Dissertation

Search for the Standard Model Higgs Boson in Hadronic $\tau^+\tau^-$ Decays with the ATLAS Detector

---

von

Daniele Zanzi

eingereicht an der

Fakultät für Physik

der

Technischen Universität München

erstellt am

Max–Planck–Institut für Physik

(Werner–Heisenberg–Institut)

München

May 2014

MAX-PLANCK-GESELLSCHAFT

Fakultät für Physik der Technischen Universität München

Max–Planck–Institut für Physik
(Werner–Heisenberg–Institut)

Search for the Standard Model Higgs Boson
in Hadronic $\tau^+\tau^-$ Decays
with the ATLAS Detector

Daniele Zanzi

Vollständiger Abdruck der von der Fakultät für Physik der Technischen Universität München zur Erlangung des akademischen Grades eines
               Doktors der Naturwissenschaften (Dr. rer. nat.)
genehmigten Dissertation.

Vorsitzender: Univ.-Prof. Dr. A. Ibarra
Prüfer der Dissertation:

      1. Priv.-Doz. Dr. H. Kroha
      2. Univ.-Prof. Dr. L. Oberauer

Die Dissertation wurde am 05.05.2014 bei der Technischen Universität München eingereicht und durch die Fakultät für Physik am 13.05.2014 angenommen.

Abstract

The discovery of a Higgs boson in di-boson decays, the evidence of its decays into fermion pairs and the compatibility of its measured properties with the Standard Model predictions support the electroweak symmetry breaking mechanism of the Standard Model.

The topic of this thesis is the search for the Higgs boson decays into a pair of τ leptons, important for probing the coupling of the Higgs boson to fermions. The search is performed in final states where both τ leptons decay hadronically using $4.6\,\mathrm{fb}^{-1}$ and $20.3\,\mathrm{fb}^{-1}$ of data collected by the ATLAS detector in proton-proton collisions at the Large Hadron Collider at center-of-mass energies of 7 and 8 TeV, respectively.

The signal selection is optimised for events with highly boosted Higgs bosons produced via gluon fusion with additional jet or via vector boson fusion. In order to reduce systematic uncertainties, the major background contributions from $Z \to \tau\tau$ and multi-jet production processes have been measured using signal-free control data samples. The trigger and identification efficiencies of hadronically decaying τ leptons have been measured using dedicated calibration data samples.

An excess of events above the predicted background is found with observed (expected) significance of $2.3\,\sigma$ $(2.1\,\sigma)$. The observation is compatible with the production of a Higgs boson with mass of 125 GeV and corresponds to a Higgs boson production cross section times branching ratio of $\mu = 1.2\pm0.4(\mathrm{stat})^{+0.5}_{-0.4}(\mathrm{syst})$ times the Standard Model prediction. The Higgs boson mass determined from the observed invariant $\tau\tau$ mass distribution is 125^{+16}_{-7} GeV. The sensitivity of this analysis and its results are compatible with the ones obtained using a multivariate approach. The combination of searches for $H \to \tau^+\tau^-$ decays in fully leptonic and semi-leptonic and fully hadronic $\tau^+\tau^-$ final states in $\sqrt{s} = 8$ TeV data with multivariate analyses leads to a $4.1\,\sigma$ evidence for Higgs boson decays into τ lepton pairs.

Acknowledgements

Contents

Introduction

This thesis is about the search for the Higgs boson of the Standard Model of particle physics in its decay into a pair of hadronically decaying τ leptons with the ATLAS detector at the Large Hadron Collider (LHC) at CERN.

The Higgs boson is a consequence of electroweak symmetry breaking introduced in the Standard Model to give masses to the fundamental particles and to ensure the consistency of the Standard Model at high energies. While this theory has successfully passed many experimental tests, including the correct prediction of the W and Z boson masses, the last missing piece has been for a long time the direct detection of the Higgs boson. Finally in 2012, this particle has been discovered at the LHC in decays into pairs of electroweak vector bosons. All properties of the new particle are so far in agreement with the Standard Model predictions.

The search presented here is important because the Higgs boson decays into fermion pairs had not been discovered yet. These decays are required for the new particle to be also responsible for the generation of the fermion masses. Among the fermionic decays, the $\tau^+\tau^-$ channel is the most sensitive for Higgs boson searches. It leads to three distinct final states depending on the decays of the τ lepton into leptons or hadrons.

The final states studied in this thesis contain two hadronically decaying τ leptons. Compared to the other $\tau^+\tau^-$ channels, the sensitivity of this final state depends on opposing features. On the positive side, the signal acceptance profits from the large hadronic τ decay branching ratio and the relatively high $\tau^+\tau^-$ invariant mass resolution, since there are only two neutrinos in the final state. On the negative side, this channel is harmed by the high rate of background of QCD jet production mimicking the hadronically decaying τ leptons. In order to suppress this background, the search is performed in events where the Higgs boson is produced together with two highly energetic jets emitted in the proton beam directions characteristic for vector boson fusion Higgs production or where the $\tau^+\tau^-$ pair has high transverse momentum perpendicular to the beams characteristic for gluon fusion Higgs production in association with a recoiling jet.

Important for this search was the optimisation of the trigger and the measurement of the trigger and hadronic τ decay identification efficiencies.

The thesis is structured as follows. Chapter 1 introduces to the theoretical foundations for the analysis and summarises the status of the Higgs boson searches and measurements of its properties. Chapter 2 describes the LHC accelerator and the ATLAS detector. Chapter 3 discusses the trigger and reconstruction algorithms for hadronically decaying τ leptons and the measurements of their performance. Chapter 4 describes the selection

of Standard Model $H \rightarrow \tau^+\tau^-$ decays in fully hadronic final states. The results are summarised in Chapter 5.

Chapter 1

The Higgs boson of the Standard Model

1.1   The Standard Model and the EWSB

The Standard Model (SM) of particle physics [1–4] describes the known fundamental particles, fermions and bosons, and their interactions (see Fig. 1.1).

Except for the Higgs boson, which has spin zero, all bosons in the SM are vector fields with spin one which mediate the fundamental interactions between the spin 1/2 fermions. The massless photon and the massless gluons mediate the electromagnetic and the strong force, respectively, while the massive $W^{\pm}$ and Z bosons mediate the weak interaction.

The only scalar boson predicted by the SM, the Higgs boson, has recently been discovered in its decays into photon or weak boson pairs [5, 6]. The search for it in decays into τ lepton pairs is the topic of this thesis.

The interactions of the SM are described by a local $SU(3) \times SU(2)_L \times U(1)_Y$ gauge symmetry. The non-Abelian $SU(3)$ gauge symmetry determines the strong interaction between quarks and eight massless gluons, while the $SU(2)_L \times U(1)_Y$ gauge symmetry governs the electroweak interaction mediated by the photon and the $W^{\pm}$ and Z bosons. The local gauge theories predict massless intermediate vector bosons. The global $SU(2)$ symmetry, together with parity violation, prohibits fermion masses.

In the SM, the masses of the fundamental fermions and bosons are generated by spontaneous breaking of the electroweak $SU(2)_L \times U(1)_Y$ gauge symmetry via the introduction of a scalar $SU(2)_L$ doublet field, the Higgs field, with a ground state invariant only under the electromagnetic $U(1)_{\text{EM}}$ and the strong $SU(3)$ gauge symmetry.

The problem that at first prevented the application of spontaneous symmetry breaking (SSB) in particle physics is the Goldstone theorem [7]. This theorem states that if in a relativistic quantum field theory, like the SM, the field equations are covariant under a continuous symmetry, then either the ground state is invariant under the same symmetry or there must exist scalar particles with zero mass, the Goldstone bosons, corresponding to the broken symmetry generators. Since such scalar massless particles have not been observed, SSB in the SM seemed to violate the Goldstone theorem.

The solution first suggested by Anderson [8] based on similar effects occurring in su-

Families

	I	II	III	Bosons	

| **Quarks** | | | | | |

Table of particle properties:

Quarks (orange):

Family I, up-type:
1/2 2/3, 2.3 MeV, u, up

Family II, up-type:
1/2 2/3, 1.27 GeV, c, charm

Family III, up-type:
1/2 2/3, 173 GeV, t, top

Photon: **1** 0, 0, γ, photon

Higgs: **0** 0, 126 GeV, H, Higgs

Family I, down-type:
1/2 -1/3, 4.8 MeV, d, down

Family II, down-type:
1/2 -1/3, 95 MeV, s, strange

Family III, down-type:
1/2 -1/3, 4 GeV, b, bottom

Gluon: **1** 0, 0, g, gluon

Leptons (violet):

Family I, charged:
1/2 1, 0.511 MeV, e, electron

Family II, charged:
1/2 1, 105 MeV, μ, muon

Family III, charged:
1/2 1, 1.777 GeV, τ, tau

Z boson: **1** 0, 91.2 GeV, Z

Family I, neutrino:
1/2 0, 0, ν_e, e neutrino

Family II, neutrino:
1/2 0, 0, ν_μ, μ neutrino

Family III, neutrino:
1/2 0, 0, ν_τ, τ neutrino

W boson: **1** ±1, 80.4 GeV, $W^\pm$

Legend box: **Spin** Charge / **Mass** / **Symbol** / Name

Figure 1.1: Overview of the fundamental particles of the SM. For each particle, its mass, spin and electric charge quantum number are indicated. Fermions, spin 1/2 particles, are arranged in three families, with a pair of leptons (violet) and a pair of quarks (orange) and the same quantum numbers, but different masses. Bosons (green) include the vector bosons mediating the fundamental interactions and the scalar Higgs boson. The massless photon mediates the electromagnetic force between all particles with electric charge. The eight massless gluons mediate the strong interaction between quarks and gluons. The massive $W^\pm$ and Z bosons mediate the weak interaction. The massive scalar Higgs boson couples to all massive particles. For each fermion there is an anti-fermion, a particle with the same mass, but opposite charges. For the electrically neutral neutrinos it is not yet known whether they are different from their anti-particles.

perconductors is that the mechanisms of giving masses to gauge bosons and of preventing the appearance of massless scalar particles are related. In 1964/65 this idea has been applied to particle physics by Brout and Englert [9], Higgs [10, 11] and Kibble, Guralnik and Hagen [12] in what is now known as the Brout-Englert-Higgs (BEH) mechanism. They proposed that in relativistic gauge theories the massless Goldstone bosons can be translated into the longitudinal polarisation states of massive gauge bosons. In 1967, Weinberg [2] and Salam [3] applied the BEH mechanism to the $SU(2)_L \times U(1)_Y$ gauge theory of the electroweak interaction introduced by Glashow [1].

In the $SU(2)_L \times U(1)_Y$ gauge theory, the Higgs field is a complex scalar weak isospin doublet with hypercharge $Y = +1$ and no electric charge. A non-vanishing vacuum expectation value $v = \langle \phi \rangle_0 / \sqrt{2}$ of the Higgs field ϕ spontaneously breaks the electroweak gauge symmetry to the remaining electromagnetic (EM) gauge symmetry $U(1)_{\mathrm{EM}}$ giving masses to the $W^\pm$ and Z bosons while leaving the photon massless.

Via the $SU(2)_L$ gauge symmetry, the three massless Goldstone boson excitations of the symmetry breaking ground state are transformed into the longitudinal polarisation states of the massive $W^\pm$ and Z fields. The fourth excitation of the ground state is the massive scalar Higgs boson. Its couplings to the weak bosons are proportional to the square of the weak boson masses which are predicted to be $m_W = \frac{1}{2}vg$ and $m_Z = \frac{1}{2}v\sqrt{g^2 + g'^2}$, where g and g' are the $SU(2)_L$ and $U(1)_Y$ gauge coupling constants, respectively.

The BEH mechanism also generates fermion masses via the introduction of Yukawa-type interactions between the fermions and the Higgs field with coupling constants $g_f = \sqrt{2}m_f/v$ proportional to the fermion masses. The test of the existence of such Yukawa couplings of the SM fermions to the Higgs field is subject of this thesis.

The Higgs boson mass m_H and the fermion masses are not predicted by the SM.

