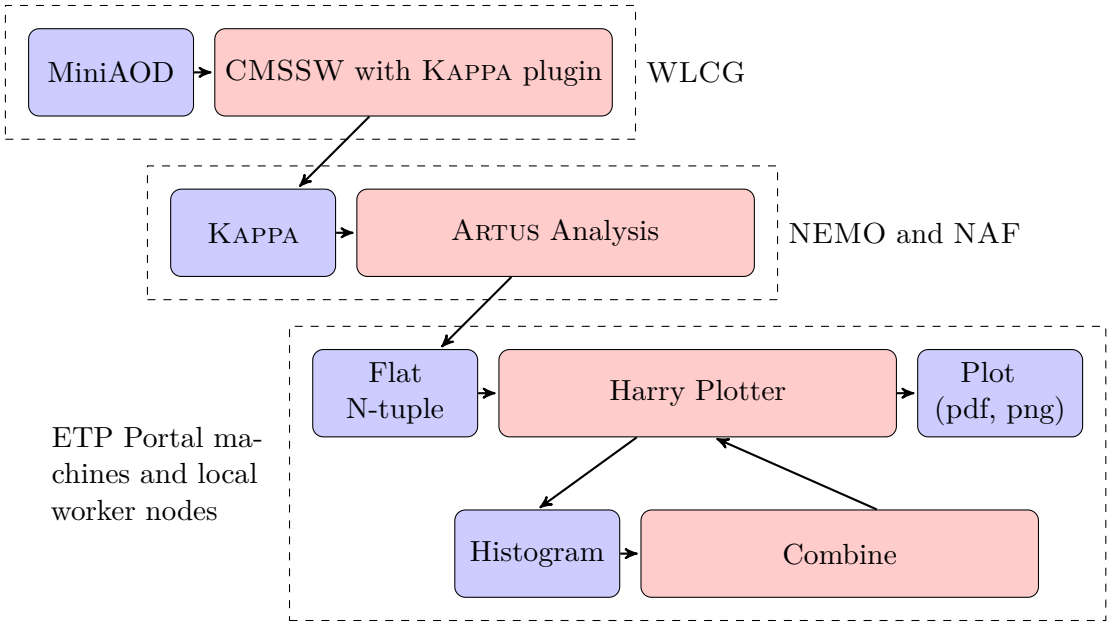
A.2.3 ETP portal machines

The development of analysis code as well as the production of plots is usually done on a so-called *portal machine*. Portal machines are shared among a few users and are usually well equipped server machines. The ETP operates several service machines. The preferred ones are equipped with 48 cores, 128 GB of main memory and several TB of SSD space. The Artus outputs have been locally stored on these SSDs for optimal performance. This setup with a locally available disk speeds up plotting by an order of magnitude compared to a file-server based solution where the network connection is the bottleneck.

A.2.4 The ETP batch system

The variety of batch system hardware like a matured batch system as well as modern server machines have been unified access through HTCondor. Since these machines have the local storage resources locally mounted, they were the only place where the parameter scans could be performed.

Figure A.1: Overview over the workflow of the $H \to \tau\tau$ analysis with the data formats (blue boxes), software (red boxes) and where they have been running (dashed boxes).



A.3 Tables

Table A.1: Datasets used in the $H \to \tau\tau$ analysis. The Run period corresponds to the data taking of the 2016 Run II.

L1 Trigger	Run Period	Number recorded of events
Muon e/γ	B	32,648,217
Muon e/γ	C	15,416,170
Muon e/γ	D	14,976,067
Single Electron	B	246,175,191
Single Electron	C	97,292,079
Single Electron	D	96,387,092
Single Muon	B	158,188,719
Single Muon	C	68,492,270
Single Muon	D	62,718,239
Tau	B	71,901,374
Tau	C	56,546,350
Tau	D	41,056,924

Table A.2: Summary of used simulated samples with their cross sections and number of generated events

process	generator	$\sigma \times BR$ (pb)	Number of events
$gg \to H$	powheg	2.78	1,498,800
$qq \to H$	powheg	0.236	1,497,468
$qq \to W^+H$	powheg	0.0039	439,400
$qq \to W^-H$	powheg	0.0061	437,580
$qq \to ZH$	powheg	0.055	597,821
$W(\to l\bar{\nu}_l) + jets$	madgraph	61526.7	28,210,360
$W(\to l\bar{\nu}_l) + 1jet$	madgraph	9644.5	39,855,520
$W(\to l\bar{\nu}_l) + 2jet$	madgraph	3144.5	29,984,239
$W(\to l\bar{\nu}_l) + 3jet$	madgraph	954.8	19,869,053
$W(\to l\bar{\nu}_l) + 4jet$	madgraph	485.6	9,174,756
$Z(\to ll)$, $(M = 50)$	madgraph	5765.4	49,877,138
$Z(\to ll)$, $(M = 10 - 50)$	madgraph	18610.0	35,079,788
$Z(\to ll)$, $(M = 150)$	madgraph	6.657	6,108,651
$Z(\to ll) + 1jet$, $(M = 50)$	madgraph	1012.5	65,485,168
$Z(\to ll) + 2jet$, $(M = 50)$	madgraph	332.8	19,695,514
$Z(\to ll) + 3jet$, $(M = 50)$	madgraph	101.8	5,753,813
$Z(\to ll) + 4jet$, $(M = 50)$	madgraph	54.8	4,101,383
Single $\bar{t}$ (t-channel)$\to l$	powheg	80.95	1,682,400
Single t (t-channel)$\to l$	powheg	136.02	3,279,200
Single $\bar{t}$	powheg	35.6	985,000
Single t	powheg	35.6	998,400
$t\bar{t}$	powheg	831.76	182,123,200
$VV \to 2l2n\bar{u}$	amcatnlo	11.95	2,944,584
$WW \to 1l1n\bar{u}2q$	amcatnlo	1.212	5,235,265
$WZ \to 1l1n\bar{u}2q$	amcatnlo	10.71	19,500,618
$WZ \to 1l3n\bar{u}$	amcatnlo	3.05	1,654,964
$WZ \to 2l2q$	amcatnlo	5.595	25,996,157
$ZZ \to 2l2q$	amcatnlo	3.22	15,498,581

Table A.3: The nuisance parameters with pulls in the signal plus background fit below 0.5 and above 0.02. The larger ones can be found in the $H \to \tau\tau$ analysis chapter in section 4.9.

Nuisance parameter	s+b fit	b-only fit
W + Jets estimation $\mu\tau$ VBF-tag	0.48	0.85
W + Jets estimation $\mu\tau$ 0-Jet	-0.46	-0.47
Electron Energy Scale	-0.42	-0.07
$\not{E}_{\mathrm{T}}$ Recoil H/DY/W $\mu\tau$ 2-Jet	-0.42	-0.39
$\not{E}_{\mathrm{T}}$ Response H/DY/W $\mu\tau$ 2-Jet	-0.41	-0.28
SS/OS factor $e\tau$ WithJets	-0.37	-0.37
W + Jets estimation $e\tau$ VBF-tag	-0.33	-0.17
$\not{E}_{\mathrm{T}}$ Response H/DY/W $\tau\tau$ 2-Jet	0.33	0.67
$e \to \tau_h$ rate	0.32	0.7
Electron efficiency	0.32	0.39
QCD estimation $\tau\tau$ 1-Jet high	-0.32	-0.11
$\not{E}_{\mathrm{T}}$ Recoil H/DY/W $\tau\tau$ 2-Jet	0.29	0.54
SS/OS factor $\mu\tau$ WithJets	0.27	0.26
$\not{E}_{\mathrm{T}}$ Recoil H/DY/W $e\tau$ 2-Jet	-0.27	-0.32
Z Boson production Cross-section	0.26	0.36
QCD estimation $\tau\tau$ 2-Jet low	0.25	0.51
Di-Boson Cross-section	-0.25	-0.18
$\not{E}_{\mathrm{T}}$ Recoil $\mathrm{t\bar{t}}$ $\mu\tau$ 2-Jet	-0.24	-0.33
QCD estimation $\tau\tau$ VBF-tag	0.23	0.87
Energy scale misreconstructed $e \to \tau$	0.23	0.08
Energy scale misreconstructed $m \to \tau$	-0.23	0.05
W + Jets estimation $e\tau$ 2-Jet low	0.21	0.14
QCD estimation $\tau\tau$ 1-Jet low	-0.2	-0.0
$\not{E}_{\mathrm{T}}$ Recoil H/DY/W $\tau\tau$ 1-Jet	-0.2	-0.18
QCD estimation $e\tau$ 1-Jet high	-0.19	-0.15
QCD estimation $e\tau$ 2-Jet low	0.19	0.12
$\mu \to \tau_h$ rate	-0.16	-0.16
QCD estimation $\mu\tau$ 2-Jet low	0.15	0.06
QCD estimation $\mu\tau$ 1-Jet low	-0.14	-0.15
W + Jets estimation $\mu\tau$ 1-Jet low	-0.14	-0.16
W + Jets estimation $e\tau$ 0-Jet	-0.13	-0.06
$\not{E}_{\mathrm{T}}$ Recoil H/DY/W $\mu\tau$ 1-Jet	-0.13	-0.1
W + Jets estimation $e\tau$ 1-Jet low	0.11	0.05
$\not{E}_{\mathrm{T}}$ Recoil $\mathrm{t\bar{t}}$ $\mu\tau$ 1-Jet	0.11	0.14
$\not{E}_{\mathrm{T}}$ Recoil $\mathrm{t\bar{t}}$ $e\tau$ 1-Jet	0.1	0.08
QCD scale 2-Jet	-0.1	0.0
QCD estimation $\mu\tau$ 0-Jet	-0.09	-0.09
$\not{E}_{\mathrm{T}}$ Recoil H/DY/W $e\tau$ 1-Jet	0.08	0.24
QCD estimation $e\tau$ 1-Jet low	0.06	0.02
W + Jets estimation $\mu\tau$ 1-Jet high	-0.06	0.09
W + Jets estimation $\mu\tau$ 2-Jet low	0.06	0.02
τ_h energy scale	-0.06	0.08
QCD scale 0-Jet	0.05	0.0
QCD estimation $e\tau$ VBF-tag	-0.04	-0.02
$\not{E}_{\mathrm{T}}$ Recoil $\mathrm{t\bar{t}}$ $\tau\tau$ 2-Jet	-0.04	-0.03
PDF unc. $qq \to H$ $e\tau$ VBF-tag	-0.04	0.0
PDF unc. $gg \to H$ $\tau\tau$ VBF-tag	0.04	0.0
SS/OS factor $e\tau$ NoJets	0.03	0.04
PDF unc. $gg \to H$ $\mu\tau$ VBF-tag	-0.03	0.0
PDF unc. $gg \to H$ $\tau\tau$ 2-Jet low	0.03	0.0
QCD estimation $\mu\tau$ VBF-tag	0.02	0.03
Jet Energy Scale	-0.02	0.07

Table A.4: Numbers of observed events and their expected composition (pre-fit) in the final categorization.

category	observed	$Z \to \tau\tau$	$Z \to l(l \to \tau_h)$	$Z \to l(\text{jet} \to \tau_h)$	W	$t\bar{t}(\text{jet} \to \tau_h)$	$t\bar{t}(t \to \tau_h)$	$\text{Di} - \text{Boson}$	QCD	$gg \to H$	$qq \to H$	$qq \to ZH$	$qq \to W^{\pm}H$
$e\tau_h$ 0 jet	29683	14116.5	2694.1	576.2	4297.5	22.3	4.9	129.5	6318.2	86.8	1.0	0.2	0.1
$e\tau_h$ 1 jet low	9034	3546.4	501.9	469.5	1487.9	166.7	66.5	208.4	2266.2	32.4	4.1	0.5	0.1
$e\tau_h$ 1 jet high	984	516.7	69.3	23.7	156.3	70.1	16.4	45.9	98.9	8.8	1.9	0.2	0.0
$e\tau_h$ 2 jet low	6268	1842.6	194.7	194.4	700.4	1157.8	893.1	327.9	788.9	25.7	7.8	1.6	0.3
$e\tau_h$ 2 jet high	91	25.8	1.9	1.5	8.4	20.6	12.4	7.1	11.8	1.5	3.9	0.0	0.0
$\mu\tau_h$ 0 jet	75725	46440.4	3716.4	1463.9	8490.8	45.9	13.9	286.0	15506.4	219.3	2.7	0.5	0.1
$\mu\tau_h$ 1jet low	18372	8793.4	498.0	559.0	2797.3	291.7	133.5	395.0	4168.5	60.3	7.4	0.9	0.2
$\mu\tau_h$ 1jet high	3205	1811.7	157.2	67.5	378.9	170.2	55.7	136.4	218.0	30.2	5.3	0.6	0.1
$\mu\tau_h$ 2jet low	12042	4271.9	190.6	174.4	1181.6	2022.6	1730.0	581.0	1655.1	54.4	17.5	3.1	0.6
$\mu\tau_h$ 2jet high	151	40.8	0.2	2.2	11.3	20.5	11.7	9.0	8.0	2.0	6.0	0.0	0.0
$\tau_h\tau_h$ 0 jet	10881	1544.2	111.1	33.6	272.5	5.0	2.2	13.2	8829.9	50.1	0.6	0.3	0.0
$\tau_h\tau_h$ 1jet low	3109	616.9	24.9	13.8	73.4	13.7	7.2	13.2	2177.5	17.2	2.3	0.3	0.1
$\tau_h\tau_h$ 1jet high	846	291.1	8.2	3.5	35.9	11.0	4.5	9.1	411.2	7.1	1.5	0.2	0.0
$\tau_h\tau_h$ 2jet low	2280	756.6	14.5	18.7	46.5	66.2	67.9	38.8	1015.3	20.8	5.3	1.3	0.2
$\tau_h\tau_h$ 2jet high	111	36.4	0.6	0.2	0.4	2.6	1.9	2.1	28.8	2.6	4.9	0.0	0.0

A.4 Control distributions

A.4.1 Kinematic variables

This chapter shows the pre-fit distribution of the imporant variables of this analysis. All error bands represent the full systematic and statistical uncertainties.

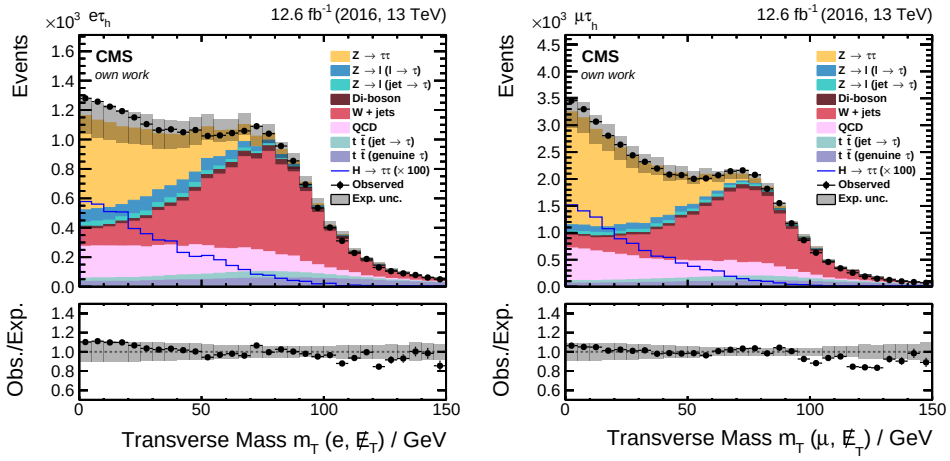


Figure A.2: Distribution of m_{T} in the $e\tau_h$ channel (left) and $\mu\tau_h$ channel (right). The transverse mass of the lepton-$\not{E}_{\mathrm{T}}$ system m_{T} has its highest value around $80\,\mathrm{GeV}/c^2$ for the $W+\mathrm{Jets}$ process.