1.2   Indirect Constraints on the Higgs Boson Mass

Upper and lower bounds on the Higgs boson mass follow from consistency requirements in the SM. An upper bound is introduced by the requirement of unitarity [13–16], for instance for the amplitude of longitudinal W scattering $W_L W_L \to W_L W_L$, which would grow proportionally to the center-of-mass energy of the scattering, eventually violating unitarity, if there is no contribution from the exchange of a virtual Higgs boson with $m_H \lesssim 800$ GeV or new physics beyond the SM at the TeV scale. For this reason, direct SM Higgs boson searches have been devised to explore the mass range up to 1 TeV.

Other theoretical arguments concerning the finiteness of the Higgs self-coupling λ under radiative corrections and the stability of the Higgs ground state [17–21] lead to even more stringent upper and lower bounds. Quantum effects of Higgs boson loops lead to a divergence of λ at high energies if the Higgs boson is heavier than about 160 GeV [22]. If the Higgs boson is too light, quantum corrections from top quark loops drive λ to negative values making the SM vacuum unstable. To ensure vacuum stability, the Higgs boson has to be heavier than 129.4 ± 1.8 GeV [23].

Indirect constraints on the Higgs boson mass can be derived from the precision measurements of electroweak observables which depend logarithmically on m_H via virtual

Higgs boson radiations. The constraint on the Higgs boson mass from the combined electroweak precision measurements at LEP, SLC, Tevatron and LHC is $m_H = 94^{+25}_{-22}$ GeV [24] as shown in Fig. 1.2 which is in agreement with the direct measurements of the Higgs boson mass performed by the ATLAS and CMS experiments [5,6] within $1.3\,\sigma$.
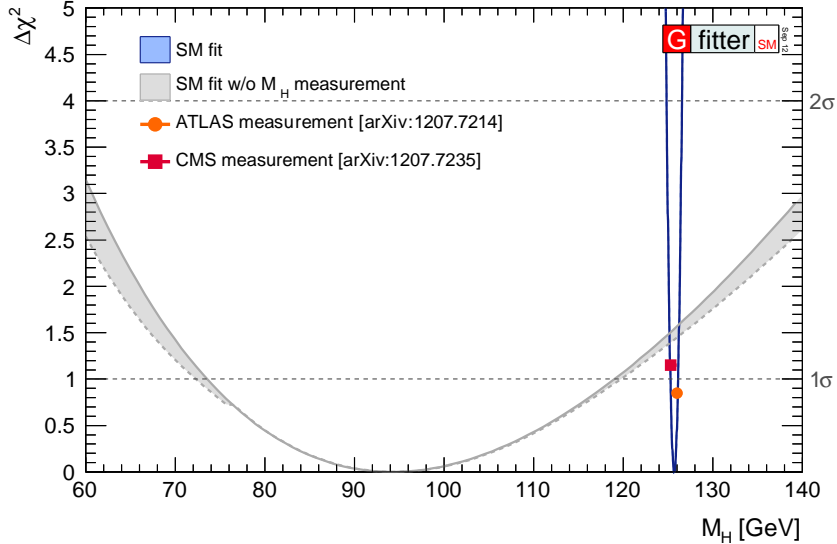

Figure 1.2: $\Delta\chi^2$ of the global fit to the electroweak precision measurements as a function of the Higgs boson mass [24]. The data points show the direct measurements of m_H by the ATLAS and CMS experiments.

1.3   Higgs Boson Production and Decay Modes

The SM Higgs boson couplings to fermions, proportional to the fermion masses, and to the weak vector bosons, proportional to the square of the vector boson masses, as well as the Higgs boson self-coupling are illustrated in Table 1.1 [25]. The dominant couplings are to the top quark and to the W and Z bosons.

The main Higgs boson production processes at hadron colliders are gluon fusion (ggF), the vector boson fusion (VBF), associated production with a vector boson (VH) and the associated production with a top quark pair ($t\bar{t}H$) as illustrated in Table 1.2.   Fig. 1.3 shows the cross sections of these processes in pp collisions as a function of the Higgs boson mass at a center-of-mass energy of 8 TeV [26].

Gluon fusion is the dominant Higgs production process at the LHC with the main contribution from the top quark loop. The cross section depends on the distribution function (PDF) of the gluon momentum fraction in the proton and on QCD radiative corrections. The QCD corrections have been computed in NLO perturbation theory [28,29] with NNLO contributions calculated in the large-m_t approximation[1] [30–37] which is expected to deviate from the complete result by less than 1% for $m_H \leq 300$ GeV [38–43].

---

[1]In the large-m_t approximation, the $gg \to H$ loop is described by an effective point interaction between

Table 1.1: SM Higgs boson couplings [25].
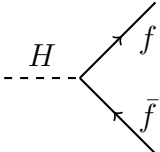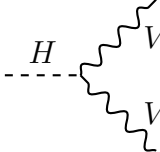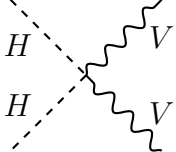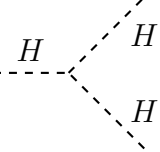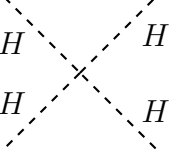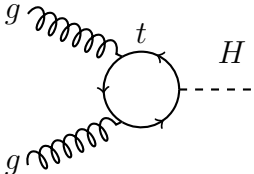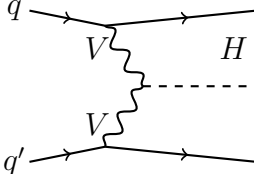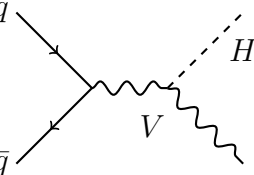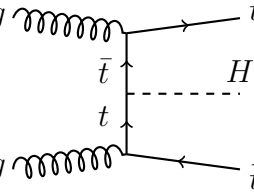
Yukawa coupling to fermions	 $g_{Hf\bar{f}} = \sqrt{2}\frac{m_f}{v}$	
Couplings to weak vector bosons ($V = W, Z$)	 $g_{HWW} = \frac{2m_W^2}{v}$ $g_{HZZ} = \frac{m_Z^2}{v}$	 $g_{HHWW} = \frac{m_W^2}{v^2}$ $g_{HHZZ} = \frac{m_Z^2}{2v^2}$
Higgs self-coupling	 $g_{HHH} = \frac{m_H^2}{2v}$	 $g_{HHHH} = \frac{m_H^2}{8v^2}$

Table 1.2: SM Higgs boson production modes at the LHC. The predicted cross sections for $m_H = 125$ GeV at $\sqrt{s} = 8$ TeV [26,27], calculated to the higher orders in perturbation theory as indicated, are given.

Production process	LO diagram	Cross section [pb]	Order in perturbation theory
ggF		$19.52^{+14.7\%}_{-14.7\%}$	NNLO+NNLL QCD, NLO EW
VBF		$1.578^{+2.8\%}_{-3.0\%}$	NLO QCD+EW, approx. NNLO QCD
VH		WH: $0.697^{+3.7\%}_{-4.1\%}$ ZH: $0.394^{+5.1\%}_{-5.0\%}$	NNLO QCD, NLO EW
$t\bar{t}H$		$0.130^{+11.6\%}_{-17.1\%}$	NLO QCD
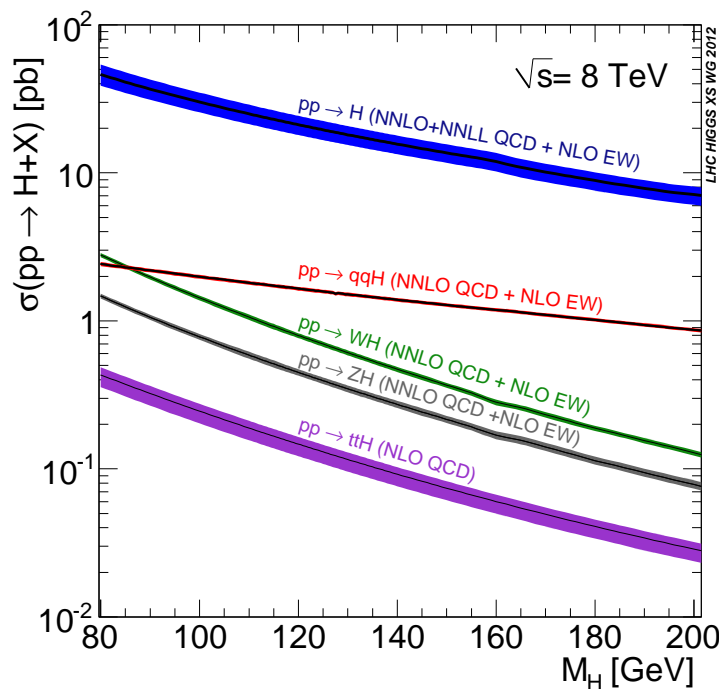
Figure 1.3: SM Higgs boson production cross sections in pp collisions at $\sqrt{s} = 8$ TeV as a function of the Higgs boson mass [26]. The bands represents the theoretical uncertainties.

It is necessary to take into account higher-order QCD corrections in the ggF cross section prediction because of the slow convergence of the perturbative expansion in α_S. The LO cross section is increased by 80-100% at NLO and by an additional 25% at NNLO. Fig. 1.4 shows the SM Higgs boson production cross section for ggF in different orders of perturbation theory as a function of the renormalisation scale μ_R [44]. The large NLO and NNLO corrections are related to the large scale dependence of the LO and NLO cross sections. A moderate scale dependence is reached only at NNLO suggesting that the size of the N^3LO corrections should be smaller. However, approximated N^3LO calculations [44, 45] indicate that the cross section may still increase by as much as 17% [44].

Other sizeable corrections are due to soft-gluon radiation computed in NNLL approximation [46], and to EW corrections computed at NLO [47–51].

Because of the large contribution of higher order QCD processes to the production cross section, many Higgs boson searches, including the one presented in this thesis, take into account the production of one or two additional jets. In such events, the Higgs boson recoils against the jet(s) acquiring significant momentum. The high transverse momentum p_T^{Higgs} of the Higgs boson in higher order ggF production provides strong discrimination between signal and background. The p_T^{Higgs} dependence of the cross section has been computed at NNLL+NNLO for the correct top and bottom quark masses [52]. Heavy quark mass effects are important for the differential ggF cross section as a function of p_T^{Higgs} although the total cross section is little affected. Fig. 1.5 shows the impact of the top and bottom quark loop contributions on the Higgs boson transverse momentum spectrum at

the gluons and the Higgs boson and contributions from W boson and b quark loops are neglected.

Higgs hadron-level cross section
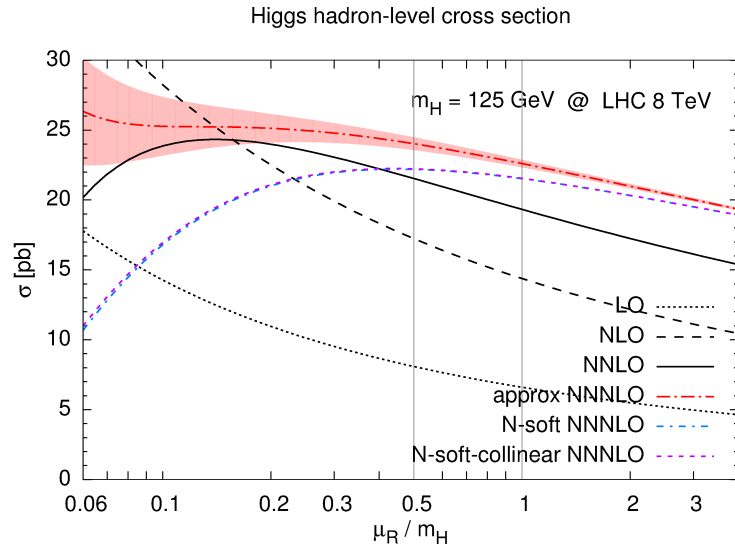
Figure 1.4: Gluon fusion production cross section of the SM Higgs boson with $m_\mathrm{H} = 125$ GeV at $\sqrt{s} = 8$ TeV as a function of the renormalisation scale μ_R in different orders of perturbation theory [44].

NLO compared to the large-m_t approximation. The bottom quark contribution is important in the low-p_T region, while at high-p_T the top quark contribution has a significant impact and the large-m_t approximation becomes invalid. The correct description of the Higgs boson transverse momentum distribution is important for the analysis presented here because the high momentum Higgs boson production in ggF is used in the event selection.

VBF production, even though it has an order of magnitude smaller cross section than ggF production, is very important for discriminating signal from background in pp collisions due to its characteristic final state. The two quarks in the initial state in this case each radiate vector bosons which annihilate with each other producing the Higgs boson. In the final state, the two quarks hadronize into two jets emitted in the forward regions of the detector close to the proton beams while the Higgs boson decays in the central part of the detector. Hadronic emission around the Higgs boson is suppressed due to the fact that the two quarks are not connected by colour fields and the hadron production, therefore, mostly develops along the quark directions. This provides an additional signature for the VBF event selection, which can hardly be imitated by other SM processes.

VBF at LO is an electroweak process. QCD corrections are only on the order of 5%. The cross section has been computed with full NLO QCD and EW corrections [53, 54] and approximate NNLO QCD corrections [55].

The ggF and VBF processes are complementary for the test of the SM because ggF is determined by the Higgs Yukawa couplings to fermions while VBF depends on the Higgs boson couplings to the weak vector bosons. It is, therefore, important to explore both production processes in order to test the role of the Higgs field in the SM.
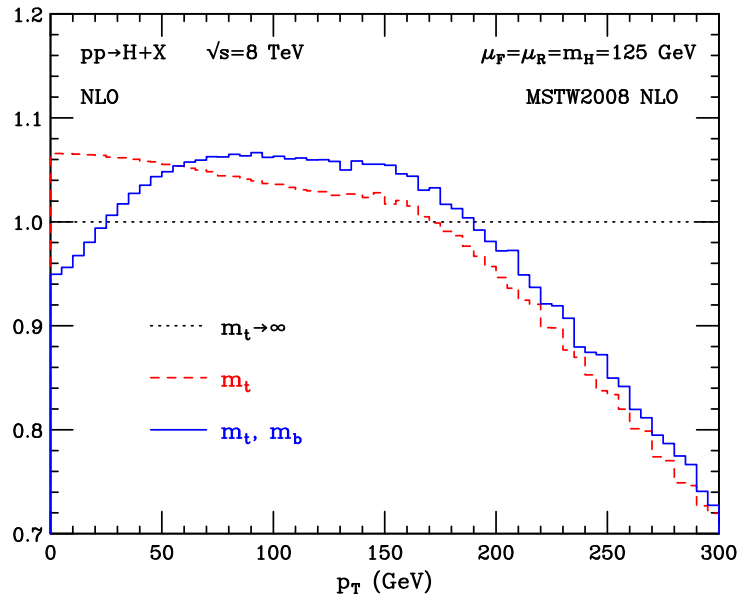
Figure 1.5: Transverse momentum distribution of the SM Higgs boson with $m_H = 125$ GeV at NLO in the large-m_t approximation (dotted black line), for the exact top quark loop contribution (red dashed line) and with both top and bottom quark loop contributions (solid blue line) [52].

The cross section for Higgs boson production in association with a weak vector boson, also known as Higgs-strahlung, has been computed including NNLO QCD [56, 57] and NLO EW corrections [58]. The production in association with a $t\bar{t}$ pair has the smallest cross section, but is relevant because it allows for the direct measurement of the top quark Yukawa coupling. Only NLO QCD corrections have been calculated so far [59–62].

The branching ratios (BR) of the Higgs boson decays into different final states are determined by the couplings of the Higgs boson to the final state particles. As shown in Fig. 1.6 [27], the SM Higgs boson mainly decays into two weak vector bosons for m_H above the threshold. For m_H below about 140 GeV, these decay modes are suppressed since only one of the two bosons can be produced on-shell.

In the low-mass region around the measured Higgs boson mass of 125 GeV, the main decay mode is into $b\bar{b}$ pairs because b quarks are the heaviest particles which can be pair-produced on-shell in Higgs decays. The next largest BR into fermions is for decays into $\tau^+\tau^-$. Decays into $c\bar{c}$ are very difficult to distinguish from QCD di-jet events. Decays into $\mu^+\mu^-$ have a very small BR, but provide high $\mu^+\mu^-$ invariant mass resolution.

The observable bosonic decay modes in the low-mass region are $H \to WW^*, ZZ^*$ and $\gamma\gamma$. BR($H \to WW^*$) is larger than the BR($H \to ZZ^*$) because of the twice larger number of degrees of freedom compared to Z and because of the smaller W mass and, therefore, the larger phase space for the decay. BR($H \to \gamma\gamma$) is very small because the decay occurs only in second order perturbation theory through W boson and top quark loops. However, this decay mode is important for Higgs boson searches due to the clear signature of two highly energetic photons and the high $\gamma\gamma$ invariant mass resolution. The
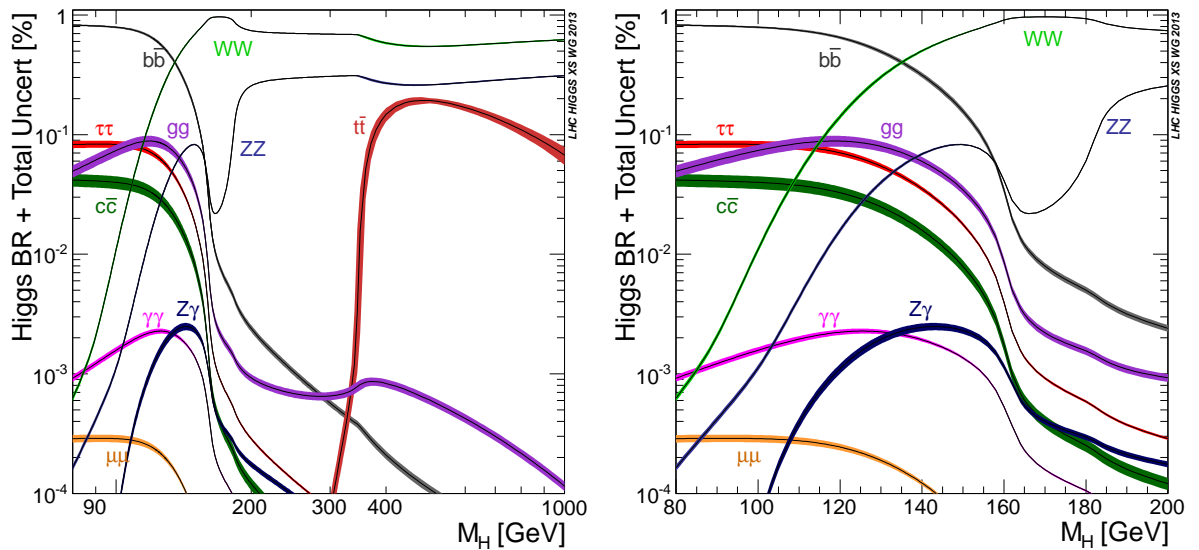
Figure 1.6: SM Higgs boson decay branching ratios as a function of the Higgs boson mass in the whole range up to 1 TeV allowed for consistency of the SM (left) and in the low mass range (right) [27]. The bands represent the theoretical uncertainties.

$H \to gg$ decay mode cannot be exploited due to the large QCD di-jet background.

1.4 Direct Higgs Boson Searches

A lower bound on the Higgs boson mass of 114.4 GeV at 95% CL has been obtained before LHC at the Large Electron-Positron collider (LEP) at CERN at center-of-mass energies of up to $\sqrt{s} = 209$ GeV [63].

At the Tevatron, the proton-antiproton collider at the Fermi National Accelerator Laboratory, direct Higgs boson searches have been carried out at a center-of-mass energy of $\sqrt{s} = 1.96$ TeV in the mass range of $90 \lesssim m_H \lesssim 200$ GeV. In 2013, an excess of events around $m_H = 125$ GeV mainly from searches in the $VH \to Vb\bar{b}$ channel has been found with a significance of $3.1\,\sigma$ [64].

The direct searches carried out at the LHC are based on data collected in 2011 at $\sqrt{s} = 7$ TeV and in 2012 at $\sqrt{s} = 8$ TeV corresponding to integrated luminosities of about $5\,\mathrm{fb}^{-1}$ and $20\,\mathrm{fb}^{-1}$, respectively. The search programme covered the mass range from the LEP lower mass limit up to about 1 TeV [65]. In the low mass range, despite the small branching ratios, the channels with highest sensitivity for the Higgs boson search are Higgs boson decays into a pair of vector bosons, namely $H \to \gamma\gamma$, $H \to ZZ^* \to l^+l^-l'^+l'^-$ ($H \to 4l$) and $H \to WW^* \to l\nu l'\nu$. The other two decay channels accessible in the low-mass range, but with lower sensitivity, are the fermionic decays $H \to \tau^+\tau^-$ and $H \to b\bar{b}$. In the high mass range, the most sensitive channels are $H \to WW \to l\nu l'\nu, l\nu qq'$ and $H \to ZZ \to l^+l^-l'^+l'^-, l^+l^-q\bar{q}$.

1.5   Higgs Boson Discovery and Properties

The discovery of a new particle compatible with the SM Higgs boson has been published by both ATLAS and CMS on July 4th, 2012 [5, 6]. The discovery is based on integrated luminosities of $4.8 - 5.1\,\text{fb}^{-1}$ at $\sqrt{s} = 7$ TeV and $5.8 - 5.3\,\text{fb}^{-1}$ at $\sqrt{s} = 8$ TeV. The most sensitive channels in both experiments are $H \to 4l$, $H \to \gamma\gamma$ and $H \to WW^* \to e\nu\mu\nu$. The significance of the excess of events observed around $m_H = 125$ GeV is above $5\,\sigma$. Since then, the discovery has been confirmed in the di-boson final states with increased precision including all data collected in 2012. In November 2013, also evidence for Higgs boson decays into fermions has been found with more than $3\,\sigma$ significance in the $H \to \tau^+\tau^-$ and $H \to b\bar{b}$ channels by both experiments compatible with the SM Higgs boson at $m_H = 125$ GeV [66–68]. The search presented here contributes to the result in the $\tau^+\tau^-$ channel.

Using the di-boson decay channels, the properties of the new particle, production and decay rates, mass, couplings to other SM particles and spin and CP quantum numbers have been measured. The results of the two experiments ATLAS and CMS [69–72] are compatible. Here, only the results published by the ATLAS collaboration are discussed.
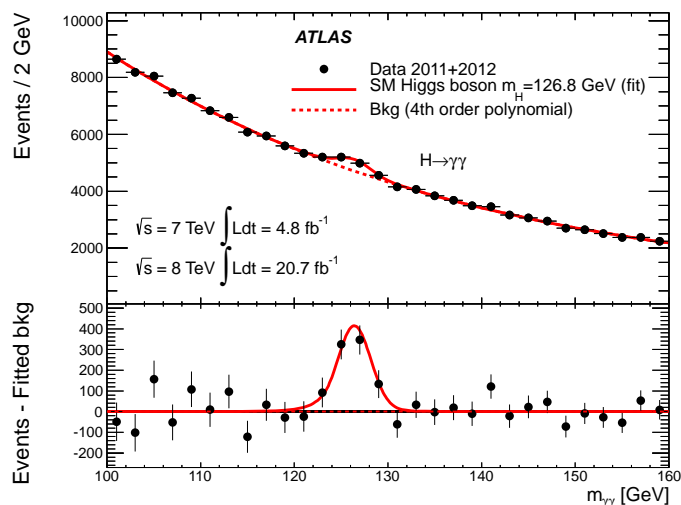
Figure 1.7: Invariant mass distribution of di-photon candidates in the ATLAS inclusive $H \to \gamma\gamma$ search. The fit of a SM Higgs boson signal (with $m_H = 126.8$ GeV) on top of a smooth background parametrisation is superimposed on the data points. The bottom panel shows the residual distribution of data with respect to the fitted background [71].