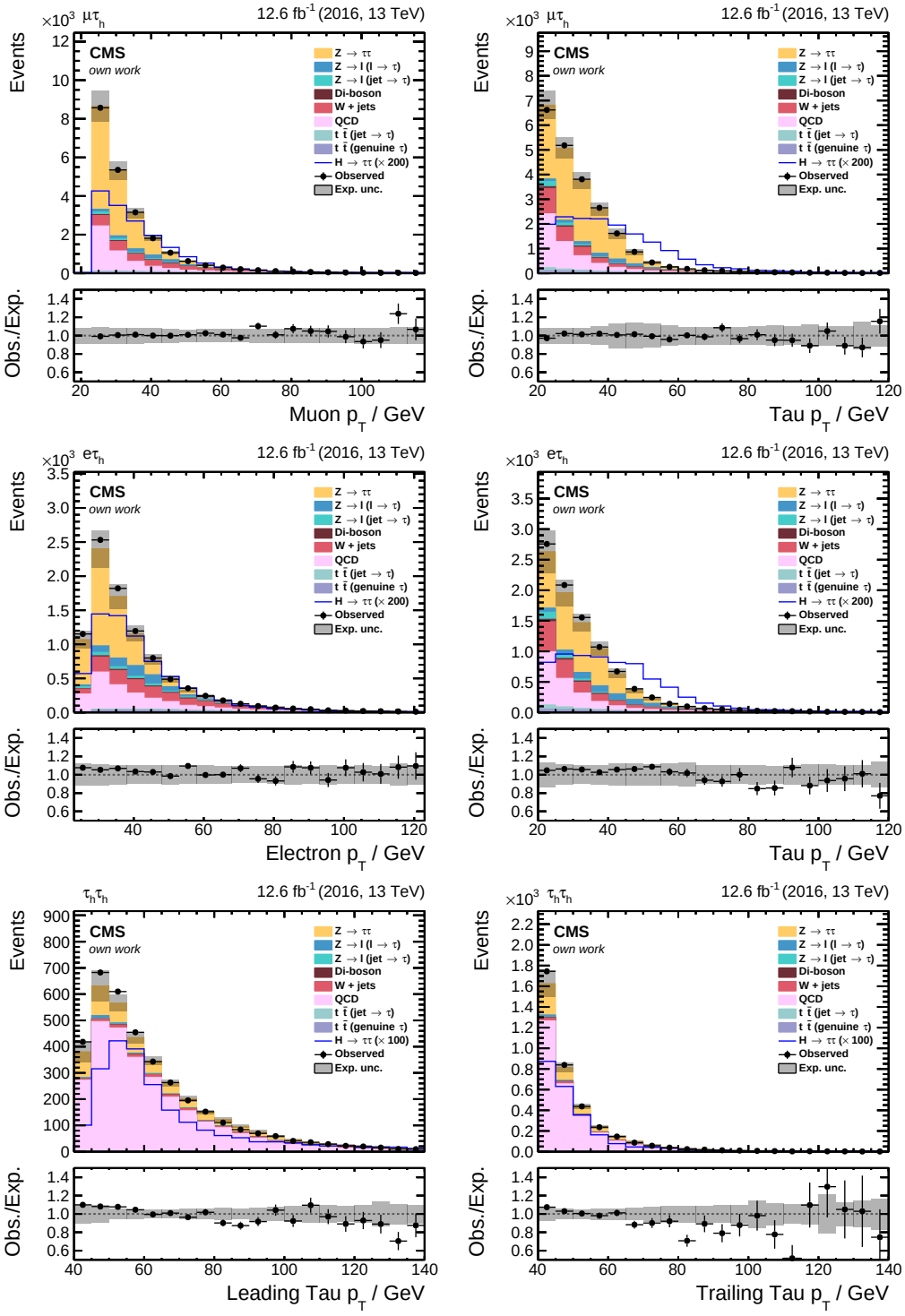


Figure A.3: Di-lepton pair transverse momenta in the $\mu\tau_h$ (first row), $e\tau_h$ (middle row) and $\tau_h\tau_h$ (bottom row) inclusive channels.

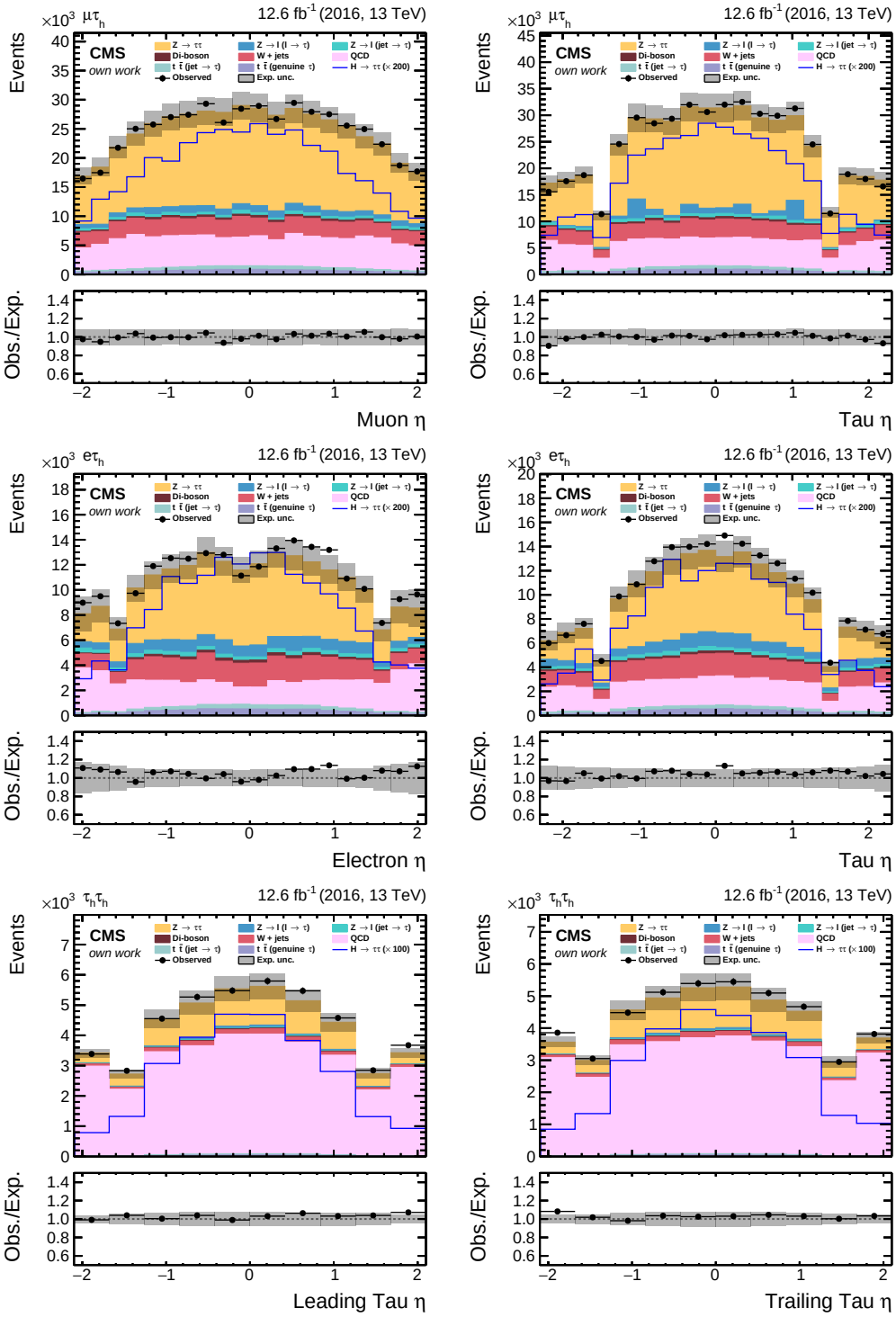


Figure A.4: Di-lepton pair pseudorapidity η in the $\mu\tau_h$ (first row), $e\tau_h$ (middle row) and $\tau_h\tau_h$ (bottom row) inclusive channels.

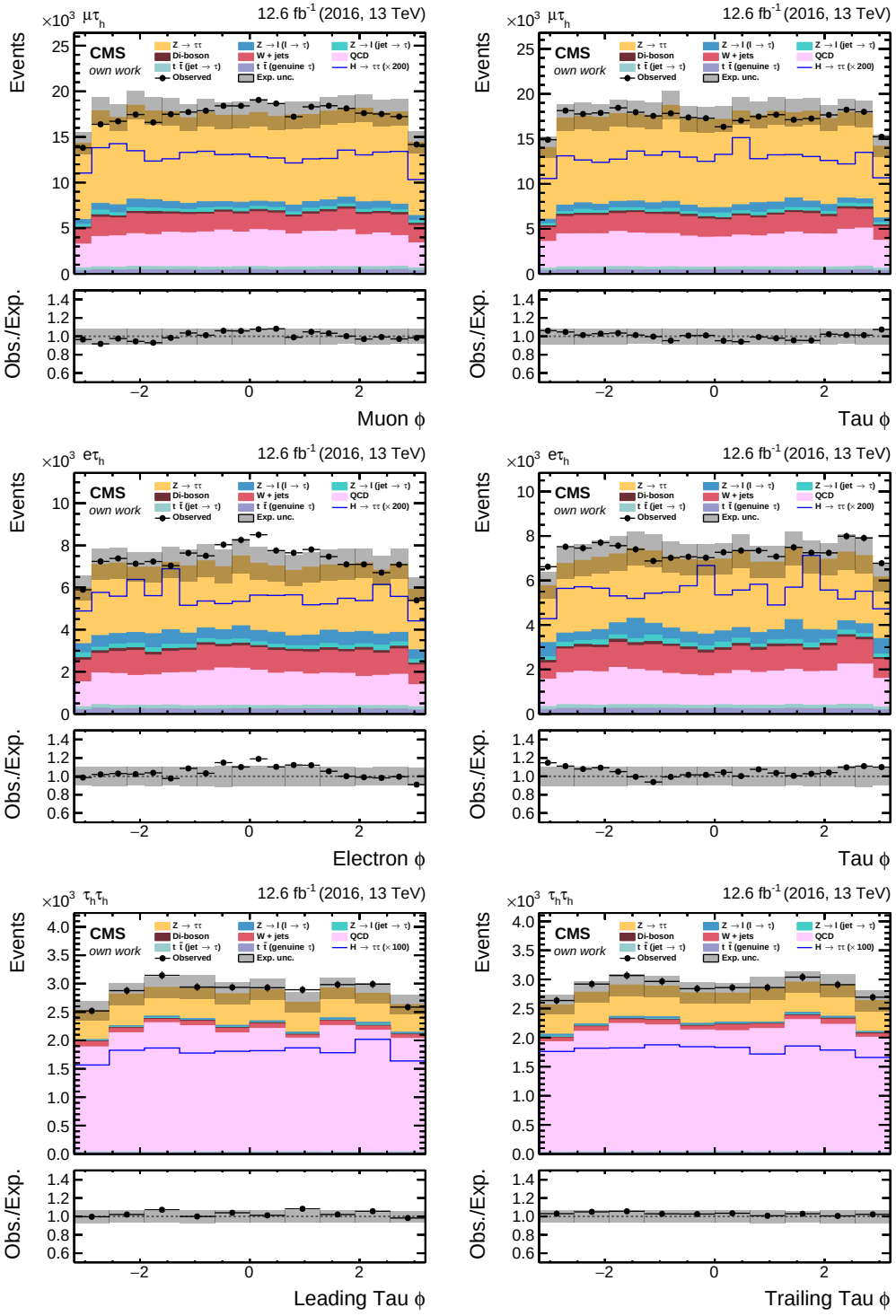


Figure A.5: Di-lepton pair ϕ in the $\mu\tau_h$ (first row), $e\tau_h$ (middle row) and $\tau_h\tau_h$ (bottom row) inclusive channels.

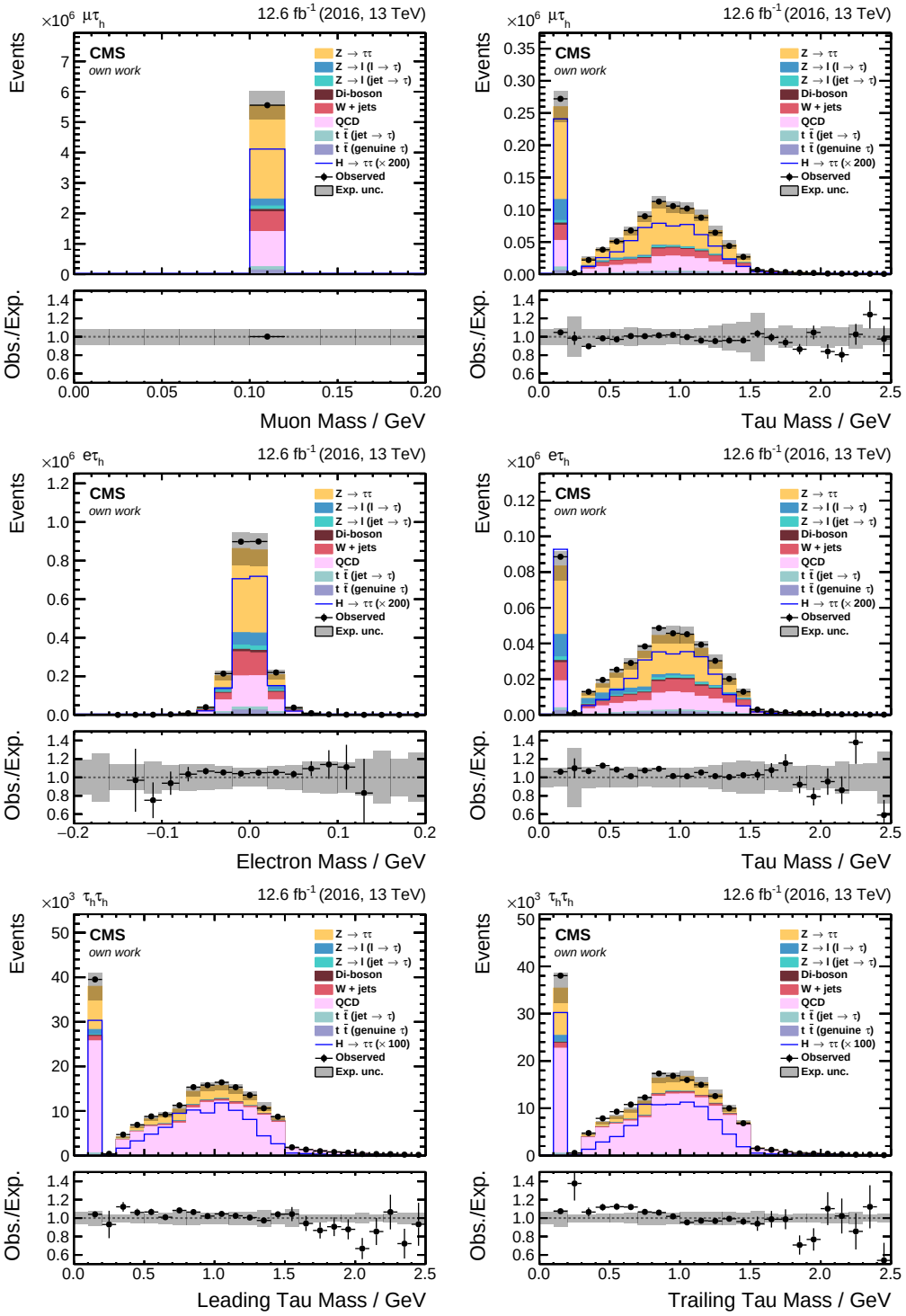


Figure A.6: Lepton masses in the $\mu\tau_h$ (first row), $e\tau_h$ (middle row) and $\tau_h\tau_h$ (bottom row) inclusive channels.