Figs. 1.7-1.9 show the mass distributions used for the evaluation of the signals in the $H \to \gamma\gamma$, $H \to 4l$ and $H \to WW^* \to l\nu l'\nu$ channels, respectively, for the full 7 and 8 TeV data sets [71]. Signals with significances of 7.4, 6.6 and 3.8 standard deviations, respectively, have been found.
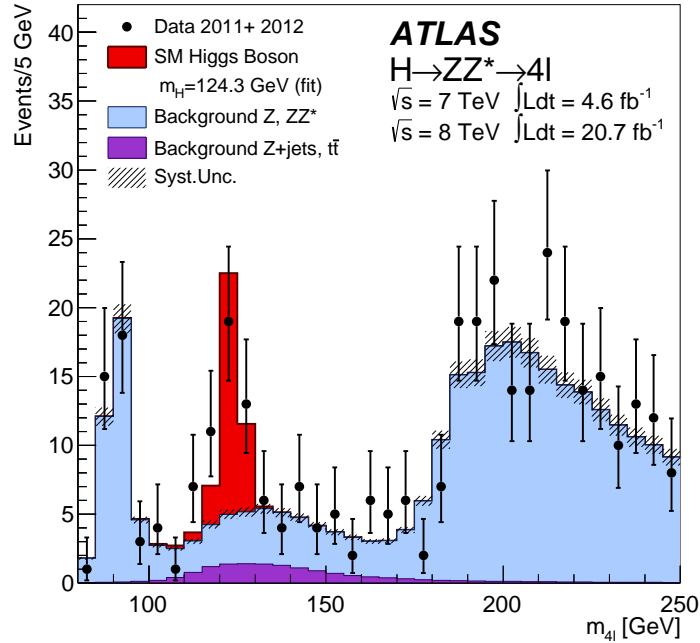
Figure 1.8: Four-lepton invariant mass distribution of data events passing the $H \rightarrow 4l$ selection together with the estimated background contributions and the fitted SM Higgs boson signal at $m_H = 124.3$ GeV [71].

The Higgs boson mass measurement is performed using the final states with high mass resolution, $H \rightarrow \gamma\gamma$ and $H \rightarrow 4l$ [71]. The mass resolution in the $H \rightarrow \gamma\gamma$ channel [73] is $1.4 - 2.5$ GeV at $m_H = 126.5$ GeV, depending on the event selection category. The precision of the measurement is limited by the photon energy scale uncertainty. In the $H \rightarrow 4l$ [74] channel a mass resolution of $1.6 - 2.4$ GeV is achieved at $m_H = 125$ GeV depending on the lepton flavours. The uncertainty in the measurement is dominated by statistics. The measured mass value is $126.8 \pm 0.2(\text{stat}) \pm 0.7(\text{syst})$ GeV in the $H \rightarrow \gamma\gamma$ channel and $123.4^{+0.6}_{-0.5}(\text{stat})^{+0.5}_{-0.3}(\text{syst})$ GeV in the $H \rightarrow 4l$ channel. The two measurements agree within $2.4\,\sigma$. The combined value is

$$m_H = 125.5 \pm 0.2(\text{stat})^{+0.5}_{-0.6}(\text{syst}) \text{ GeV}.$$

The Higgs boson production cross sections times branching ratios have been measured in the di-boson final states $H \rightarrow \gamma\gamma$, $H \rightarrow 4l$ and $H \rightarrow WW^* \rightarrow l\nu l'\nu$ and in the fermionic channels $H \rightarrow \tau^+\tau^-$ and $H \rightarrow b\bar{b}$. Combining all channels the signal strength, the ratio of the measured signal yield to the one predicted by the SM, for $m_H = 125.5$ GeV is

$$\mu = \sigma/\sigma_{\text{SM}} = 1.30^{+0.18}_{-0.17}$$

compatible with the SM prediction within 11% [75]. The signal strengths measured in the individual channels are summarised in Fig. 1.10.
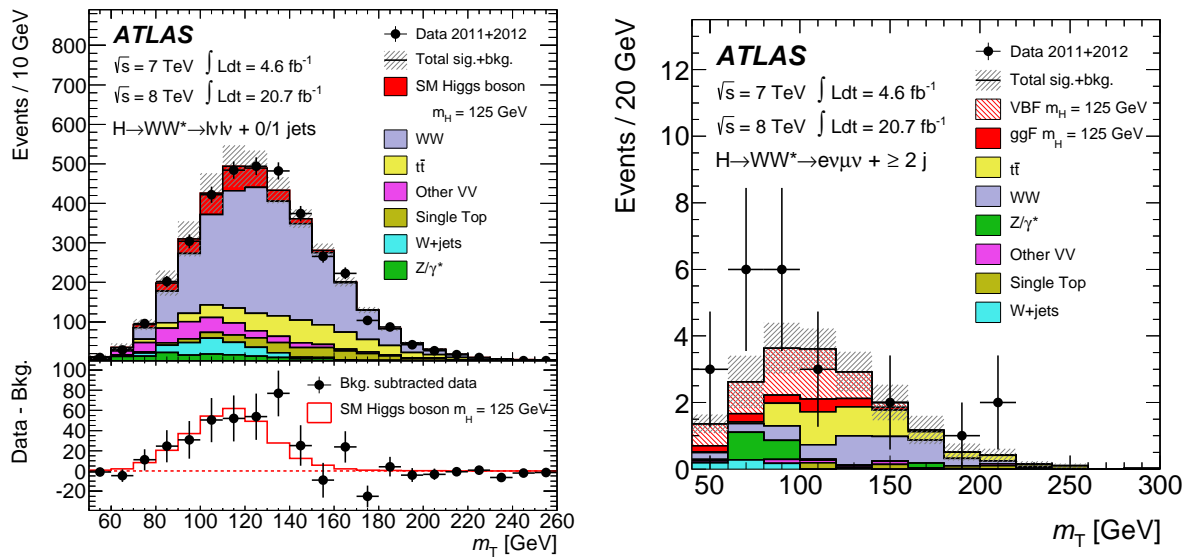
Figure 1.9: Transverse mass distributions of events after the $H \rightarrow WW^* \rightarrow l\nu l'\nu$ selection with $N_{jet} \leq 1$ (left) and $N_{jet} \geq 2$ (right) additional jets. The signal is stacked on top of the background contributions. In the right plot, VBF and ggF production contributions are shown separately. The hatched area shows the total uncertainty on the signal and background yields. The lower panel in the left plot shows the residual distribution of the data with respect to the estimated background in comparison with the m_T distribution expected for a SM Higgs boson with $m_H = 125$ GeV [71].
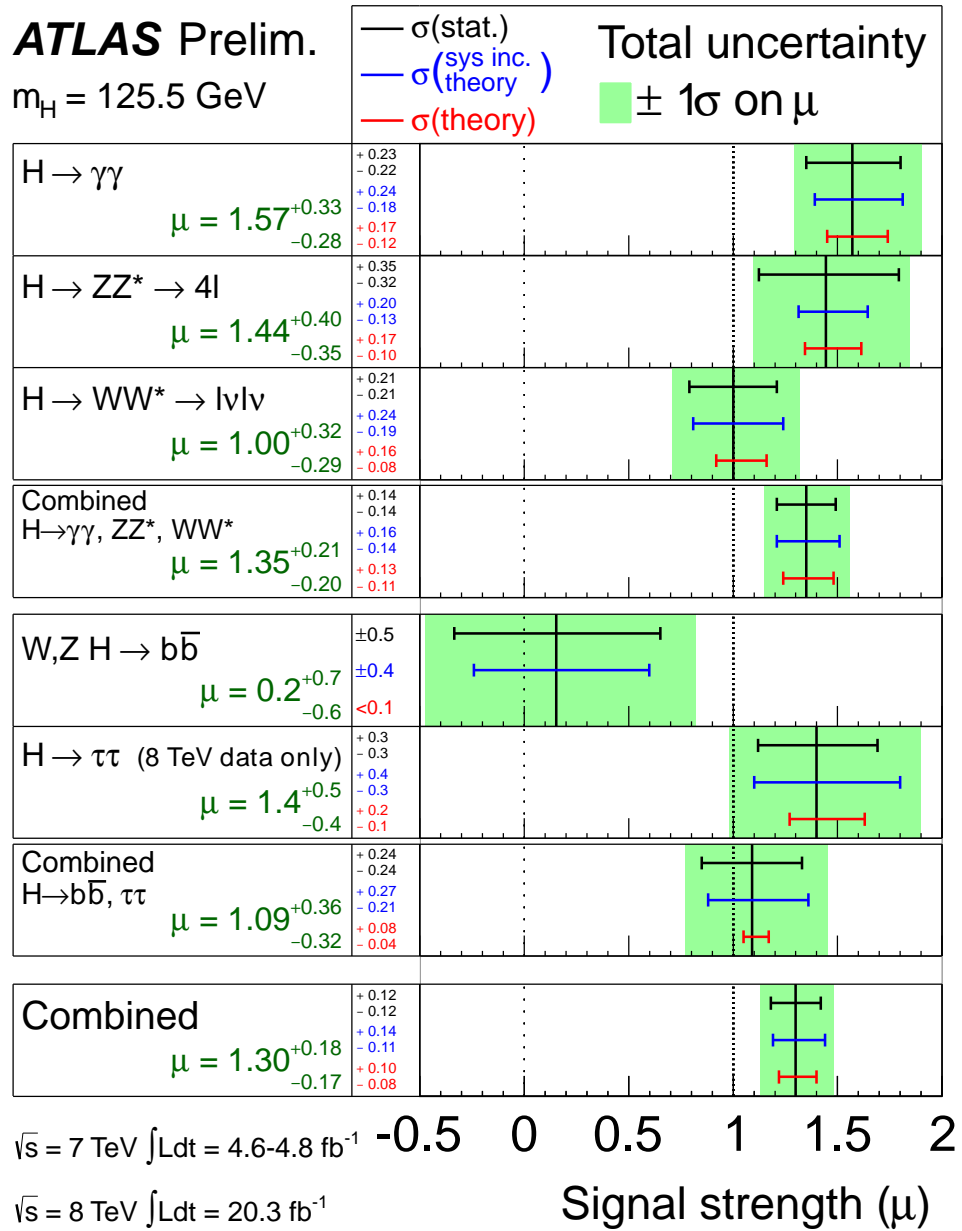
Figure 1.10: Measured signal strengths for different decay channels of the Higgs boson with mass $m_H = 125.5$ GeV [66, 67, 71] (updates in Ref. [75]).

The signal strengths μ for different production and decay modes are shown in Fig. 1.11. The vector boson mediated production processes VBF and VH are distinguished from the fermion mediated processes ggF and $t\bar{t}H$. Since the branching ratio scale factors B/B_{SM} can be different for the different channels, the contours in the μ_{VBF+VH} vs. $\mu_{ggF+t\bar{t}H}$ plane cannot be directly compared. Only the ratios $\mu_{VBF+VH} \times B/B_{SM}/\mu_{ggF+ttH} \times B/B_{SM}$ for the different channels can be compared and show good agreement among each other and with the SM expectation of unity. A $4.1\,\sigma$ evidence for VBF Higgs boson production
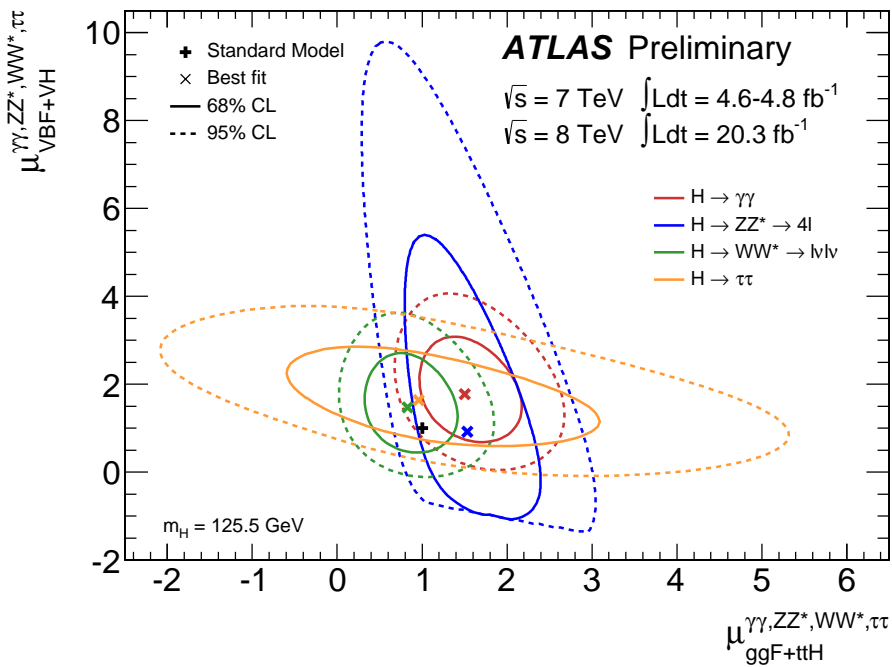
Figure 1.11: 68% and 95% CL contours of the VBF and VH vs. ggF and ttH production strengths of the Higgs boson measured in the $H \to \gamma\gamma$, $H \to ZZ^* \to 4l$, $H \to WW^* \to l\nu l'\nu$ and $H \to \tau^+\tau^-$ channels for a Higgs boson mass of 125.5 GeV [75].