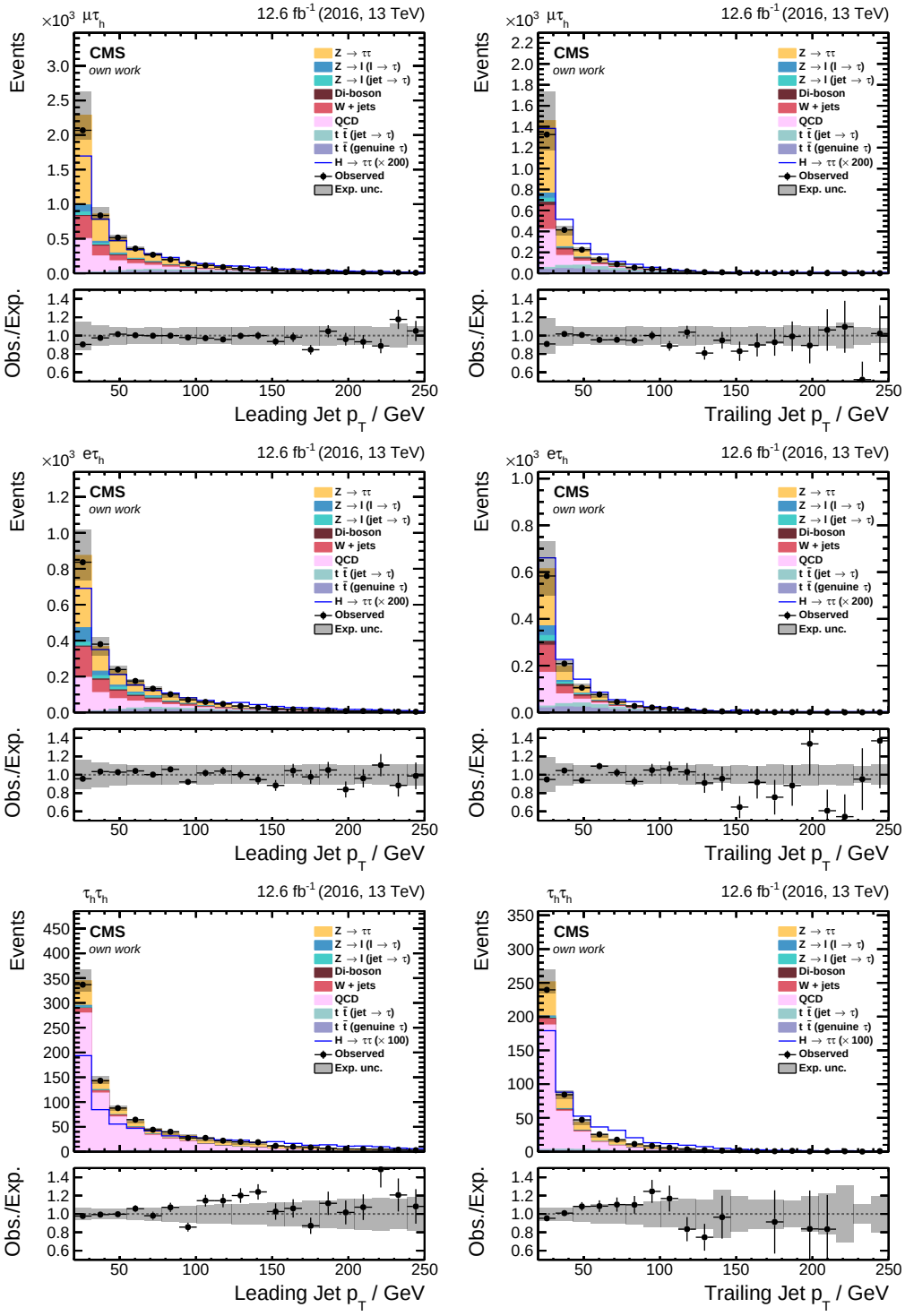


Figure A.7: Leading and trailing jet p_{T} in the $\mu\tau_h$ (first row), $e\tau_h$ (middle row) and $\tau_h\tau_h$ (bottom row) inclusive channels.

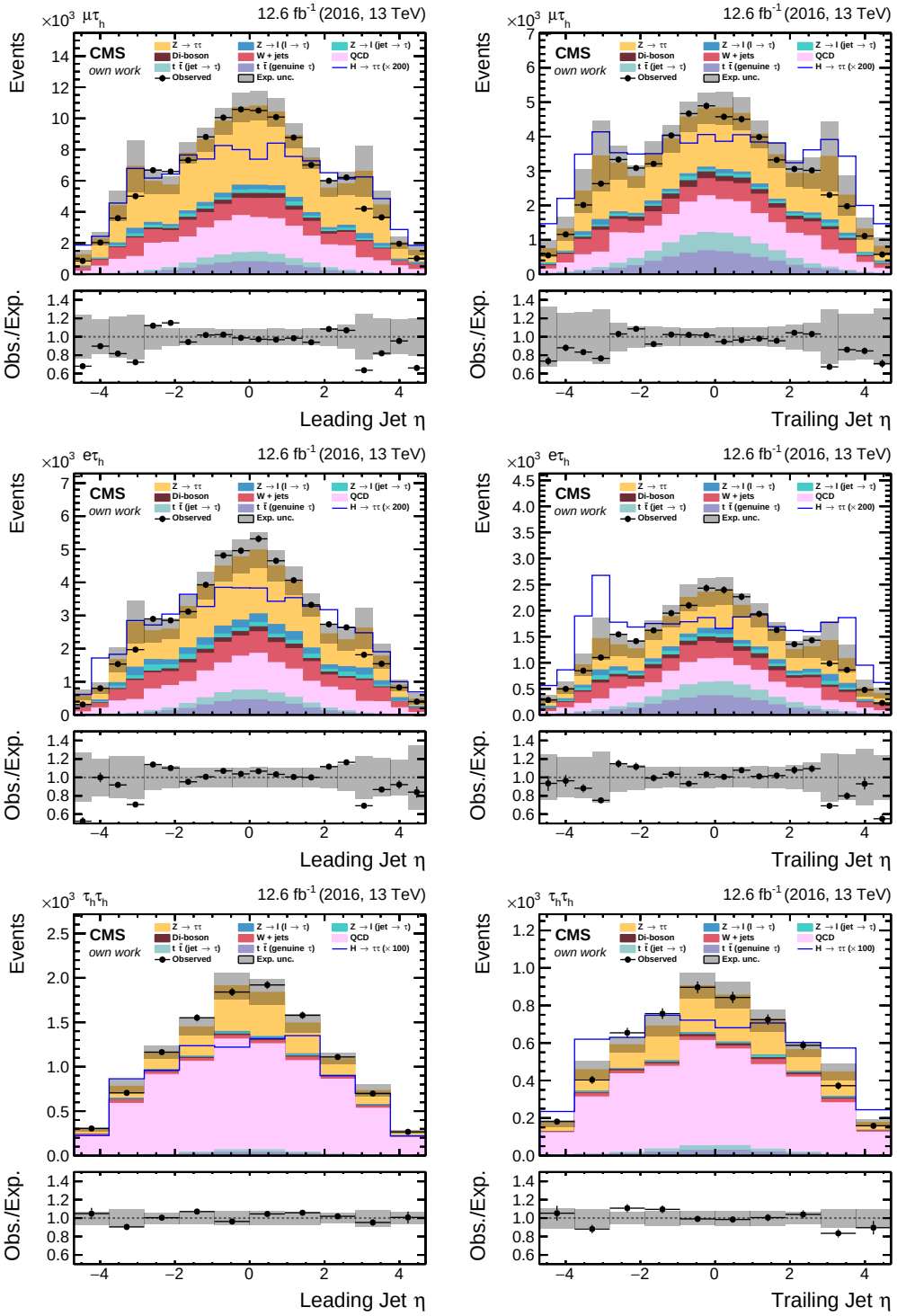


Figure A.8: Leading and trailing jet pseudorapidity η in the $\mu\tau_h$ (first row), $e\tau_h$ (middle row) and $\tau_h\tau_h$ (bottom row) inclusive channels.

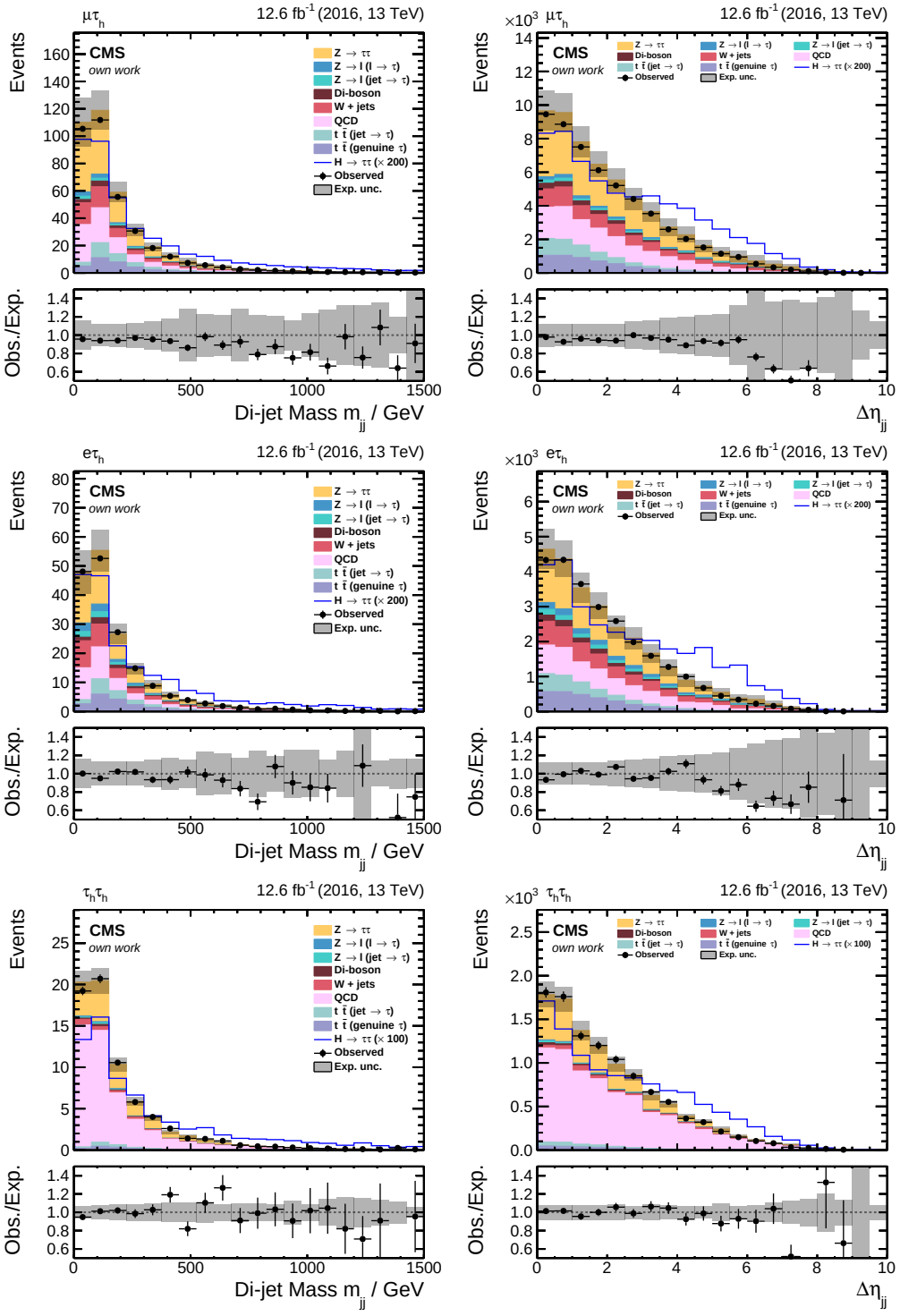


Figure A.9: Di-jet mass m_{jj} and $\Delta\eta_{jj}$ in the $\mu\tau_h$ (first row), $e\tau_h$ (middle row) and $\tau_h\tau_h$ (bottom row) inclusive channels.

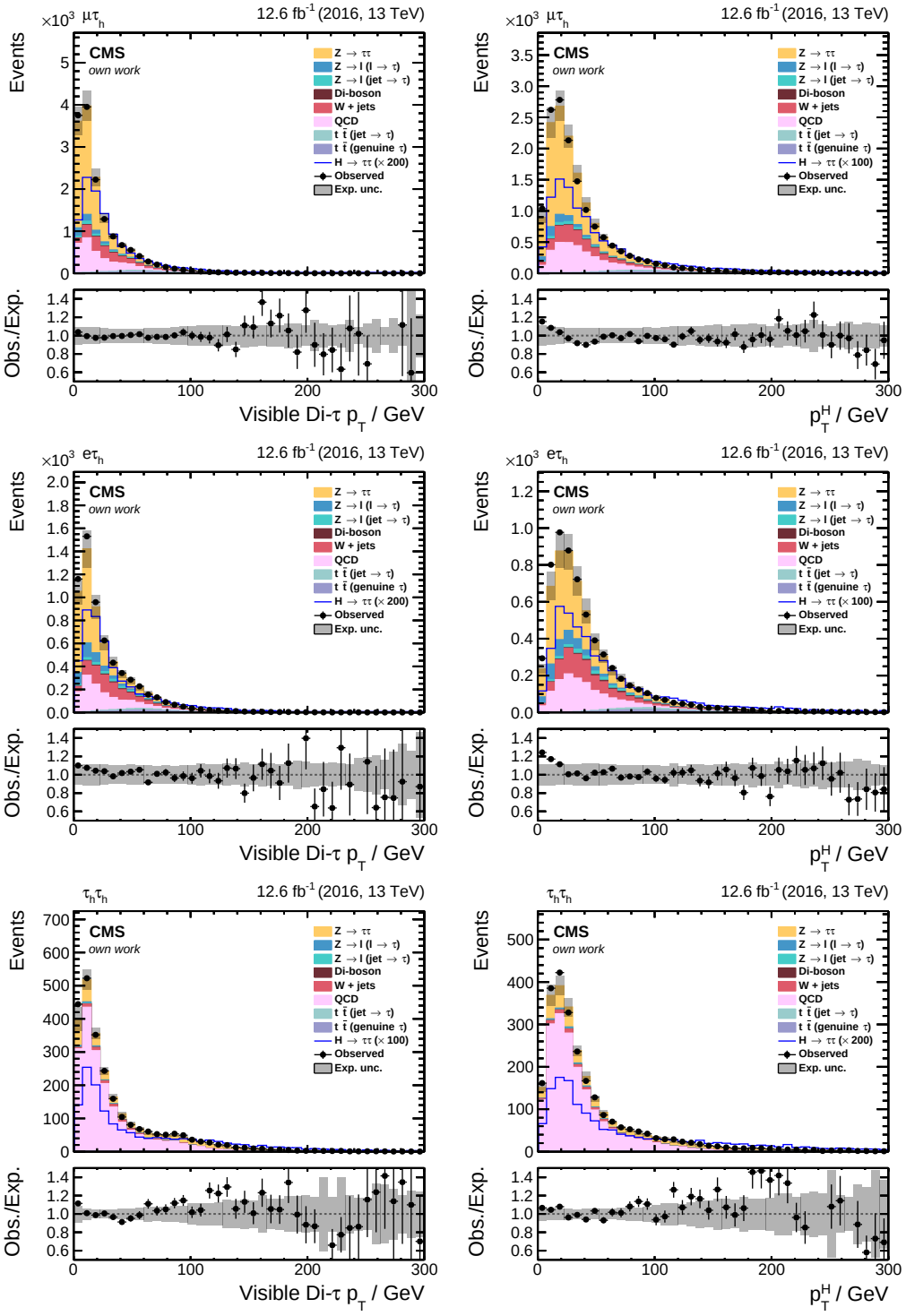


Figure A.10: Di-Lepton transverse momentum (left) and Di-Lepton plus $\not{E}_{\mathrm{T}}$ transverse momentum (right) in the $\mu\tau_h$ (first row), $e\tau_h$ (middle row) and $\tau_h\tau_h$ (bottom row) inclusive channels.

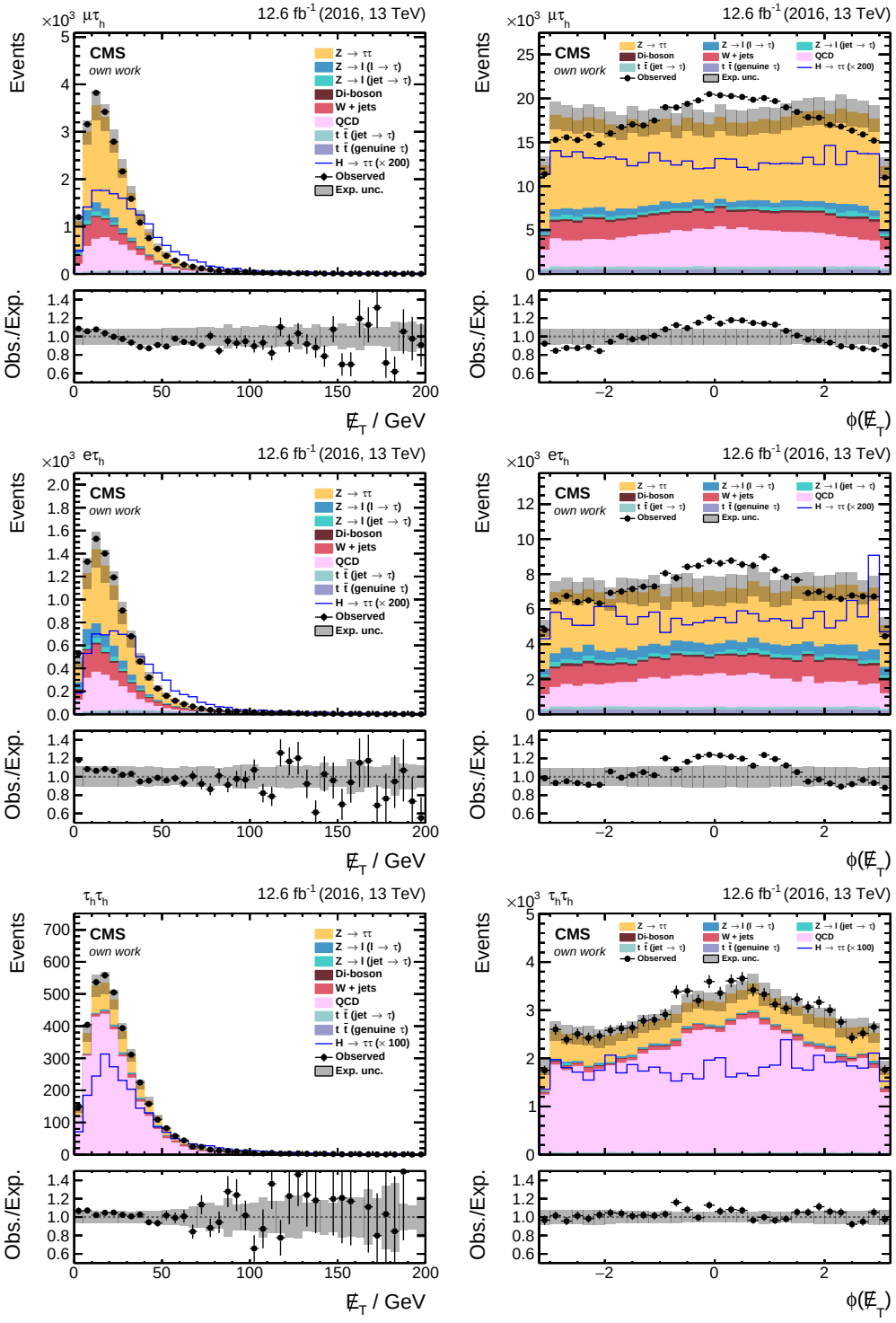


Figure A.11: $\not\!\!E_{\mathrm{T}}^{\mathrm{MVA}}$ and $\phi_{\not\!\!E_{\mathrm{T}}}^{MVA}$ in the $\mu\tau_h$ (first row), $e\tau_h$ (middle row) and $\tau_h\tau_h$ (bottom row) inclusive channels.