is found from the measurement of [75]

$$\mu_{VBF}/\mu_{ggF+ttH} = 1.4^{+0.5}_{-0.4}(\text{stat})^{+0.4}_{-0.3}(\text{syst}).$$

The Higgs boson coupling strengths have been determined from combined fits to the measurements of the signal strengths $\mu_{ij} = \sigma_i BR_j/\sigma_i^{SM} BR_j^{SM}$ with $\sigma_i BR_j = \Gamma_i \Gamma_j/\Gamma_{tot}$ and the partial and total widths Γ_i and Γ_{tot} of the Higgs boson in the $H \to \gamma\gamma$, $H \to 4l$, $H \to WW^* \to l\nu l'\nu$, $H \to b\bar{b}$ and $H \to \tau^+\tau^-$ channels. Based on the assumption that the signals observed in the different final states originate from the same narrow resonance with $m_H = 125.5$ GeV, the partial widths Γ_i and, therefore, the squares of the couplings y_i for the production and decay modes are measured relative to the SM LO predictions $\kappa_i = y_i^2/y_{i,SM}^2 = \Gamma_i/\Gamma_{i,SM}$, where $i = W, Z, t, b, \tau$. The κ_γ and κ_g scale factors for the loop

processes $H \to \gamma\gamma$ and $gg \to H$, respectively, are functions of the other coupling scale factors as predicted by the SM.

To test the couplings to fermions and vector bosons, universal coupling scale factors $\kappa_V = y_V^2/y_{V,SM}^2$ and $\kappa_F = y_F^2/y_{F,SM}^2$ as in the SM are assumed for the weak gauge bosons and for the fermions, respectively. As shown in Fig. 1.12 [75], the results for the individual decay channels and their combination are in good agreement with the SM expectation. Among the channels considered only the $H \to \gamma\gamma$ decay is sensitive to the relative sign of κ_V and κ_F via the interference of W boson and top quark loops.
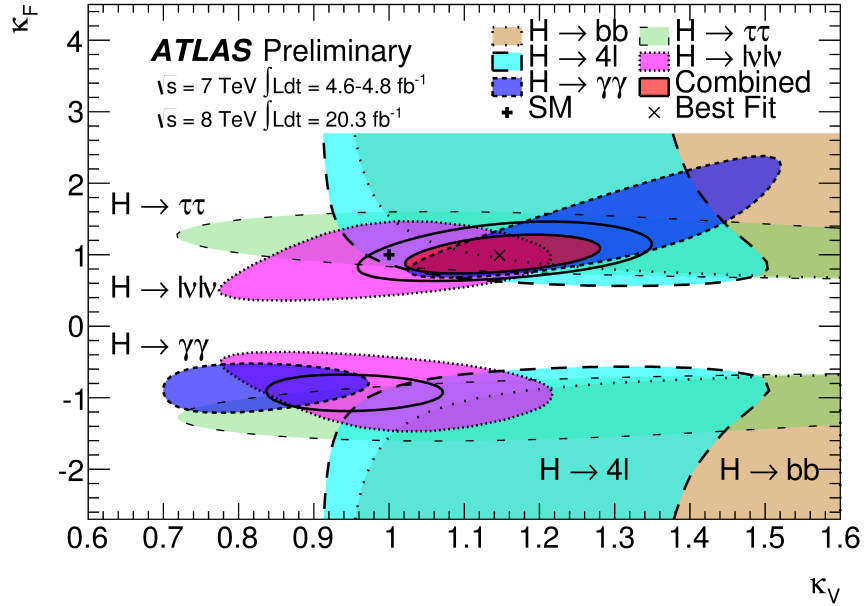
Figure 1.12: 68% CL contours of the universal scale factors κ_V and κ_F of the couplings of the Higgs boson to weak gauge bosons and fermions in the SM [75].

For the test of the spin and parity quantum numbers of the Higgs boson, the SM hypothesis $J^P = 0^+$ has been compared to the alternative hypotheses $0^-, 1^+, 1^-$ and 2^+ in fits to distributions of kinematical variables of the di-boson final states [72, 76]. Fig. 1.13 summarises the tests of the spin-parity hypotheses. In all cases, the SM $J^P = 0^+$ quantum numbers are favoured while the $0^-, 1^+, 1^-$ and 2^+ hypotheses are rejected at $97.8\%, 99.97\%, 99.7\%$ and $> 99.9\%$ CL$_s$, respectively.

The measurements of the couplings, the spin and the parity of the new particle strongly support the hypothesis that it is the SM Higgs boson. It still needs to be investigated however whether the Higgs boson properties show effects of physics beyond the SM.
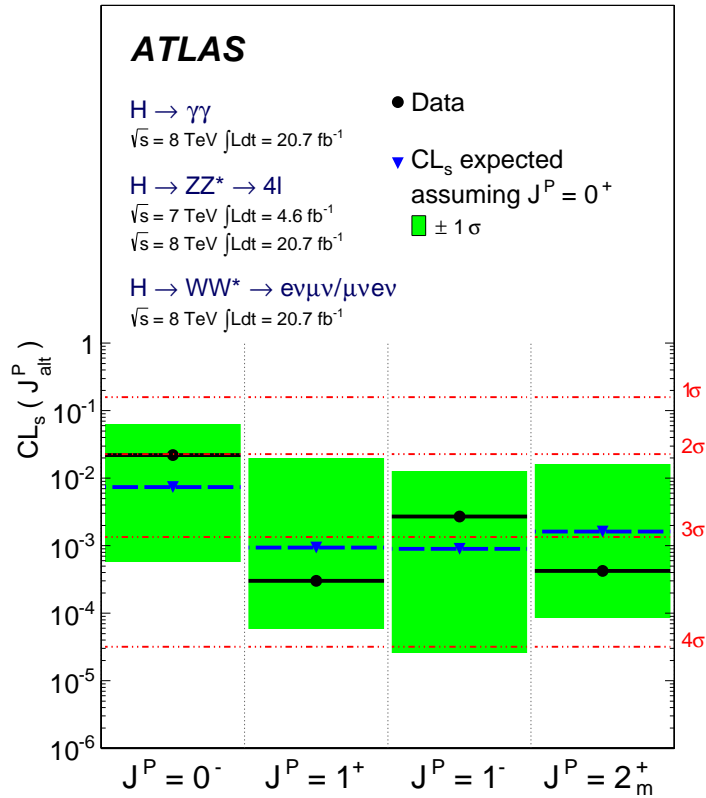
Figure 1.13: Expected (full blue triangles) and observed (full black circles) CL_s confidence levels for the rejection of alternative spin-parity hypotheses for the Higgs boson with respect to the SM $J^P = 0^+$ hypothesis [72]. CL_s is a modified confidence level to account for possible downward fluctuations of the background estimate.

Chapter 2

The Large Hadron Collider and the ATLAS Experiment

The Large Hadron Collider (LHC) started operation end of 2009 and since then delivered more than $25\,\mathrm{fb}^{-1}$ of data at the highest center-of-mass energies to the ATLAS and CMS experiments for the search for the Higgs boson and new physics beyond the Standard Model.

This chapter is devoted to an overview of the LHC (Section 2.1) and of the ATLAS detector (Section 2.2), with which the data analysed in this thesis have been recorded.

2.1 The Large Hadron Collider

The Large Hadron Collider (LHC) at CERN (Geneva, Switzerland) [77] is a proton storage ring of 27 km circumference in the tunnel of the LEP accelerator designed to provide colliding proton beams for a wide physics programme which includes the search for the Higgs boson and the measurement of its properties as well as the exploration of the TeV energy scale in the search for physics beyond the SM. The LHC has been designed for a center-of-mass energy of $\sqrt{s} = 14$ TeV and a maximum instantaneous luminosity of $\mathcal{L} = 10^{34}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$. The proton beams are kept in their orbit by superconducting dipole magnets providing a magnetic field of up to 8 T.

In 2012, two years after the start of operations, the LHC collided protons at the record center-of-mass energy of $\sqrt{s} = 8$ TeV and already reached almost the design luminosity (see Table 2.1 for details).

The acceleration of protons to such high energies is achieved by a complex chain of accelerators, as sketched in Fig. 2.1, where the energy and intensity of the beams is increased in steps. In the first step, the linear accelerator LINAC 2 accelerates the protons to 50 MeV, followed by then three synchrotrons, the Booster, the Proton Synchrotron (PS) and the Super Proton Synchrotron (SPS), which store the protons and accelerate them to 1.4, 25 and 450 GeV, respectively. In these rings, the proton beams are bunches to provide stable beams and collisions with high luminosity. Once the protons have been injected into the LHC, their energies are ramped up to the collision energy where they are collided. One fill of the LHC lasts for several hours, until the beam intensity has degraded too much and the beams are dumped on a dedicated target.

Table 2.1: Parameters of the *pp* collisions delivered by LHC in 2011 and 2012. The average number of interactions per bunch crossing is given by the mean of the Poisson distribution of the number of inelastic interactions per bunch crossing $\mu = \mathcal{L}\sigma_{\text{inel}}/n_{\text{bunch}}f_{\text{r}}$, where $\mathcal{L}$ is the luminosity, σ_{inel} the *pp* inelastic cross section (71.5 mb at 7 TeV and 73.0 mb at 8 TeV), n_{bunch} the number of bunches and f_{r} the proton revolution frequency [78].

Parameter	2011	2012
$\sqrt{s}$ [TeV]	7	8
Number of colliding bunches	1380	1380
Bunch spacing [ns]	50	50
Maximum bunch intensity [protons/bunch]	1.45×10^{11}	1.7×10^{11}
Peak luminosity [cm^{-2}s^{-1}]	3.7×10^{33}	7.7×10^{33}
Maximum average number of interactions per bunch crossing	32	70
Longest beam lifetime [hours]	26	23
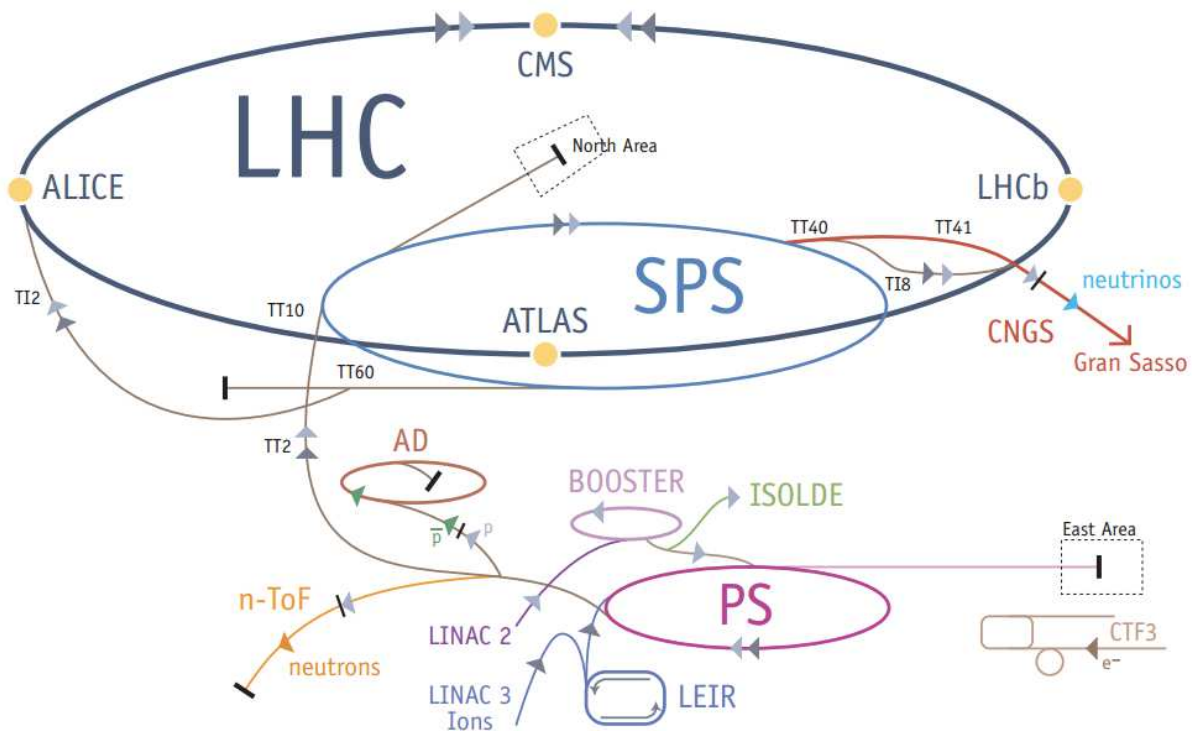
Figure 2.1: Sketch of the LHC accelerator and storage rings (© CERN).

2.2 The ATLAS Detector

ATLAS [79] is one of the four experiments installed along the LHC ring and, like CMS [80], is designed as a multi-purpose detector to test the SM and search for new physics at the TeV scale. The other two experiments are ALICE [81], which is specialised in heavy ion physics, and LHCb [82], which focuses on B-meson physics.

ATLAS is designed to reconstruct and identify a wide range of signatures, like missing transverse momentum $E_{\mathrm{T}}^{\mathrm{miss}}$, secondary vertices and high-p_{T} leptons and jets. The guiding principles of the detector design are as follows:

▷ Good momentum resolution and high reconstruction efficiency of the tracking system.

▷ Accurate electromagnetic and hadronic energy measurements in the calorimeters for the reconstruction and identification of muons, electrons, photons, jets, hadronic τ decays and $E_{\mathrm{T}}^{\mathrm{miss}}$.

▷ High granularity and solid angle coverage.

▷ Efficient reconstruction of secondary vertices.

▷ Radiation-hard detectors and front-end electronics.

Figure 2.2 shows a schematic view of the ATLAS detector. It consists of a high precision silicon tracking detector surrounded by a straw tube tracker in a 2 T solenoidal magnetic field. The superconducting magnet coil is surrounded by the electromagnetic and hadronic calorimeters. The outer-most part of the detector, defining its size of 44 m in length and

more than 25 m in height, is the muon spectrometer with its own air-core toroidal magnetic field. The precision muon tracking and trigger detectors are installed on the eight toroidal coils in the barrel and on three wheels in the two endcap regions.
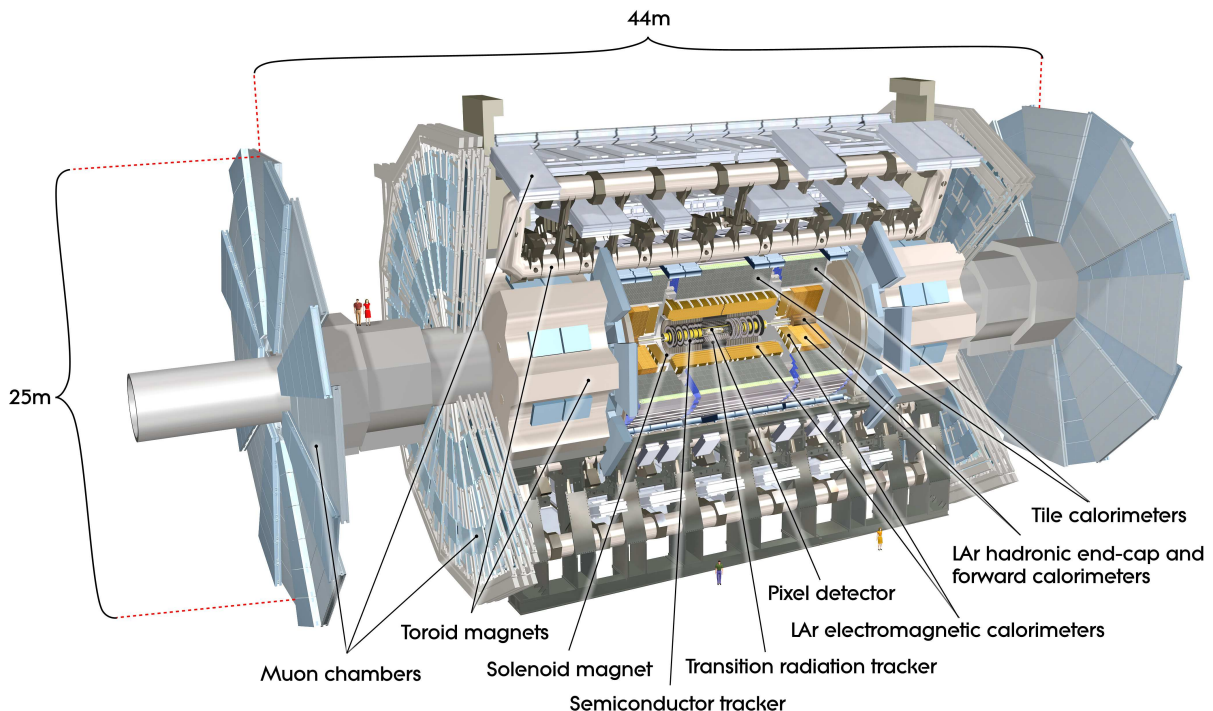
Figure 2.2: Schematic overview of the ATLAS detector (ATLAS Experiment © CERN).

In the ATLAS coordinate system [79] the z axis points along the beam direction, the y axis upwards and the x axis towards the center of the LHC ring. Paths of particles crossing the ATLAS detector are usually described in polar coordinates with the azimuthal angle ϕ between $-\pi$ and π and $\phi = 0$ on the positive x axis and with the polar angle θ between 0 and π and $\theta = 0$ on the positive z axis. Instead of the polar angle frequently the pseudo-rapidity $\eta = -\ln\tan(\theta/2)$ is used, which equals the Lorentz invariant rapidity $y = \frac{1}{2}\ln\left(\frac{E+p_z}{E-p_z}\right)$ in the limit of massless particles, where p_z is the momentum component along the beam axis. Angular separations between particle tracks are usually measured by the distance $\Delta R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2}$ in the $\eta - \phi$ plane.

2.2.1 The Detector Components

At the center of the ATLAS detector is the Inner Detector (ID) shown in Fig. 2.3. It is composed of the silicon pixel and silicon microstrip trackers (SCT) closest to the interaction point and of the Transition Radiation Tracker (TRT).

The first two detectors cover the region $|\eta| < 2.5$ and are designed for the measurement of the tracks and momenta of charged particles with $p_{\mathrm{T}} > 0.5$ GeV in a solenoidal magnetic field of 2 T and for the reconstruction of proton-proton interaction and secondary
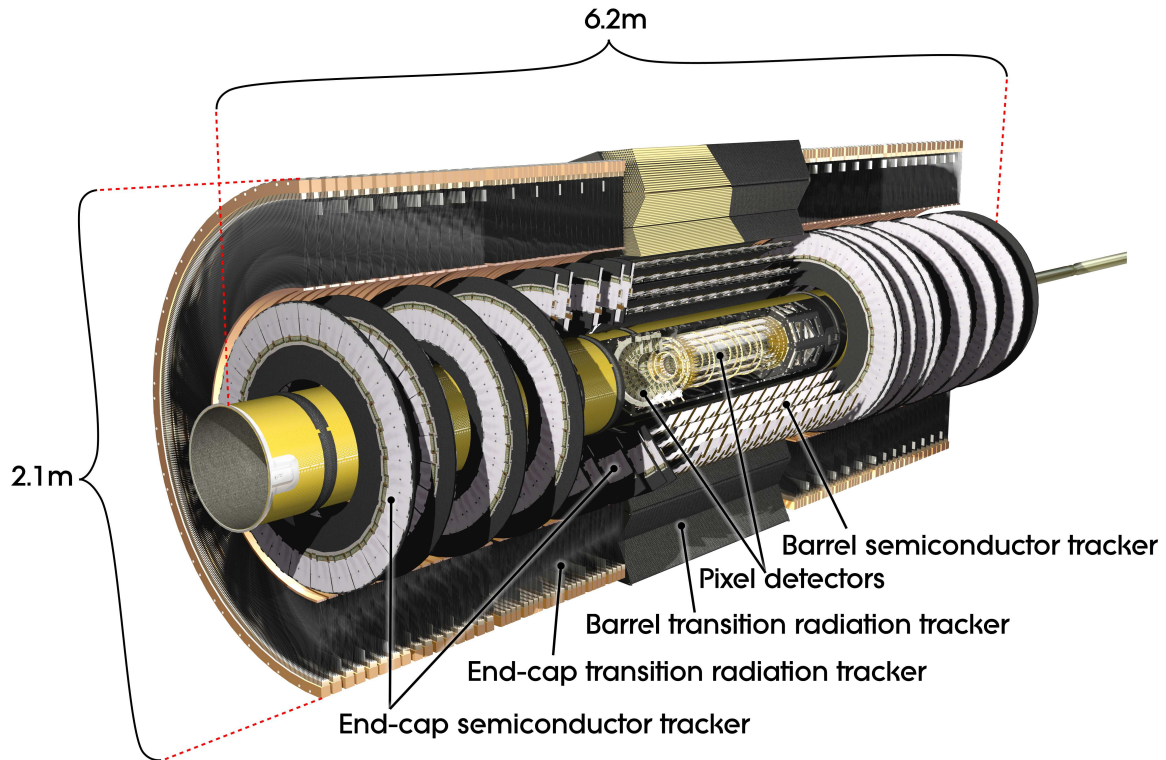
Figure 2.3: Schematic overview of the ATLAS Inner Detector. The pixel detector consists of three layers of pixels with intrinsic accuracies of 10 μm $(R - \phi)$ and 115 μm (z). The SCT has eight strip layers that provides four space points per track with accuracies of 17 μm $(R - \phi)$ and 580 μm (z). The TRT provides $R - \phi$ information with accuracy of 130 μm per straw (ATLAS Experiment © CERN).

decay vertices. The silicon trackers reconstruct tracks from seven space points along the trajectory. The tracking information is also used for the identification of jets originating from b quarks or hadronically decaying τ leptons.

The TRT surrounds the SCT covering the solid angle region $|\eta| < 2.0$. It consists of gas-filled straw tubes embedded in radiator material generating transition radiation. A track passing through the TRT leaves on average 36 hits. The transition radiation is used for electron identification.

The ID reconstruct transverse and longitudinal track impact parameters with resolutions of about 10 μm and 90 μm, respectively, combing the information from the pixel detector, SCT and TRT.