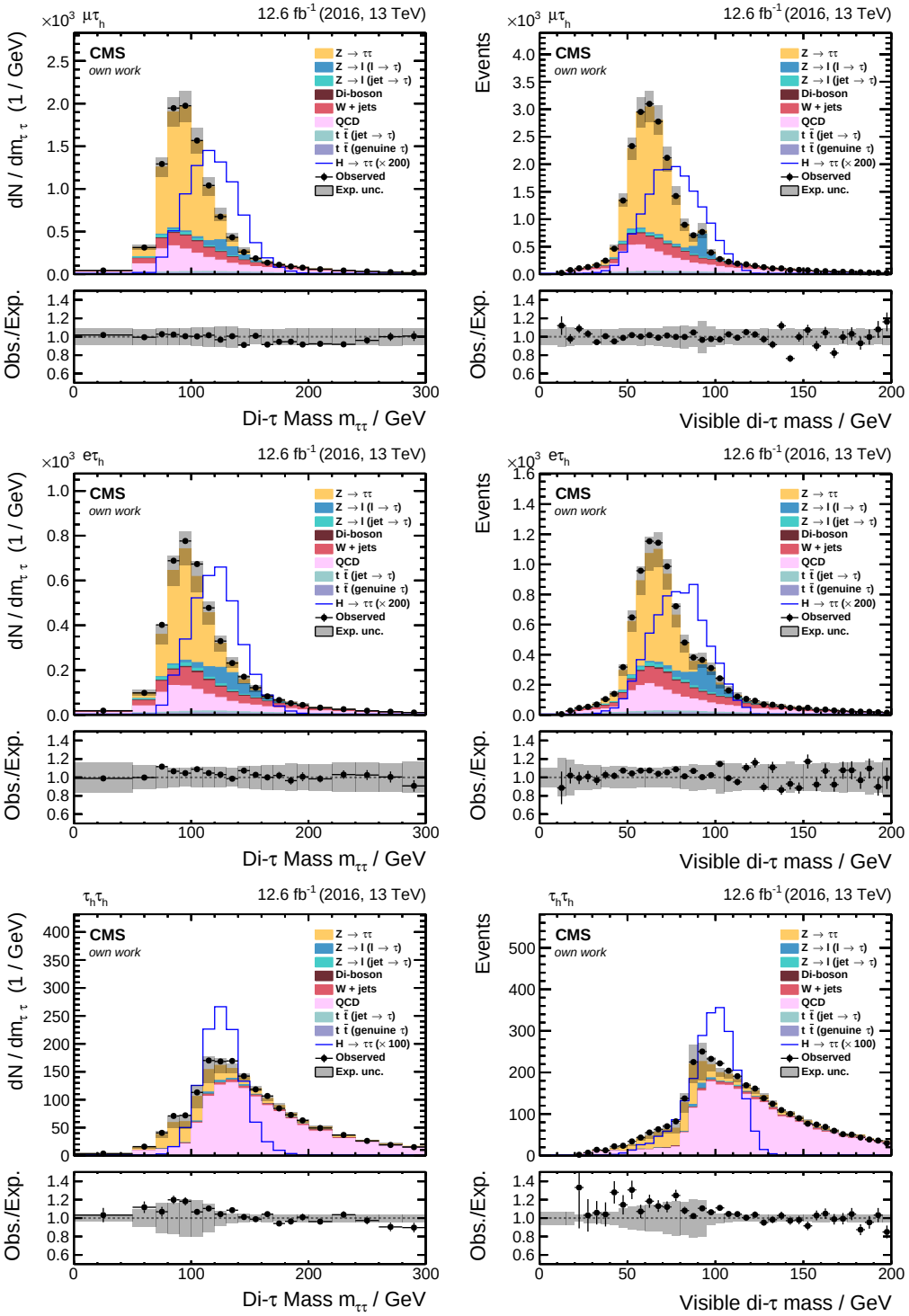
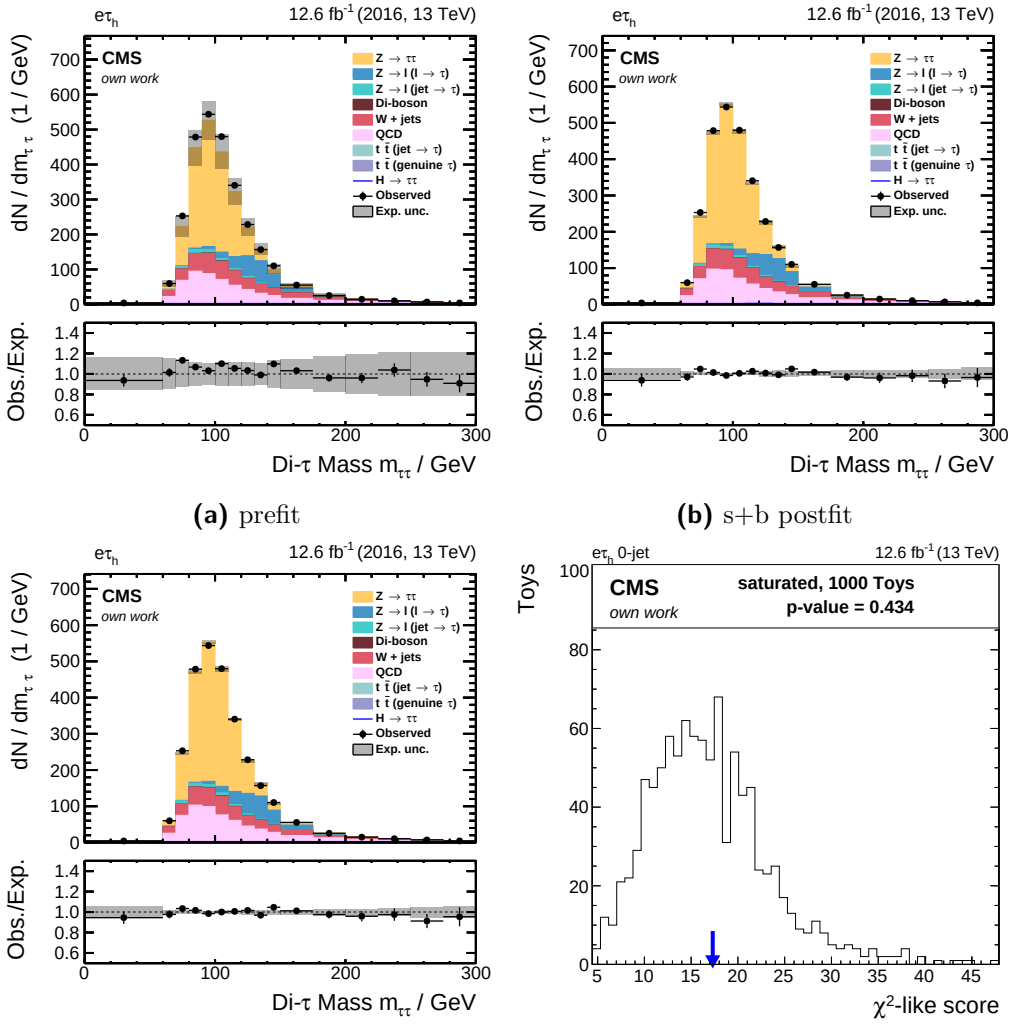


Figure A.12: Mass of the fully reconstructed di-tau system $m_{\tau\tau}^{SVFit}$ and the mass of the visible decay components $m_{\tau\tau}^{vis}$ in the $\mu\tau_h$ (first row), $e\tau_h$ (middle row) and $\tau_h\tau_h$ (bottom row) inclusive channels.

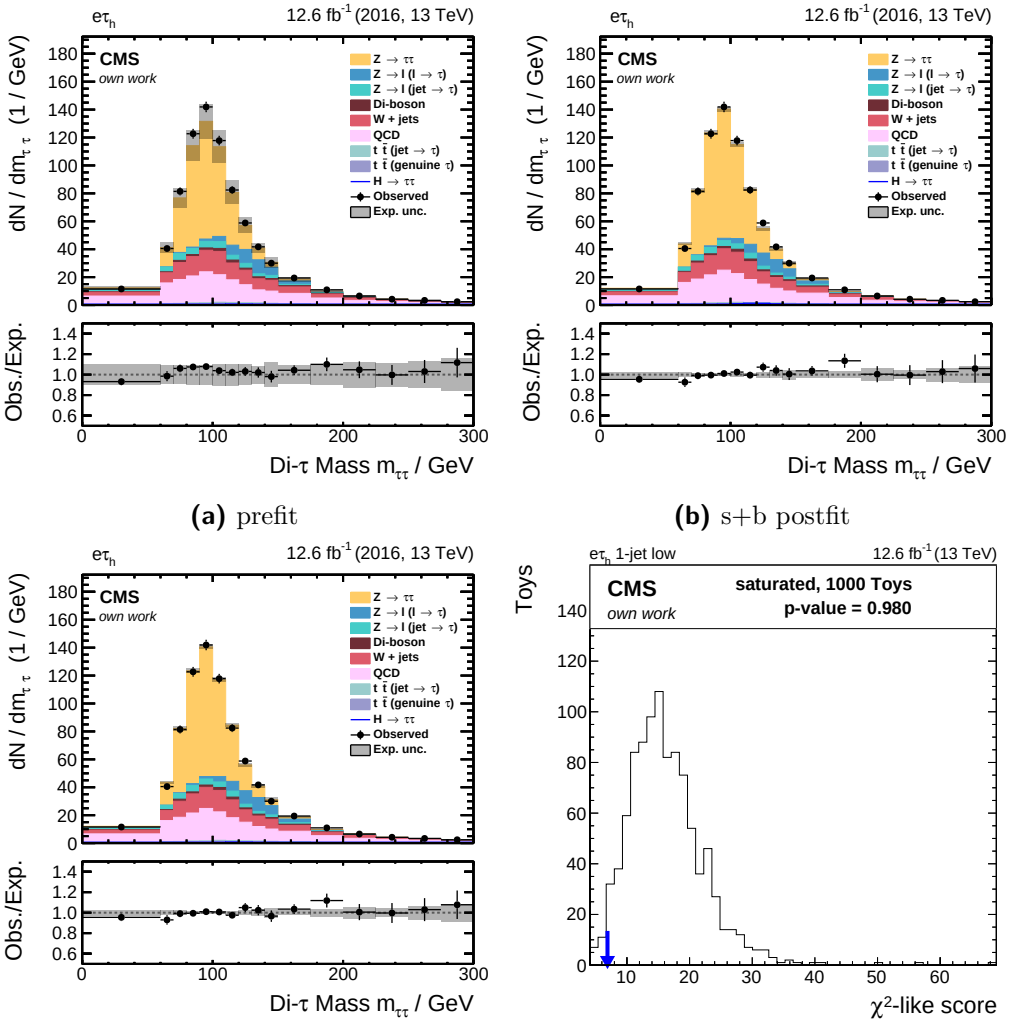
A.4.2 Di-tau mass $m_{\tau\tau}^{SVFit}$



(a) prefit **(b)** s+b postfit

(c) b postfit **(d)** Goodness-of-fit test with saturated model

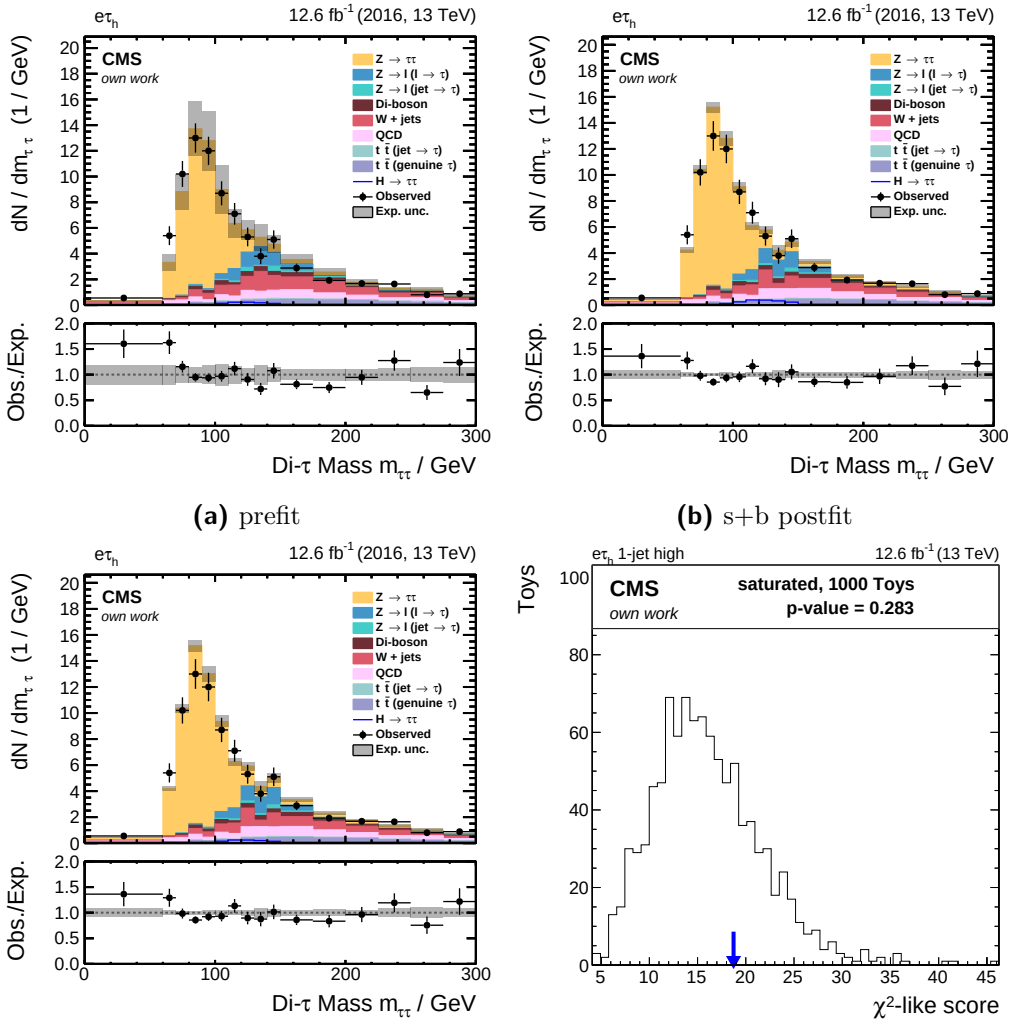
Figure A.13: Control distributions of the di-tau mass $m_{\tau\tau}^{SVFit}$ in the $e\tau_h$ 0-jet category. (a) pre-fit, (b) under the signal plus background hypothesis and (c) under the background-only hypothesis. (d) goodness-of-fit test control plot.