Outside of the superconducting coil of the ID are electromagnetic and hadronic calorimeters shown in Fig. 2.4. The electromagnetic calorimeter is a sampling calorimeter with liq-
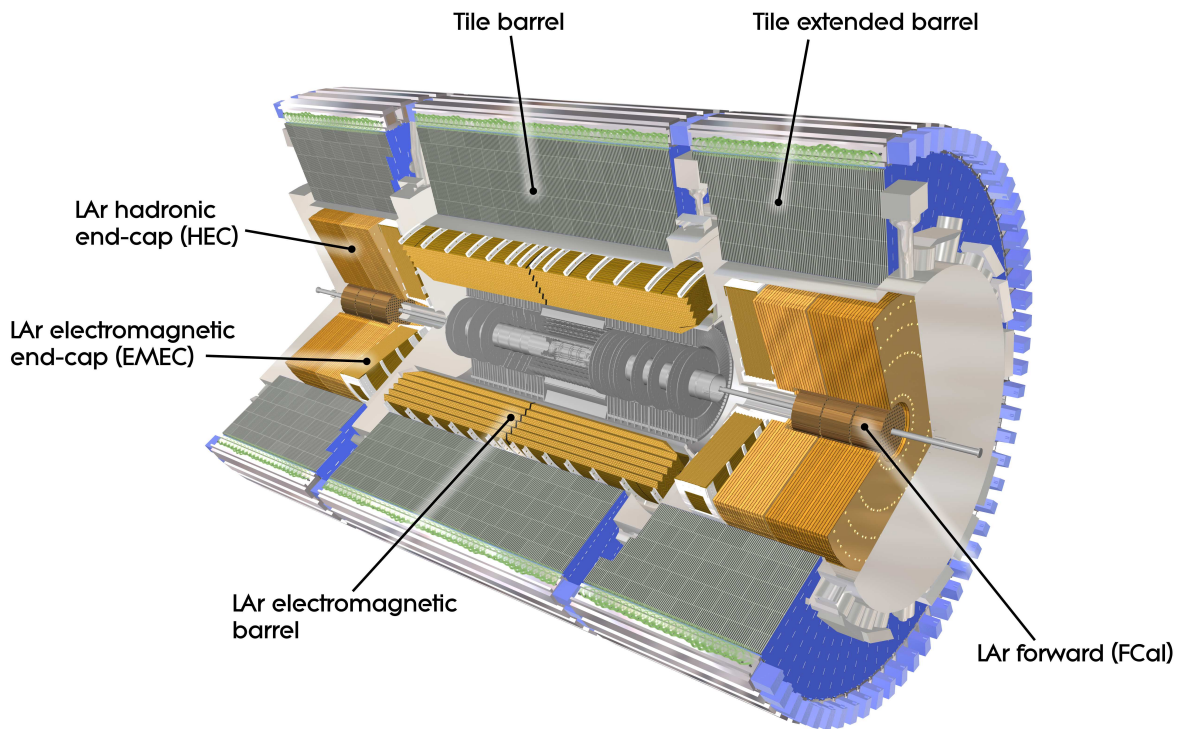
Figure 2.4: Schematic overview of the ATLAS calorimeters. The electromagnetic calorimeter is segmented into three layers with different granularity in depth and has a total thickness of more than $22X_0$. The typical cell size is $\Delta\eta \times \Delta\phi = 0.025 \times 0.025$. The hadronic calorimeter has a thickness more than 10 hadronic interaction lengths and a typical cell size of $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ (ATLAS Experiment © CERN).

uid argon (LAr) as active medium and lead as absorber material. The hadronic calorimeter is a scintillating tile sampling calorimeter with iron absorber plates in the central region and a LAr calorimeter with copper and tungsten absorber plates in the endcap

and forward regions, respectively.

Both electromagnetic and hadronic calorimeters cover the region $|\eta| < 4.9$, but the highest-granularity part of the electromagnetic calorimeter only extends up to $|\eta| = 3.2$. The high granularity provides accurate information on both magnitude and position of energy deposits. The matching or non-matching of tracks and energy deposits in the calorimeters is used for the reconstruction of electrons and jets and for the identification of neutral particles like photons and neutral pions which deposit energy in the calorimeters but leave no tracks in the ID. The hadronic calorimeters have lower granularity than the electromagnetic calorimeters, which is sufficient for jet reconstruction and E_T^{miss} measurement.

The total thickness of the calorimeter system at $|\eta| = 0$ is 11 interaction lengths, minimising punch-through of hadrons into the muon system and providing good E_T^{miss} resolution.

The outer part of ATLAS is the Muon Spectrometer (Fig. 2.5). It provides momentum
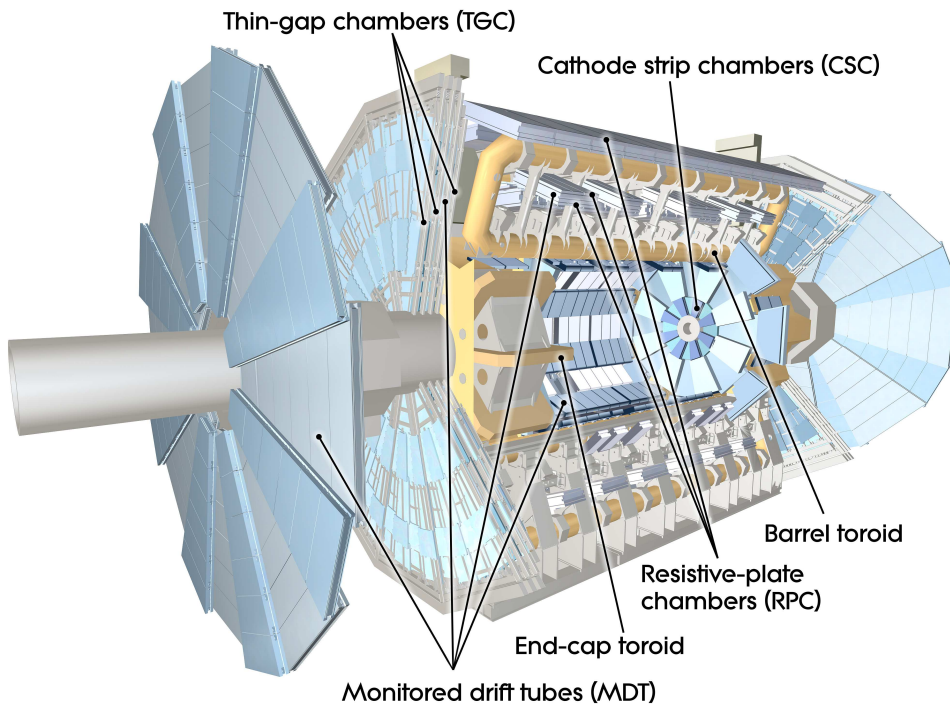
Figure 2.5: Schematic overview of the ATLAS Muon Spectrometer. Three layers of MDT and CSC chambers covering the range $|\eta| < 2.7$ each provide a track position resolution of better than 40 μm. The muon trigger chambers cover the solid angle range of $|\eta| < 2.4$ (ATLAS Experiment © CERN).

measurement of charged particles penetrating the calorimeters in the range $|\eta| < 2.7$. It is composed of three layers of Monitored Drift Tube (MDT) chambers or Cathode Strip Chambers (CSC) which measure the sagitta of charged tracks in a toroidal magnetic field of superconducting air-core magnets minimising multiple scatterings.

In addition to MDT and CSC, the muon tracking chambers, there are Resistive Plate Chambers (RPC) and Thin Gap Chambers (TGC) with high time resolution which are used for triggering on muons.

2.2.2   The Trigger System

The trigger system is essential for particle detectors at hadron colliders. Due to the high luminosity at LHC, the event rate reaches 1 GHz which requires fast and reliable algorithms to decide whether a given event contains an interesting signature and should be stored or not. Since the stored event rate is several orders of magnitude smaller than the collision rate, the triggers need to be very selective, i.e. they need to reject most of the events without loosing the few interesting ones.

The ATLAS trigger system is split into three levels. Level-1 (L1) is a hardware trigger based on the muon detector and the calorimeters using signatures of high-p_T muon, energy deposits in the calorimeters and E_T^{miss}. The L1 trigger has a time latency of $2.5\,\mu$s. In 2012, the typical L1 input and output rates were 20 MHz and 65 kHz, respectively. The level-2 (L2) trigger can use longer processing time and data with full granularity including tracking information. It has a latency of up to 100 ms with an output rate of 6.5 kHz. The trigger software is implemented on commercial computer farms. The final stage of the trigger system, the event filter (EF), has enough time to process the full event information using algorithms as for the offline event reconstruction. The event processing time can be as long as 1 s and the event rate is reduced down to the stored event rate of about 1 kHz corresponding to about 1 GB/s data rate.

2.2.3   The Luminosity Measurement

The accurate determination of the recorded luminosity is needed for all cross section measurements. In ATLAS the luminosity is determined indirectly by several detectors by the measurement of the pp interaction rate per bunch crossing μ_{vis} which is related to the luminosity as [83]

$$\mathcal{L} = \frac{\mu_{vis} n_b f_r}{\sigma_{vis}} \tag{2.1}$$

where n_b is the number of protons per bunch, f_r is the proton revolution frequency (11245.5 Hz) and $\sigma_{vis} = \epsilon\,\sigma_{inel}$ is the total inelastic pp cross section within the detector acceptance ϵ. The two detectors dedicated to the measurement of the luminosity are LUCID and BCM. Lucid is a Cherenkov detector placed at a distance of 17 m from the interaction point covering the solid angle range $5.6 < |\eta| < 6.0$, while the BCM consists of diamond sensors arranged around the beam pipe at 184 cm from the interaction point. Both LUCID and BCM measure the luminosity for each bunch crossing. The current measurement in the central and forward calorimeters are also sensitive to the luminosity. The calibration of σ_{vis} is performed using scans of the beam separation, known as van der Meer scans.

Fig. 2.6 [78] shows the integrated luminosities delivered by the LHC in 2011 and 2012 and recorded by ATLAS before and after requirements of good quality data for physics analyses. The difference between the delivered and the recorded luminosity is due to inefficiencies in the data acquisition and to the time needed by the detectors

to become operative once the beams in the accelerator are stable. Fig. 2.7 [83] shows
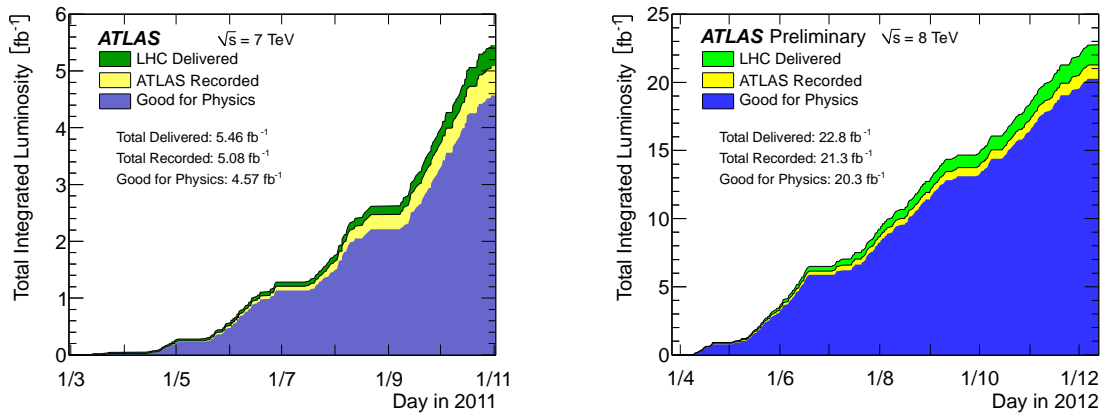
Figure 2.6: Cumulative integrated luminosity delivered to (green) and recorded by ATLAS before (yellow), and after requirements for good quality data for physics analyses (blue) in 2011 at $\sqrt{s} = 7$ TeV (left) and in 2012 at $\sqrt{s} = 8$ TeV (right) [78].

the distributions of the average number of interactions per bunch crossing for 2011 and 2012 data. The presence of multiple interactions, the so-called "pile-up", can seriously impact the performance of the detector from trigger and data acquisition to the event reconstruction and needs to be taken into account in the data processing and analysis. The overlay of several interactions in the detector can be due to either multiple collisions in the same bunch crossing, the "in-time" pile-up, or to the overlay of detector hits from interactions in different bunch crossings, the "out-of-time" pile-up.
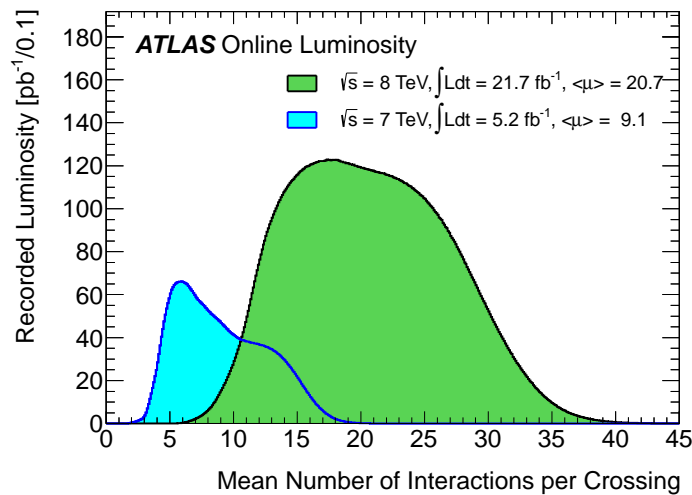
Figure 2.7: Luminosity-weighted distributions of the average number of interactions per bunch crossing in 2011 and 2012 data taking [83].