(a) prefit **(b)** s+b postfit

(c) b postfit **(d)** Goodness-of-fit test with saturated model

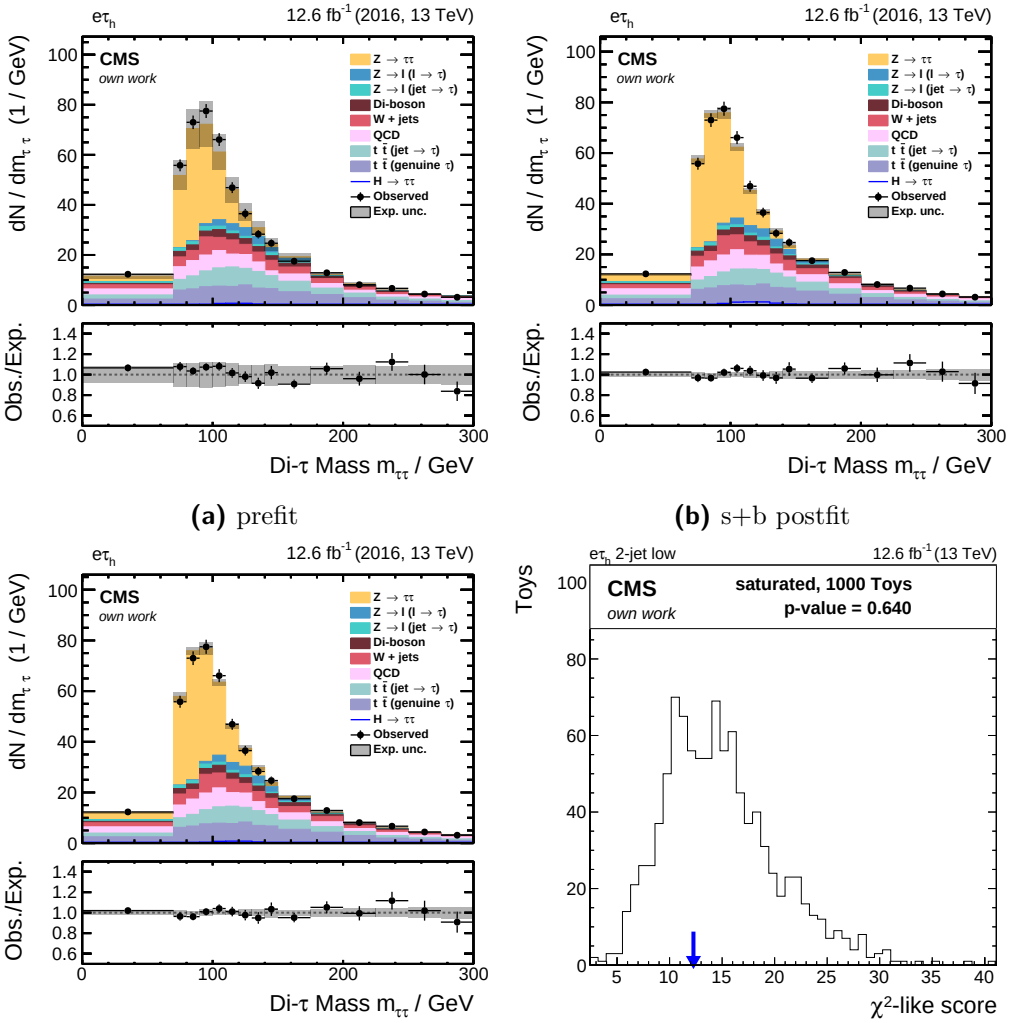
Figure A.14: Control distributions of the di-tau mass $m_{\tau\tau}^{SVFit}$ in the $e\tau_h$ 1-jet low category. (a) pre-fit, (b) under the signal plus background hypothesis and (c) under the background-only hypothesis. (d) goodness-of-fit test control plot.



(a) prefit **(b)** s+b postfit

(c) b postfit **(d)** Goodness-of-fit test with saturated model

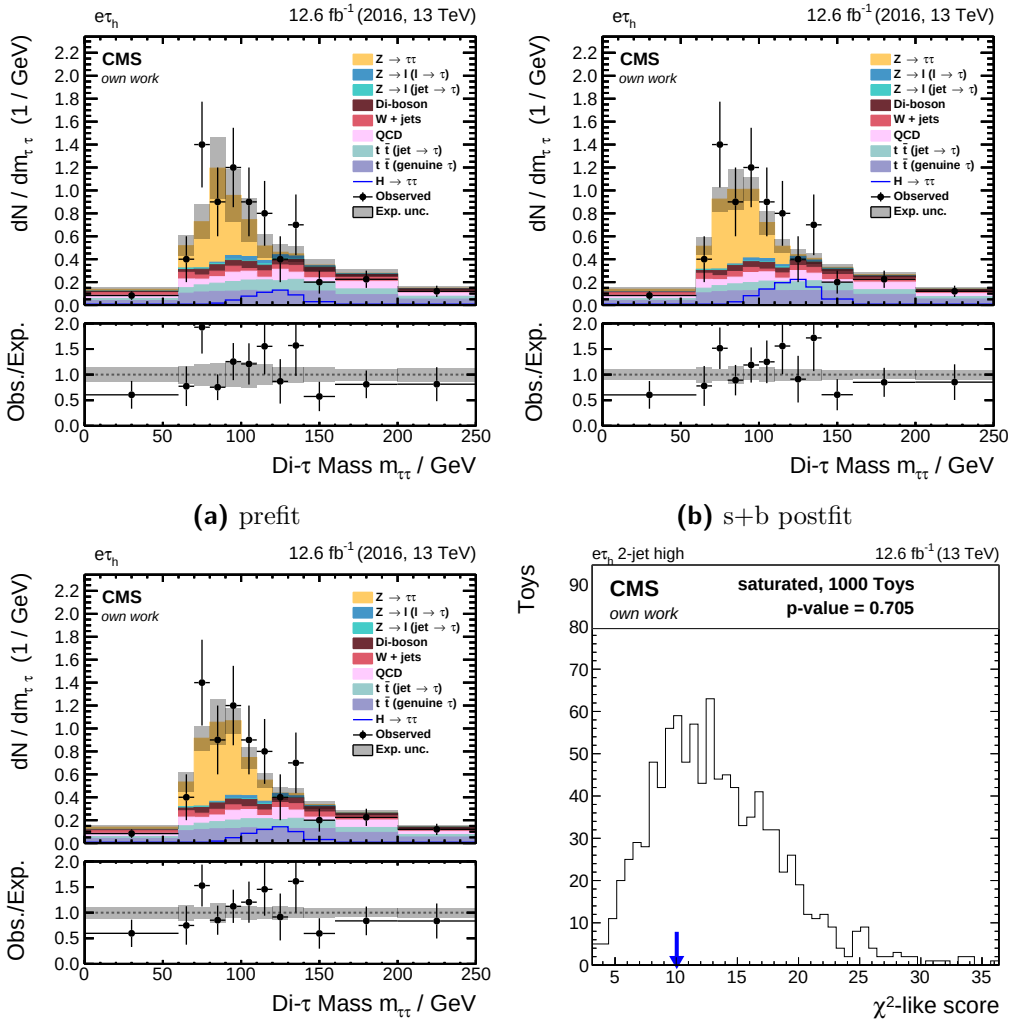
Figure A.15: Control distributions of the di-tau mass $m_{\tau\tau}^{SVFit}$ in the $e\tau_h$ 1-jet high category. (a) pre-fit, (b) under the signal plus background hypothesis and (c) under the background-only hypothesis. (d) goodness-of-fit test control plot.



(a) prefit **(b)** s+b postfit

(c) b postfit **(d)** Goodness-of-fit test with saturated model

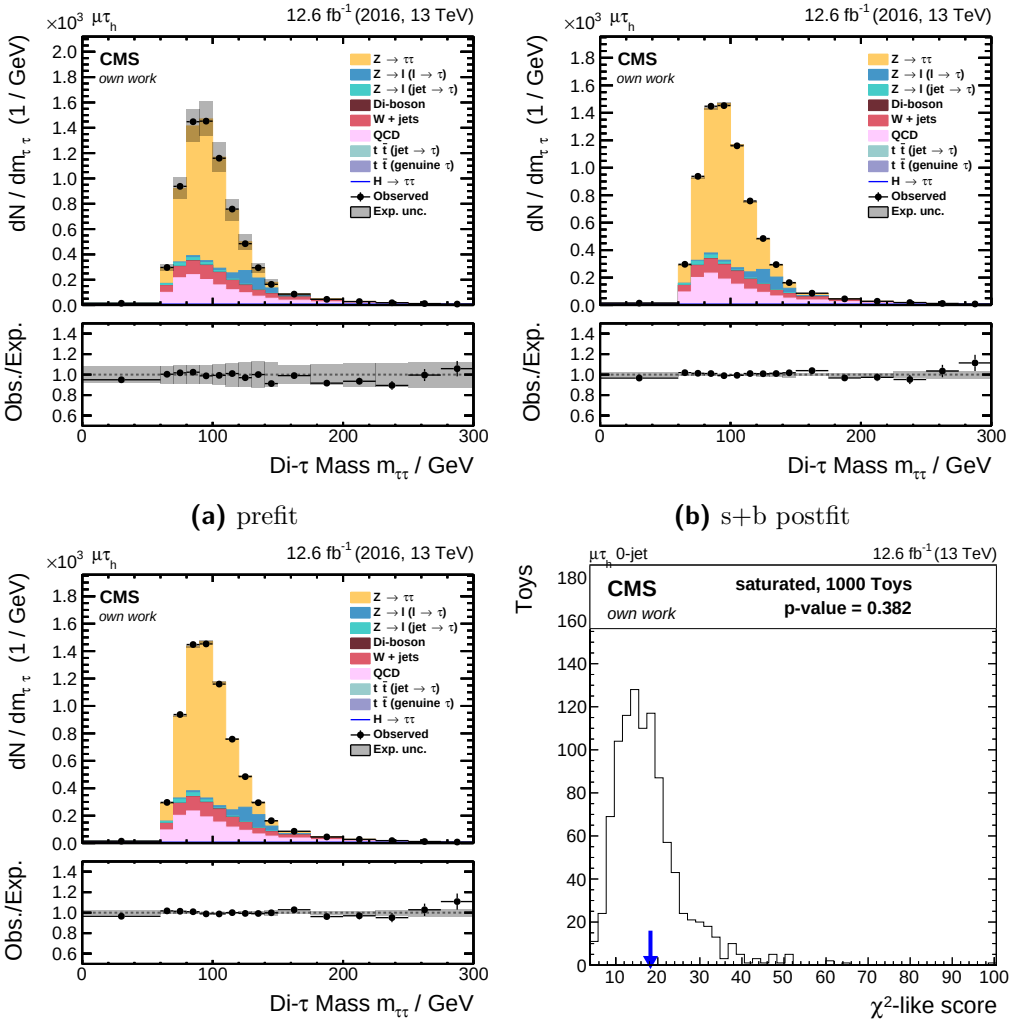
Figure A.16: Control distributions of the di-tau mass $m_{\tau\tau}^{SVFit}$ in the $e\tau_h$ 2-jet low category. (a) pre-fit, (b) under the signal plus background hypothesis and (c) under the background-only hypothesis. (d) goodness-of-fit test control plot.



(a) prefit **(b)** s+b postfit

(c) b postfit **(d)** Goodness-of-fit test with saturated model

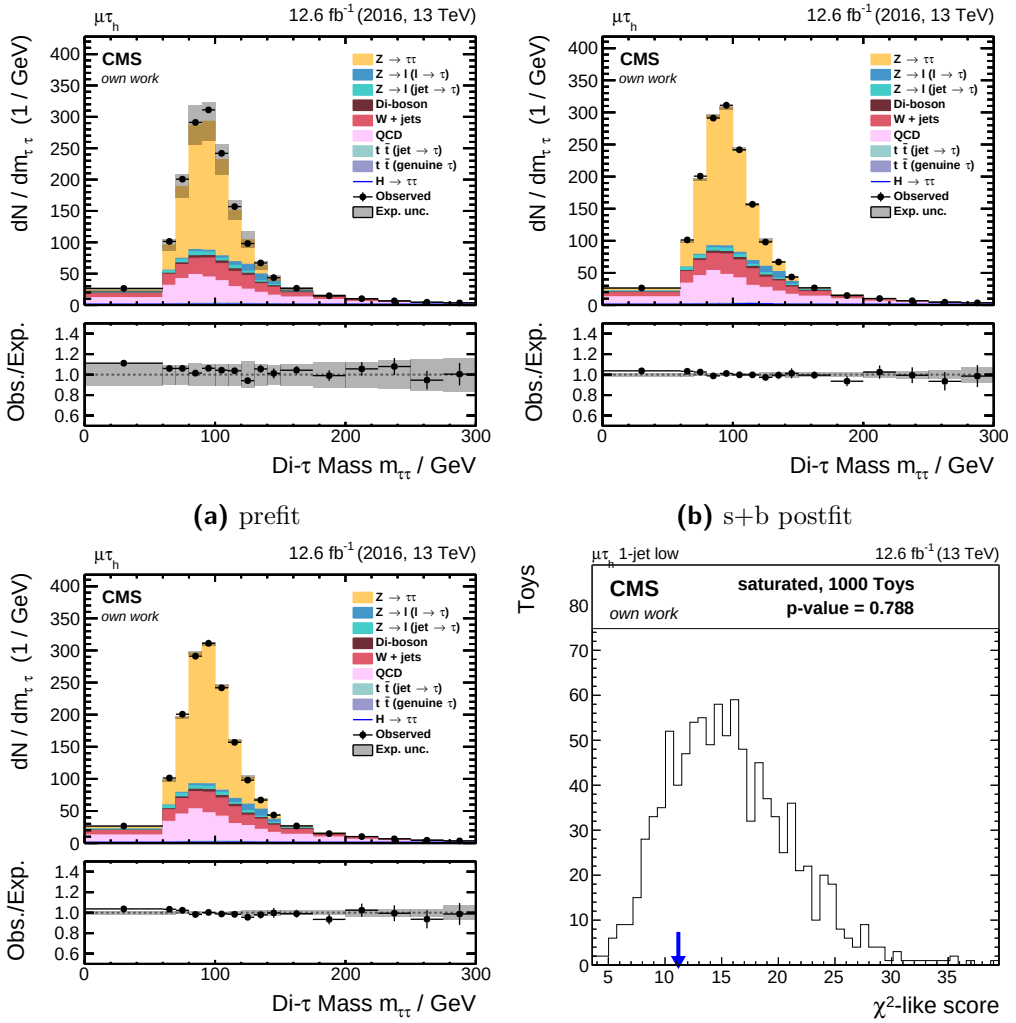
Figure A.17: Control distributions of the di-tau mass $m_{\tau\tau}^{SVFit}$ in the $e\tau_h$ 2-jet high category. (a) pre-fit, (b) under the signal plus background hypothesis and (c) under the background-only hypothesis. (d) goodness-of-fit test control plot.



(a) prefit **(b)** s+b postfit

(c) b postfit **(d)** Goodness-of-fit test with saturated model

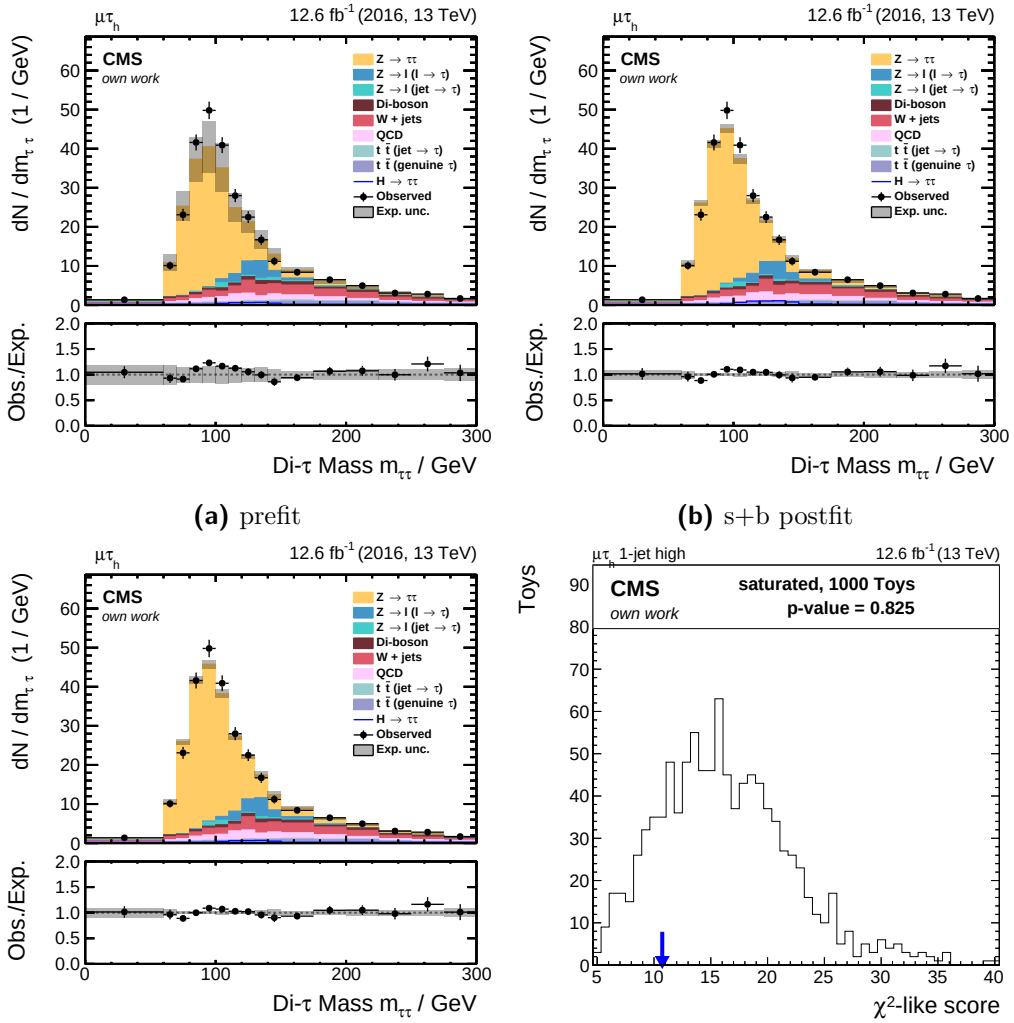
Figure A.18: Control distributions of the di-tau mass $m_{\tau\tau}^{SVFit}$ in the $\mu\tau_h$ 0-jet category. (a) pre-fit, (b) under the signal plus background hypothesis and (c) under the background-only hypothesis. (d) goodness-of-fit test control plot.



(a) prefit **(b)** s+b postfit

(c) b postfit **(d)** Goodness-of-fit test with saturated model

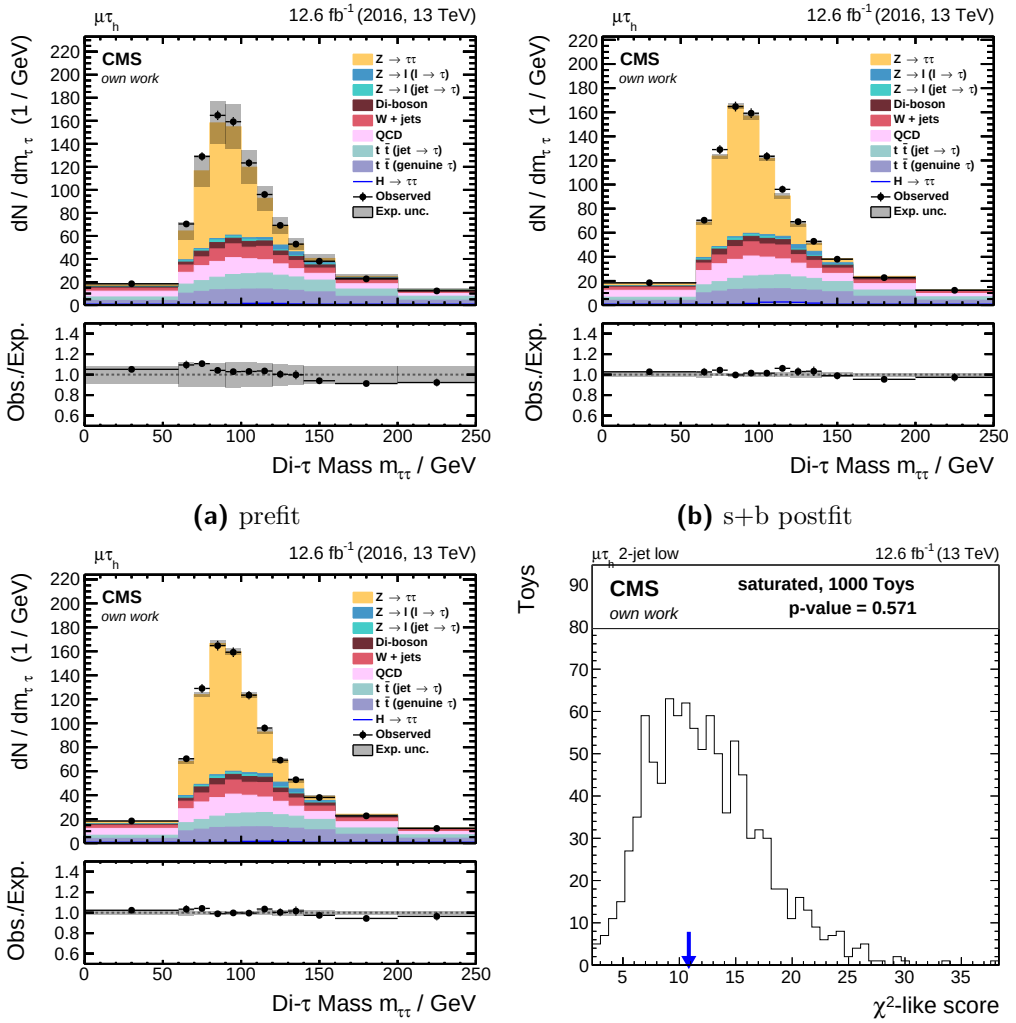
Figure A.19: Control distributions of the di-tau mass $m_{\tau\tau}^{SVFit}$ in the $\mu\tau_h$ 1-jet low category. (a) pre-fit, (b) under the signal plus background hypothesis and (c) under the background-only hypothesis. (d) goodness-of-fit test control plot.



(a) prefit **(b)** s+b postfit

(c) b postfit **(d)** Goodness-of-fit test with saturated model

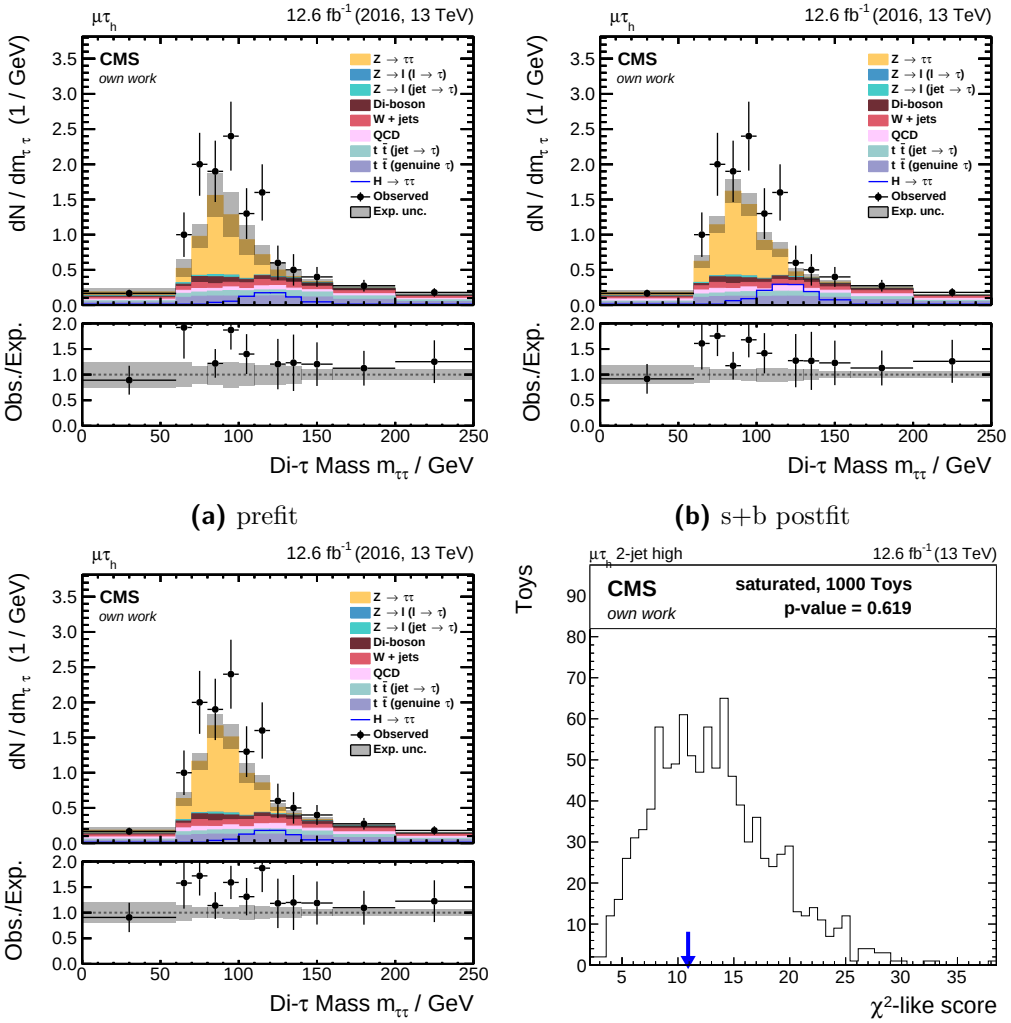
Figure A.20: Control distributions of the di-tau mass $m_{\tau\tau}^{SVFit}$ in the $\mu\tau_h$ 1-jet high category. (a) pre-fit, (b) under the signal plus background hypothesis and (c) under the background-only hypothesis. (d) goodness-of-fit test control plot.



(a) prefit **(b)** s+b postfit

(c) b postfit **(d)** Goodness-of-fit test with saturated model

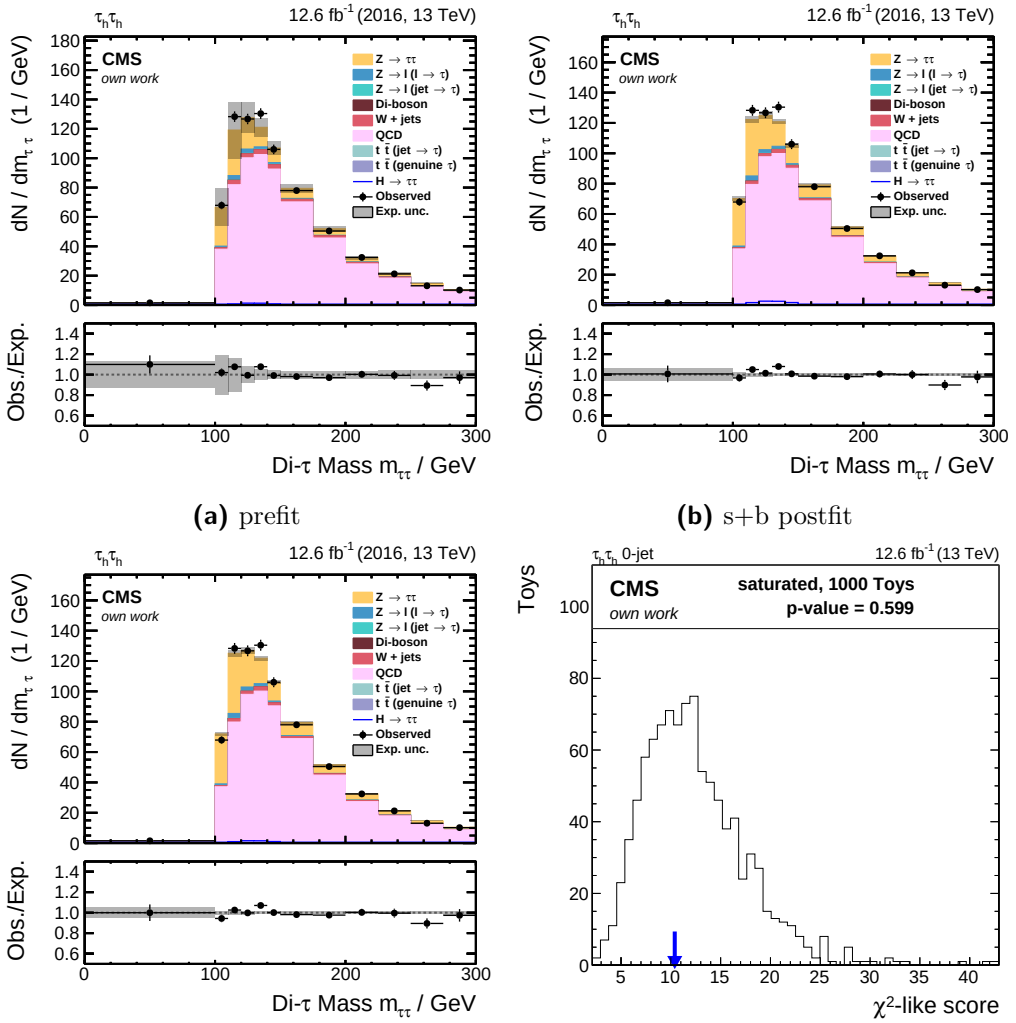
Figure A.21: Control distributions of the di-tau mass $m_{\tau\tau}^{SVFit}$ in the $\mu\tau_h$ 2-jet low category. (a) pre-fit, (b) under the signal plus background hypothesis and (c) under the background-only hypothesis. (d) goodness-of-fit test control plot.



(a) prefit **(b)** s+b postfit

(c) b postfit **(d)** Goodness-of-fit test with saturated model

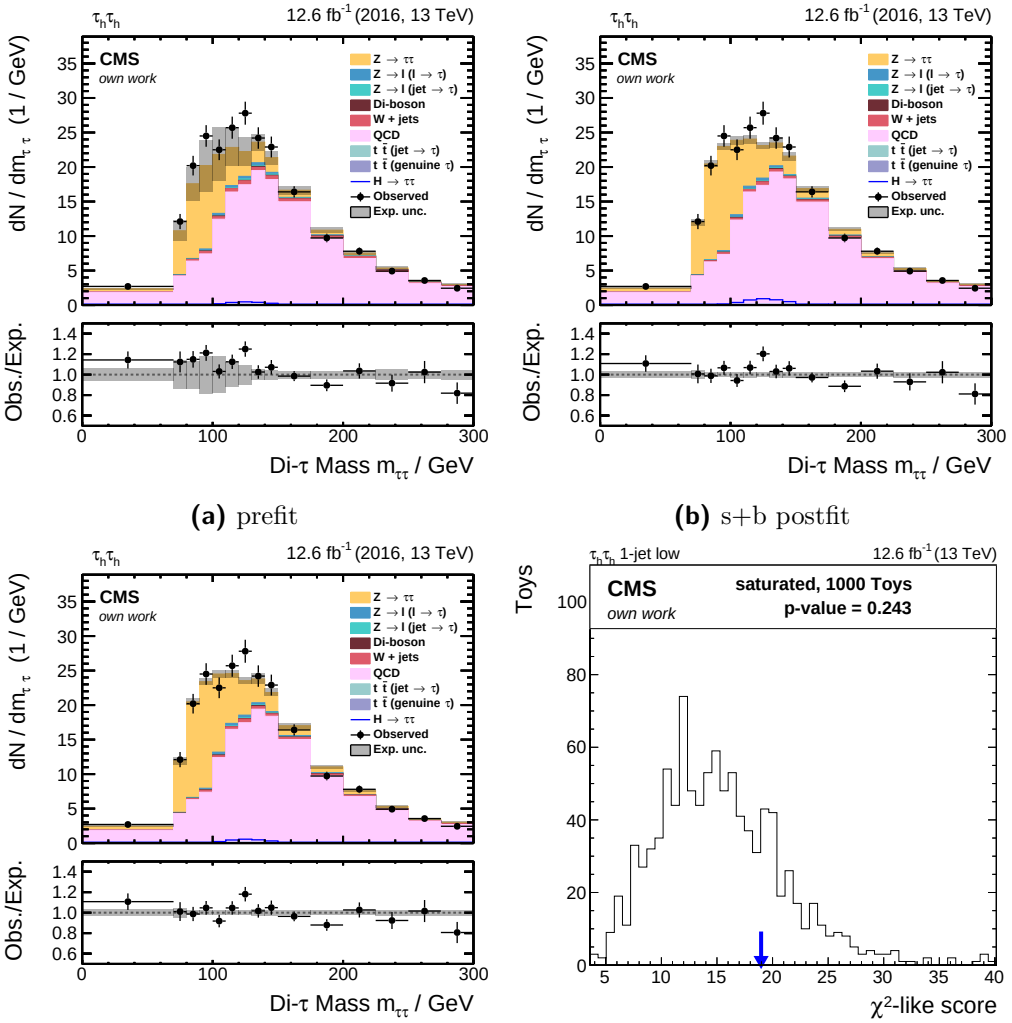
Figure A.22: Control distributions of the di-tau mass $m_{\tau\tau}^{SVFit}$ in the $\mu\tau_h$ 2-jet high category. (a) pre-fit, (b) under the signal plus background hypothesis and (c) under the background-only hypothesis. (d) goodness-of-fit test control plot.



(a) prefit **(b)** s+b postfit

(c) b postfit **(d)** Goodness-of-fit test with saturated model

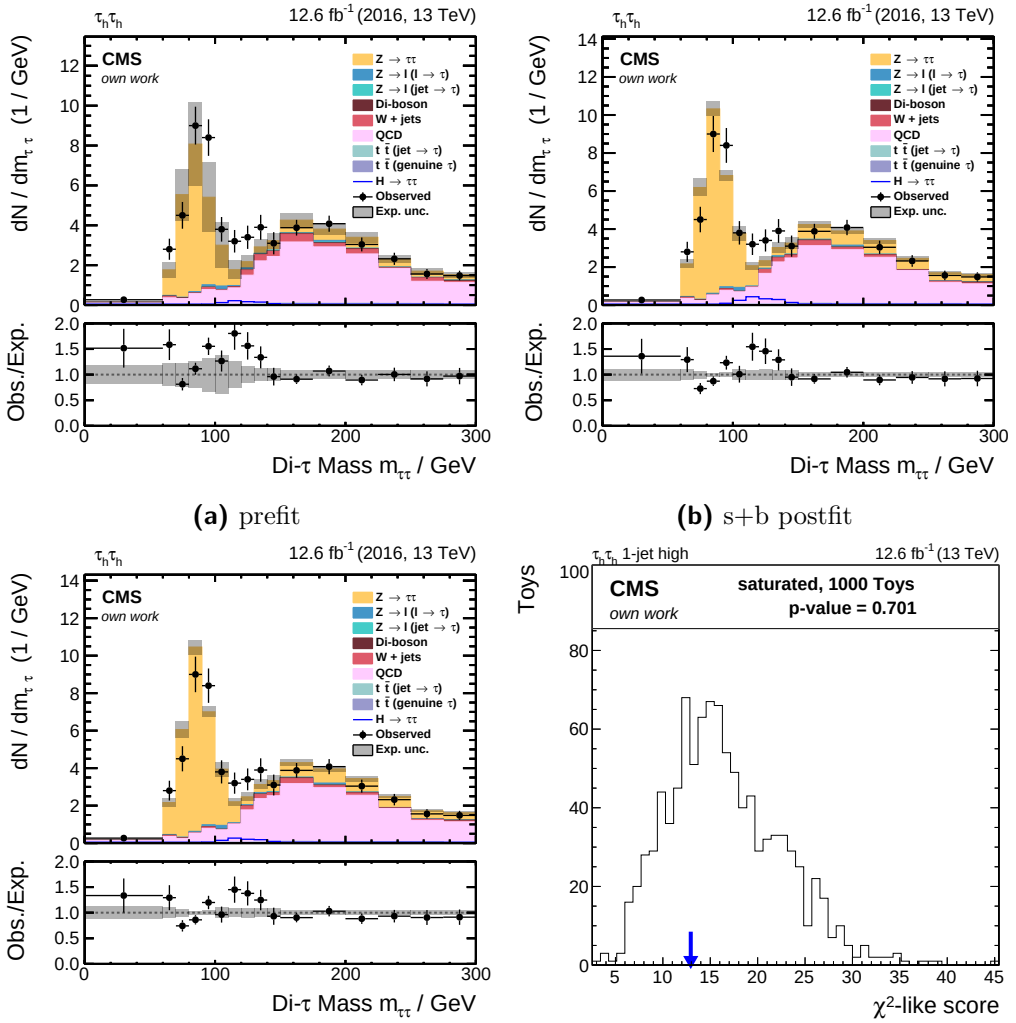
Figure A.23: Control distributions of the di-tau mass $m_{\tau\tau}^{SVFit}$ in the $\tau_h\tau_h$ 0-jet category. (a) pre-fit, (b) under the signal plus background hypothesis and (c) under the background-only hypothesis. (d) goodness-of-fit test control plot.



(a) prefit **(b)** s+b postfit

(c) b postfit **(d)** Goodness-of-fit test with saturated model

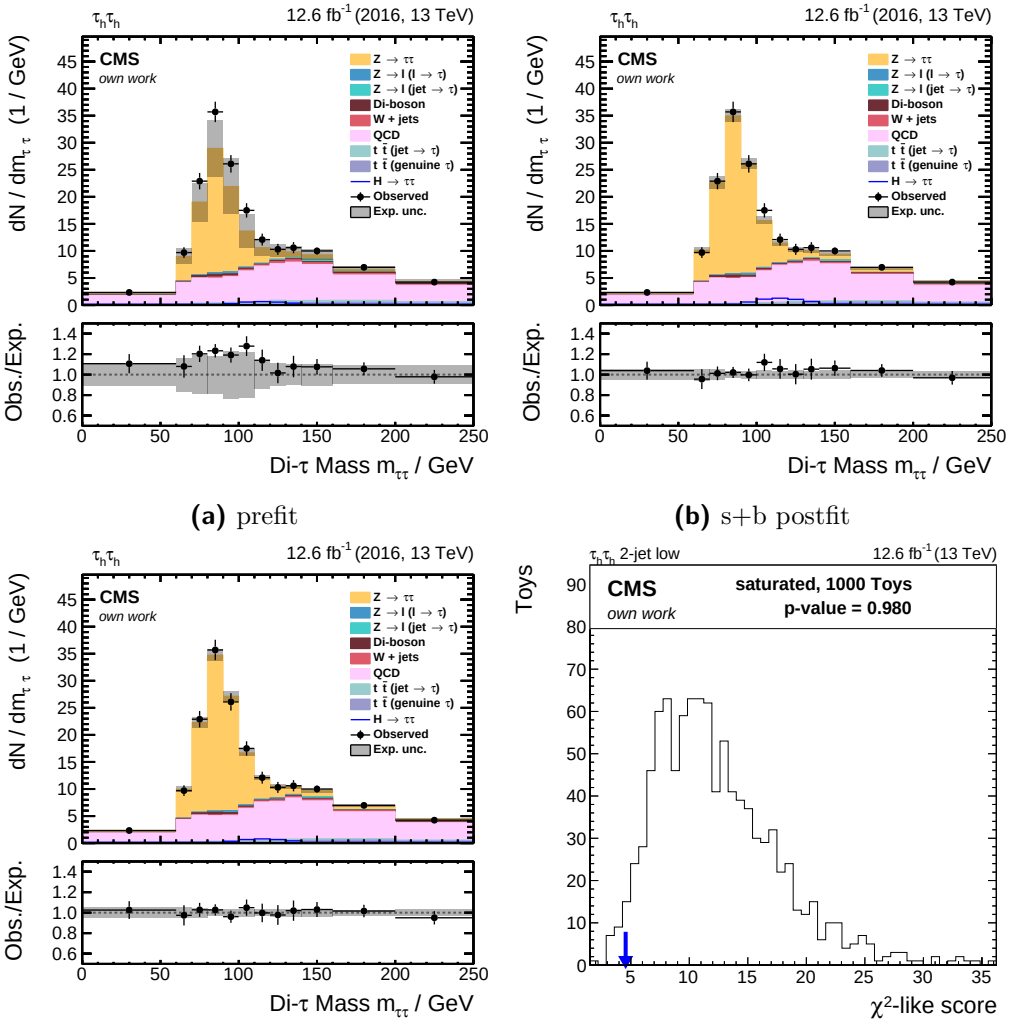
Figure A.24: Control distributions of the di-tau mass $m_{\tau\tau}^{SVFit}$ in the $\tau_h\tau_h$ 1-jet low category. (a) pre-fit, (b) under the signal plus background hypothesis and (c) under the background-only hypothesis. (d) goodness-of-fit test control plot.



(a) prefit **(b)** s+b postfit

(c) b postfit **(d)** Goodness-of-fit test with saturated model

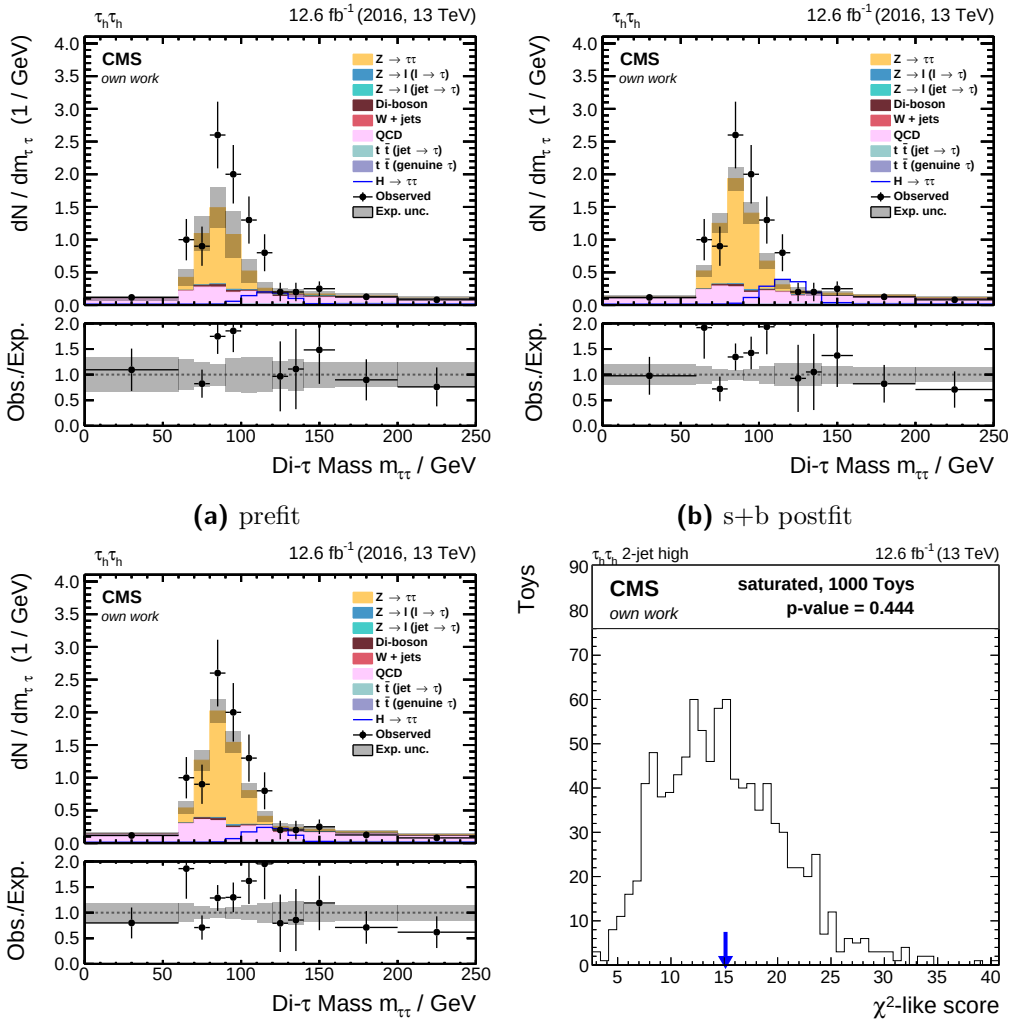
Figure A.25: Control distributions of the di-tau mass $m_{\tau\tau}^{SVFit}$ in the $\tau_h\tau_h$ 1-jet high category. (a) pre-fit, (b) under the signal plus background hypothesis and (c) under the background-only hypothesis. (d) goodness-of-fit test control plot.



(a) prefit **(b)** s+b postfit

(c) b postfit **(d)** Goodness-of-fit test with saturated model

Figure A.26: Control distributions of the di-tau mass $m_{\tau\tau}^{SVFit}$ in the $\tau_h\tau_h$ 2-jet low category. (a) pre-fit, (b) under the signal plus background hypothesis and (c) under the background-only hypothesis. (d) goodness-of-fit test control plot.



(a) prefit **(b)** s+b postfit

(c) b postfit **(d)** Goodness-of-fit test with saturated model

Figure A.27: Control distributions of the di-tau mass $m_{\tau\tau}^{SVFit}$ in the $\tau_h\tau_h$ 2-jet high category. (a) pre-fit, (b) under the signal plus background hypothesis and (c) under the background-only hypothesis. (d) goodness-of-fit test control plot.

Bibliography

All references to CERN publications can be found on the CERN Document Server: cds.cern.ch. References to theses published at the Institut für Experimentelle Teilchenphysik are available at ekp.kit.edu/invenio. For arXiv preprints, please refer to arxiv.org. The digital object identifier (DOI) is used to uniquely identify referenced papers and can be resolved on dx.doi.org.

[1] David Griffiths, *Introduction to Elementary Particles*, Wiley-VCH GmbH & Co. KGaA, Weinheim, 2014, ISBN: 978-3527406012.

[2] Roger Wolf, *The Higgs Boson Discovery at the Large Hadron Collider*, Springer, 2015. Aufl. ISBN: 978-3-319-18512-5.

[3] Peter Schmüser, *Feynman-Graphen und Eichtheorien für Experimentalphysiker*, Springer, 1995, ISBN: 978-3-642-57766-6URL: https://cds.cern.ch/record/1605928.

[4] P.W. Higgs, "Broken symmetries, massless particles and gauge fields", *Physics Letters* **12** (1964) 132–133, DOI: 10.1016/0031-9163(64)91136-9.

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

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

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

[8] T. W. B. Kibble, "Symmetry Breaking in Non-Abelian Gauge Theories", *Phys. Rev.* **155** (1967) 1554–1561, DOI: 10.1103/PhysRev.155.1554.

[9] C. Grojean et al., "Very boosted Higgs in gluon fusion", *Journal of High Energy Physics* (2014) p. 22, DOI: 10.1007/JHEP05(2014)022.

[10] Matthew R. Buckley and Michael J. Ramsey-Musolf,
"Diagnosing spin at the LHC via vector boson fusion",
Journal of High Energy Physics (2011) p. 94,
DOI: 10.1007/JHEP09(2011)094.

[11] V. Hankele et al.,
"Anomalous Higgs boson couplings in vector boson fusion at the CERN LHC",
Phys. Rev. D **74** (2006) p. 095001, DOI: 10.1103/PhysRevD.74.095001.

[12] Sayipjamal Dulat et al., "New parton distribution functions from a global analysis of quantum chromodynamics", *Phys. Rev. D* **93** (2016) p. 033006,
DOI: 10.1103/PhysRevD.93.033006.

[13] J. Baglio et al.,
"The measurement of the Higgs self-coupling at the LHC: theoretical status",
Journal of High Energy Physics **04** (2013) p. 151,
DOI: 10.1007/JHEP04(2013)151, [arXiv:1212.5581].

[14] A. Denner et al.,
"Standard model Higgs-boson branching ratios with uncertainties",
The European Physical Journal C **71** (2011) p. 1753,
DOI: 10.1140/epjc/s10052-011-1753-8.

[15] Oliver Sim Brüning et al., *LHC Design Report*, CERN, 2004,
ISBN: 9789290832249URL: http://cds.cern.ch/record/782076.

[16] Stephen Myers,
The LEP Collider, from design to approval and commissioning,
John Adams' Lecture, CERN, 1991,
ISBN: 9789290830405URL: http://cds.cern.ch/record/226776.

[17] CERN, *The HL-LHC project*, 2016,
URL: http://hilumilhc.web.cern.ch/about/hl-lhc-project.

[18] Stephen Myers, "LHC Guide", 2017,
URL: http://cds.cern.ch/record/2255762.

[19] The CMS collaboration, *CMS Web*, 2016,
URL: https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults#Luminosity_versus_week.

[20] Stephen Myers, *The CMS magnet project: Technical Design Report*,
Technical Design Report CMS, CERN, 1997,
ISBN: 9789290831013URL: http://cds.cern.ch/record/331056.

[21] David Varney, *CMS Document Server*, 2012, URL: https://cms-docdb.cern.ch/cgi-bin/PublicDocDB/ShowDocument?docid=5581.

[22] V Karimäki et al.,
"The CMS tracker system project: Technical Design Report",
Technical Design Report CMS (1997),
URL: http://cds.cern.ch/record/368412.

[23] CMS Collaboration, "Commissioning and performance of the CMS silicon strip tracker with cosmic ray muons",
Journal of Instrumentation, Technical Design Report CMS **5** (2010) T03008,
URL: http://stacks.iop.org/1748-0221/5/i=03/a=T03008.

[24] Katja Klein, "Lessons learned during CMS tracker end cap construction",
Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, Technical Design Report CMS **579** (2007) 731–735, DOI: 10.1016/j.nima.2007.05.283.

[25] The CMS Collaboration, "The CMS experiment at the CERN LHC",
Journal of Instrumentation, Technical Design Report CMS **3** (2008) S08004,
URL: http://stacks.iop.org/1748-0221/3/i=08/a=S08004.

[26] The CMS Collaboration, "Description and performance of track and primary-vertex reconstruction with the CMS tracker",
Journal of Instrumentation, Technical Design Report CMS **9** (2014) P10009,
URL: http://stacks.iop.org/1748-0221/9/i=10/a=P10009.

[27] The CMS collaboration, "CMS Physics: Technical Design Report Volume 1: Detector Performance and Software", Technical Design Report CMS (2006),
URL: http://cds.cern.ch/record/922757.

[28] Stephen Myers, *Particle-Flow Event Reconstruction in CMS and Performance for Jets, Taus, and MET*, tech. rep., CERN, 2009,
URL: http://cds.cern.ch/record/1194487.

[29] Daniele Bertolini et al., "Pileup per particle identification",
Journal of High Energy Physics, Technical Design Report CMS (2014) p. 59,
DOI: 10.1007/JHEP10(2014)059.

[30] Matteo Cacciari, Gavin P. Salam and Gregory Soyez,
"The anti-k_T jet clustering algorithm",
Journal of High Energy Physics, Technical Design Report CMS (2008) p. 063,
URL: http://stacks.iop.org/1126-6708/2008/i=04/a=063.

[31] Yu.L. Dokshitzer et al., "Better jet clustering algorithms",
Journal of High Energy Physics, Technical Design Report CMS (1997) p. 001,
URL: http://stacks.iop.org/1126-6708/1997/i=08/a=001.

[32] Stephen D. Ellis and Davison E. Soper,
"Successive combination jet algorithm for hadron collisions",
Phys. Rev. D, Technical Design Report CMS **48** (1993) 3160–3166,
DOI: 10.1103/PhysRevD.48.3160.

[33] Gregory Soyez Matteo Cacciari Gavin P. Salam, "FastJet user manual",
Technical Design Report CMS (2011),
URL: https://arxiv.org/abs/1111.6097.

[34] The CMS collaboration, "Determination of jet energy calibration and transverse momentum resolution in CMS",
Journal of Instrumentation, Technical Design Report CMS **6** (2011) P11002,
URL: http://stacks.iop.org/1748-0221/6/i=11/a=P11002.

[35] Joram Berger, Günter Quast and Wim de Boer, "Search for the Higgs Boson Produced via Vector-Boson Fusion in the Decay Channel H $\to \tau\tau$",
Technical Design Report CMS (2014),
URL: http://cds.cern.ch/record/1747055.

[36] The CMS collaboration,
"Identification of b-quark jets with the CMS experiment",
Journal of Instrumentation, Technical Design Report CMS **8** (2013) P04013,
URL: http://stacks.iop.org/1748-0221/8/i=04/a=P04013.

[37] CMS Collaboration, *Pileup Jet Identification*, tech. rep., CERN, 2013,
URL: https://cds.cern.ch/record/1581583.

[38] CMS Collaboration,
"Performance of τ-lepton reconstruction and identification in CMS",
Journal of Instrumentation, Technical Design Report CMS **7** (2012) P01001,
URL: http://stacks.iop.org/1748-0221/7/i=01/a=P01001.

[39] The CMS collaboration,
"Evidence for the 125 GeV Higgs boson decaying to a pair of τ leptons",
Journal of High Energy Physics, Technical Design Report CMS (2014) p. 104,
DOI: 10.1007/JHEP05(2014)104.

[40] Andreas Hoecker et al., "TMVA: Toolkit for Multivariate Data Analysis",
PoS(ACAT2007)040, Technical Design Report CMS (2007)
[arXiv:physics/0703039].

[41] Friese Raphael and Günter Quast,
"A new Multivariate Approach for the $H \to \tau\tau \to \mu\mu$ analysis",
Technical Design Report CMS (2013),
URL: http://ekp-invenio.physik.uni-karlsruhe.de/record/48275/files/iekp-ka2013-14.pdf.

[42] Peter J. Huber, "Robust Estimation of a Location Parameter", *Ann. Math. Statist.* Technical Design Report CMS **35** (1964) 73–101, DOI: 10.1214/aoms/1177703732.

[43] The CMS collaboration, "Observation of a new boson with mass near 125 GeV in pp collisions at $\sqrt{s} = 7$ and 8 TeV", *Journal of High Energy Physics*, Technical Design Report CMS (2013) p. 81, DOI: 10.1007/JHEP06(2013)081.

[44] The ATLAS collaboration, "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC", *Physics Letters B*, Technical Design Report CMS **716** (2012) 1–29, DOI: http://dx.doi.org/10.1016/j.physletb.2012.08.020.

[45] A. L. Read, "Presentation of search results: the CL_s technique", *Journal of Physics G: Nuclear and Particle Physics*, Technical Design Report CMS **28** (2002) p. 2693, URL: http://stacks.iop.org/0954-3899/28/i=10/a=313.

[46] Glen Cowan et al., "Asymptotic formulae for likelihood-based tests of new physics", *Eur. Phys. J.* Technical Design Report CMS **C71** (2011) p. 1554, DOI: 10.1140/epjc/s10052-011-1554-0,10.1140/epjc/s10052-013-2501-z, [arXiv:1007.1727].

[47] The CMS collaboration, "Observation of the diphoton decay of the Higgs boson and measurement of its properties", *The European Physical Journal C*, Technical Design Report CMS **74** (2014) p. 3076, DOI: 10.1140/epjc/s10052-014-3076-z.

[48] The CMS Collaboration, "Measurement of the properties of a Higgs boson in the four-lepton final state", *Phys. Rev. D*, Technical Design Report CMS **89** (2014) p. 092007, DOI: 10.1103/PhysRevD.89.092007.

[49] The CMS Collaboration, "Measurement of Higgs boson production and properties in the WW decay channel with leptonic final states", *Journal of High Energy Physics*, Technical Design Report CMS (2014) p. 96, DOI: 10.1007/JHEP01(2014)096.

[50] The CMS Collaboration, "Search for the standard model Higgs boson produced in association with a W or a Z boson and decaying to bottom quarks", *Phys. Rev. D*, Technical Design Report CMS **89** (2014) p. 012003, DOI: 10.1103/PhysRevD.89.012003.

[51] Vardan Khachatryan et al.,
"Search for resonant pair production of Higgs bosons decaying to two bottom quark–antiquark pairs in proton–proton collisions at 8 TeV",
Phys. Lett. Technical Design Report CMS **B749** (2015) 560–582,
DOI: 10.1016/j.physletb.2015.08.047, [arXiv:1503.04114].

[52] Serguei Chatrchyan et al.,
"Missing transverse energy performance of the CMS detector",
JINST, Technical Design Report CMS **6** (2011) P09001,
DOI: 10.1088/1748-0221/6/09/P09001, [arXiv:1106.5048].

[53] CMS Collaboration, "Performance of missing energy reconstruction in 13 TeV pp collision data using the CMS detector",
Technical Design Report CMS (2016),
URL: http://cds.cern.ch/record/331056.

[54] CMS Collaboration, "Pileup per particle identification: preparation for Run 2",
Technical Design Report CMS (2015),
URL: http://cds.cern.ch/record/2051942.

[55] The CMS collaboration, "Multivariate Determination of the Missing Energy in the Transverse Plane ($\not\!\!E_{\mathrm{T}}$) at $\sqrt{s} = 13\,\mathrm{TeV}$",
Technical Design Report CMS (2015),
URL: http://cds.cern.ch/record/2048696.

[56] CMS Collaboration,
Search for a neutral MSSM Higgs boson decaying into tautau at 13 TeV,
tech. rep., CERN, 2016, URL: http://cds.cern.ch/record/2160252.

[57] The CMS collaboration, *Search for a neutral MSSM Higgs boson decaying into $\tau\tau$ with* 12.9 fb^{-1} *of data at* $\sqrt{s} = 13$ TeV, tech. rep., CERN, 2016,
URL: http://cds.cern.ch/record/2231507.

[58] CMS Collaboration, *Summaries of CMS cross section measurements*, 2017,
URL: https://twiki.cern.ch/twiki/bin/view/CMSPublic/PhysicsResultsCombined.

[59] CMS Collaboration,
CMS Luminosity Measurement for the 2015 Data Taking Period, tech. rep.,
CERN, 2016, URL: https://cds.cern.ch/record/2138682.

[60] Simone Alioli et al., "A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX",
Journal of High Energy Physics, Technical Design Report CMS (2010) p. 43,
DOI: 10.1007/JHEP06(2010)043.

[61] Paolo Nason and Carlo Oleari, "NLO Higgs boson production via vector-boson fusion matched with shower in POWHEG",
Journal of High Energy Physics, Technical Design Report CMS (2010) p. 37,
DOI: 10.1007/JHEP02(2010)037.

[62] Simone Alioli et al., "NLO Higgs boson production via gluon fusion matched with shower in POWHEG",
Journal of High Energy Physics, Technical Design Report CMS (2009) p. 002,
URL: http://stacks.iop.org/1126-6708/2009/i=04/a=002.

[63] Torbjörn Sjöstrand, Stephen Mrenna and Peter Skands,
"A brief introduction to PYTHIA 8.1", *Computer Physics Communications*,
Technical Design Report CMS **178** (2008) 852–867,
DOI: http://dx.doi.org/10.1016/j.cpc.2008.01.036.

[64] D. de Florian et al., "Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector", Technical Design Report CMS (2016),
[arXiv:1610.07922].

[65] K. A. Olive, "Review of Particle Physics",
Chin. Phys. Technical Design Report CMS **C40** (2016) p. 100001,
DOI: 10.1088/1674-1137/40/10/100001.

[66] Jon Butterworth et al., "PDF4LHC recommendations for LHC Run II",
J. Phys. Technical Design Report CMS **G43** (2016) p. 023001,
DOI: 10.1088/0954-3899/43/2/023001, [arXiv:1510.03865].

[67] Richard D. Ball et al., "Impact of Heavy Quark Masses on Parton Distributions and LHC Phenomenology",
Nucl. Phys. Technical Design Report CMS **B849** (2011) 296–363,
DOI: 10.1016/j.nuclphysb.2011.03.021, [arXiv:1101.1300].

[68] Johan Alwall et al., "MadGraph 5: going beyond",
Journal of High Energy Physics, Technical Design Report CMS (2011) p. 128,
DOI: 10.1007/JHEP06(2011)128.

[69] Torbjörn Sjöstrand, Stephen Mrenna and Peter Skands,
"A brief introduction to {PYTHIA} 8.1", *Computer Physics Communications*,
Technical Design Report CMS **178** (2008) 852–867,
DOI: https://doi.org/10.1016/j.cpc.2008.01.036.

[70] CMS Collaboration, "Event generator tunes obtained from underlying event and multiparton scattering measurements", *The European Physical Journal C*,
Technical Design Report CMS **76** (2016) p. 155,
DOI: 10.1140/epjc/s10052-016-3988-x.

[71] CMS Collaboration, *Standard Model Cross Sections for CMS at 13 TeV*, 1997, URL: https://twiki.cern.ch/twiki/bin/viewauth/CMS/StandardModelCrossSectionsat13TeV.

[72] A. Mitov M. Czakon, *NNLO+NNLL top-quark-pair cross sections*, 2013, URL: https://twiki.cern.ch/twiki/bin/view/LHCPhysics/TtbarNNLO.

[73] Michal Czakon and Alexander Mitov, "Top++: A Program for the Calculation of the Top-Pair Cross-Section at Hadron Colliders", *Comput. Phys. Commun.* Technical Design Report CMS **185** (2014) p. 2930, DOI: 10.1016/j.cpc.2014.06.021, [arXiv:1112.5675].

[74] Daniel Wicke, "Properties of the Top Quark", *Eur. Phys. J.* Technical Design Report CMS **C71** (2011) p. 1627, DOI: 10.1140/epjc/s10052-011-1627-0, [arXiv:1005.2460].

[75] J. Alwall et al., "The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations", *Journal of High Energy Physics*, Technical Design Report CMS (2014) p. 79, DOI: 10.1007/JHEP07(2014)079.

[76] Rikkert Frederix and Stefano Frixione, "Merging meets matching in MC@NLO", *Journal of High Energy Physics*, Technical Design Report CMS (2012) p. 61, DOI: 10.1007/JHEP12(2012)061.

[77] CMS Collaboration, *Top Quark Physics Summary Figures*, 2016, URL: https://twiki.cern.ch/twiki/bin/view/CMSPublic/PhysicsResultsTOPSummaryFigures.

[78] Robert D. Cousins, *Generalization of Chisquare Goodness-of-Fit Test for Binned Data Using Saturated Models, with Application to Histograms*, 2013, URL: http://www.physics.ucla.edu/~cousins/stats/cousins_saturated.pdf.

[79] M Cinquilli et al., "CRAB3: Establishing a new generation of services for distributed analysis at CMS", *Journal of Physics: Conference Series*, Technical Design Report CMS **396** (2012) p. 032026, URL: http://stacks.iop.org/1742-6596/396/i=3/a=032026.

[80] R. Brun and F. Rademakers, "ROOT: An object oriented data analysis framework", *Nucl. Instrum. Meth.* Technical Design Report CMS **A389** (1997) 81–86, DOI: 10.1016/S0168-9002(97)00048-X.

[81] Joram Berger et al., "ARTUS - A Framework for Event-based Data Analysis in High Energy Physics", Technical Design Report CMS (2015), [arXiv:1511.00852].

[82] Tim Bray, *The JavaScript Object Notation (JSON) Data Interchange Format*, Internet-Draft, Internet Engineering Task Force, 2017, 14, URL: https://datatracker.ietf.org/doc/html/draft-ietf-jsonbis-rfc7159bis-03.

[83] Patrick Fuhrmann, *dCache, the Overview*, tech. rep., DESY, 2011, URL: https://www.dcache.org/manuals/dcache-whitepaper-light.pdf.

[84] Frank Fischer, *Large-scale dynamic Provisioning of Compute Resources for High Energy Physics using Cloud Technology*, Karlsruhe, 2017, URL: https://ekp-invenio.physik.uni-karlsruhe.de/record/48879.

Acknowledgments

This PhD thesis is based on data taken with the CMS experiment, provided with colliding beams from the LHC. I want to acknowledge the wonderful work of the many physicists and engineers involved in running and constructing these complex machines as well as the support by the many institutions and countries that made it possible to built and maintain such machines in the middle of Europe at CERN.
Many thanks to both my supervisors Prof. Dr. Günter Quast and Dr. Roger Wolf. I really appreciate the faith they put in me by giving me the chance to do this PhD at KIT. Thank you both for the support over all the time and the detailed comments and suggestions on my thesis.
I'd like to thank my colleagues Stefan Wunsch, Nicola Zäh, Sebastian Wozniewski, Artur Akhmetshin, Dr. Fabio Colombo, Emanuel Pfeffer, Marcus Schmitt, Dr. Stefan Wayand, Dr. Dominik Haitz, Dr. Thomas Hauth, Thomas Berger, Josry Metwally, Daniel Savoiu and Dr. Klaus Rabbertz for the enriching discussions and important inputs.
Special thanks goes to Rene Caspart and Dr. Andrew Gilbert for sharing their deep knowledge about the statistical combination and the di-tau analysis. I appreciate the common work on the Standard Model $H \to \tau\tau$ analysis with my colleague in Aachen Alexander Nehrkorn and the productive collaboration with Dr. Alexei Raspereza.
Many thank go to Dr.Manuel Zeise for his patience in introducing me to particle physics analysis and to Dr. Joram Berger, Dr. Oliver Oberst and Dr. Fred Stober for their support in my first days at the institute.
Thanks go to the computing team of Matthias Schnepf, Frank Fischer, Christoph Heidecker, Günter Erli, Dr. Manuel Giffels and Dr. Max Fischer that made the outstanding computing resources accessible.
Thanks to especially Philip Coleman Harris, Raffaele Angelo Gerosa and in general all the people at MIT and CERN for the common work on the $\not\!\!E_{\mathrm{T}}$. A thank you goes also to Vivian O'Dell for the nice introduction to the CMS data acquisition system and the collaboration of the whole CMS operating crew.
Thanks to Bärbel Bräunling for the all-time support in all circumstances. Thanks goes to the graduate schools KSETA and KCETA and the people behind it, who support the research efforts of young physicists with budget and knowledge.
I am indebted to my wife Sabrina who joyfully followed me abroad, cared about our family all the time and showed great understanding for the peculiar working hours necessary to successfully write a PhD thesis.