2.2.4   The Reconstruction of Physics Object

The following main physics objects are reconstructed from the electronic hits in the ATLAS detector and arranged in these categories:

Electrons and Photons: Electrons and photons leave very similar signatures in the electromagnetic calorimeters. Electrons, in contrast to photons, also produce tracks in the ID. The reconstruction of these objects starts from clusters of energy deposits in the electromagnetic calorimeter, which may or may not match with tracks in the ID [84]. Based on the track-cluster matching, electrons are distinguished from photons. The detailed analysis of the energy deposits in the calorimeter, using shower shapes, isolation criteria from other objects as well as the ID track quality are used to improve the identification of these objects and increase the rejection rate of charged hadrons.

Muons: Muons are detected in the muon spectrometer leaving little energy in the calorimeters. The ATLAS muon spectrometer provides very high momentum resolution [85]. Two independent measurements of the muon tracks are performed, in the ID and in the muon spectrometer, and combined to improve the momentum resolution in the common η-range.

Jets: In this thesis jets are reconstructed using the anti-k_t algorithm with distance parameter $R = 0.4$ [86]. Input to this algorithm are topologically connected clusters of energy deposits in the calorimeters calibrated using the local cluster weighting (LC) method [87].

τ Leptons: The majority of the τ decays (64.8%) is into hadrons which form strongly collimated jets in the detectors which are identified by a dedicated algorithm. This algorithm starts from the calibrated reconstructed jets. The hadronic part of the τ decay can be distinguished from QCD jets by the track multiplicity, the shape and the isolation of the energy deposits in the calorimeter and a reconstructed τ decay vertex since τ leptons in ATLAS travel on average 100 μm before decaying. The reconstruction, calibration and identification algorithms of the hadronic τ decay products are described in Chapter 3.

Missing Transverse Energy: The missing transverse energy $E_{\mathrm{T}}^{\mathrm{miss}}$ is defined as the momentum imbalance measured in the detector in the plane perpendicular to the beam axis where the sum of transverse momenta has to be zero due to momentum conservation. The imbalance signals undetected particles like neutrinos or new weakly interacting particles. Part of it can also be due to inactive detector regions and energy mismeasurement. The $E_{\mathrm{T}}^{\mathrm{miss}}$ reconstruction [88] uses all reconstructed objects, electrons, muons, photons, hadronically decaying τ leptons and jets:

$$E_{x(y)}^{\mathrm{miss}} = E_{x(y)}^{\mathrm{miss,e}} + E_{x(y)}^{\mathrm{miss},\gamma} + E_{x(y)}^{\mathrm{miss},\tau} + E_{x(y)}^{\mathrm{miss,jets}} + E_{x(y)}^{\mathrm{miss,SoftTerm}} + E_{x(y)}^{\mathrm{miss},\mu} \qquad (2.2)$$

where each term is the negative sum of the transverse energy component of the calibrated reconstructed objects in x and y direction. The "SoftTerm" contribution contains topological clusters not associated with any of the other objects. Since the

energy resolutions of jets and soft contributions are significantly effected by pile-up events, dedicated pile-up suppression methods are applied.

Chapter 3

τ-jet Reconstruction and Trigger

The search for the SM Higgs boson presented in this thesis is performed in the $H \rightarrow \tau^+\tau^-$ decays where both τ leptons decay hadronically. The detectable (visible) products of the hadronic τ lepton decay, i.e. all decay products except for the τ neutrino, appear in the detector as a particle jet with characteristic properties, referred to in the following as τ-jet. The strong discrimination between τ-jets and gluon- or quark-initiated jets, referred to as QCD jets, is of crucial importance for the presented Higgs boson search since the QCD multi-jet production is one of the two major sources of background. In this thesis the τ-jet trigger for the selection of Higgs boson candidates has been optimised to provide the highest possible signal selection efficiency at a reasonably low rate of triggered multi-jet events. In addition, the efficiencies of the τ-jet trigger and identification have been measured using dedicated calibration data samples. These measurements are used to correct the corresponding efficiencies obtained from the τ-jet simulation. These studies are described in the following after a general description of the algorithms for the τ-jet reconstruction with the ATLAS detector.

The τ lepton is the heaviest lepton with a mass of 1776.82 ± 0.16 MeV [25]. It decays into lighter leptons or hadrons plus a τ neutrino with a lifetime of $(290.6 \pm 1.0) \times 10^{-15}$ s and an average decay length of $87.11 \, \mu$m. τ leptons from Higgs boson decays decay already in the beam pipe such that only the τ decay products can be detected. Table 3.1 summarises the main decay modes with the corresponding branching ratios (BR).

Electrons and muons produced in leptonic τ decays (BR ~ 0.35) cannot be distinguished from prompt leptons produced in pp collisions. The τ lepton identification is, however, possible in hadronic decays (BR ~ 0.65) where τ-jets can be distinguished from QCD jets due to their characteristic features. A τ-jet originates from a small number of hadrons, typically one or three charged and none, one or two neutrals pions emitted close to the direction of the τ lepton. A QCD jet, instead, has on average a higher multiplicity of hadrons emitted in a wider cone around the jet axis. The reconstructed τ-jet consists of charged particle tracks reconstructed in the inner detector and energy deposits of charged pions in the hadronic calorimeter collimated with energy deposits of neutral pions decaying into photons in the electromagnetic calorimeter. The fraction of the τ-jet energy deposited in the electromagnetic calorimeter is typically larger than the energy fraction deposited by QCD jets due to the higher fraction of neutral pions in τ-jets. Based on these

Table 3.1: Main τ^- lepton decay modes and corresponding branching ratios (BR). $h^\pm$ stands for charged pions ($\pi^\pm$) and kaons ($K^\pm$), while h^0 stands for photons (γ) and neutral pions (π^0). Intermediate resonances are indicated [25].

Decay Mode	BR
$\mu^-\bar{\nu}_\mu\nu_\tau$	0.174
$e^-\bar{\nu}_e\nu_\tau$	0.178
$h^-(h^0)\,\nu_\tau$	0.501
$\quad\pi^-\nu_\tau$	0.108
$\quad\pi^-\pi^0\nu_\tau\,[\text{via }\rho^-(770)]$	0.255
$\quad\pi^-2\pi^0\nu_\tau\,[\text{via }a_1^-(1260)]$	0.093
$h^-h^-h^+(h^0)\,\nu_\tau$	0.146
$\quad\pi^-\pi^-\pi^+\nu_\tau\,[\text{via }a_1^-(1260)]$	0.093
$\quad\pi^-\pi^-\pi^+\pi^0\nu_\tau$	0.046

features, τ-jet trigger, reconstruction and identification algorithms have been developed.

Section 3.1 introduces the τ-jet reconstruction procedure. Sections 3.2 and 3.3 describe the algorithms for the τ-jet identification and trigger, respectively, together with the corresponding efficiency measurements with calibration data samples containing real τ-jets. Finally, section 3.4 describes the τ-jet trigger and identification efficiency measurements with calibration data samples containing fake τ-jets which do not originate from a real τ-jet.

3.1   Reconstruction of τ-jet Candidates

The reconstruction of the τ-jet candidates [79] is based on the information recorded by the inner detector and the electromagnetic and hadronic calorimeters. The reconstructed τ-jet candidates are seeded by a jet reconstructed from topological clusters using the anti-k_t algorithm with distance parameter $R = 0.4$ (see Section 2.2.4). This jet is required to have $p_{\text{T,jet}} > 10$ GeV and $|\eta_{\text{jet}}| < 2.5^1$. The reconstructed four-momentum $(p_\text{T}, \eta, \phi, m)$ of the τ-jet candidate is defined by the η_{jet} and ϕ_{jet} values of the seed jet and assuming the vanishing invariant mass of the visible τ lepton decay products, $m = 0$. The transverse momentum differs from the one of the seed jet and is determined by a dedicated energy calibration procedure, described in the next section.

Tracks with $p_\text{T}^{\text{trk}} > 1$ GeV are associated to the τ-jet if they are found within a cone of radius $\Delta R < 0.2$ around the axis of the seed jet. The tracks are required to pass a set of quality criteria including upper thresholds on the transverse and longitudinal impact parameters relative to the reconstructed primary vertex.

---

[1]This range corresponds to the η-coverage of the tracking system.

In the following, a reconstructed τ-jet candidate in simulated events is called either "true" or "fake" depending on whether it originates from a hadronically decaying τ lepton or not, respectively. According to the number of associated tracks, a τ-jet is defined as 1-prong, 3-prong or multi-prong.

3.1.1 The τ-jet Energy Calibration

The τ-jet energy [89] is reconstructed from the energy deposits in the electromagnetic and hadronic calorimeters and is calibrated to equal the energy of the visible decay products of the corresponding simulated τ lepton. The τ-jet energy scale (TES) obtained from the calibration corrects for the difference between the measured and the true τ-jet energy and is different from the seed jet energy scale because of the the larger fraction of electromagnetic energy in τ-jets compared to QCD jets.

The TES is determined by combining simulation-based and data-driven calibration procedures. The data-driven calibration procedure corrects the energy of the observed reconstructed τ-jet candidates to the energy of the corresponding reconstructed τ-jets obtained from simulation. The corresponding energy calibration factor is obtained by comparing the τ-jets energies measured in a dedicated $Z \to \tau\tau \to \mu\tau_{\text{jet}} + 3\nu$ calibration data sample ($\tau_{\text{jet}} = \tau$-jet) with the predictions from simulation. The additional simulation-based calibration relates the reconstructed energy of the simulated τ-jets to the true total energy of all visible τ lepton decay products. In this way, the reconstructed τ-jet energy is tuned on average to the true energy defined in the same way for both observed and simulated τ-jets.

The momentum resolution[2] obtained after applying the simulation-based calibration procedure on simulated τ-jets in the pseudo-rapidity range of $0.8 < |\eta| < 1.3$ is shown in Fig. 3.1 in dependence on the true momentum $p_{\text{vis}}^{\tau-\text{true}}$ of all visible τ decay products. The resolution is defined as the width σ of the distribution of the difference between the calibrated (p_F^τ) and the true ($p_{\text{vis}}^{\tau-\text{true}}$) momentum for simulated τ-jets relative to the true momentum. The momentum resolution obtained for 1-prong is typically better than the one for multi-prong τ-jets because of the higher fraction of energy deposited in the electromagnetic calorimeter in the former case.

The simulation-based TES is applied both on simulated and observed τ-jets. An additional difference in energy scales can occur due to the mis-modelling of the detector response. Thus, the energy of the observed reconstructed τ-jets needs to be corrected to match the prediction from simulation. Such difference is determined with a $Z \to \tau\tau \to \mu\tau_{\text{jet}} + 3\nu$ calibration data sample based on the invariant mass distribution of the visible Z boson decay products, $m(\mu, \tau_{\text{jet}})$. A corresponding TES correction factor is defined as the scaling of the p_{T} of the simulated τ-jets which gives the best statistical agreement between the observed and expected $m(\mu, \tau_{\text{jet}})$ distributions.

τ-jet Energy Scale Uncertainties

The TES calibration procedure is affected by uncertainties of the calorimeter response to the hadronic τ decay products, the modelling of pile-up interactions and underlying

---

[2]In reconstructed τ-jets $m = 0$, so that energy and momentum are equivalent.
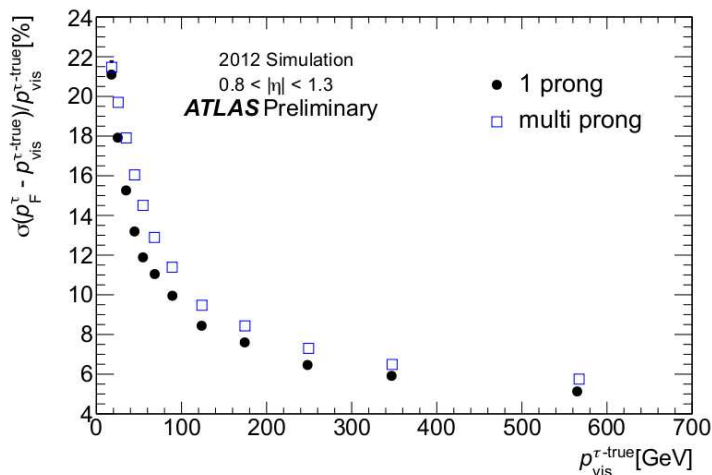
Figure 3.1: Momentum resolution for simulated 1-prong and multi-prong τ-jets in the pseudo-rapidity range $0.8 < |\eta| < 1.3$ in dependence on the true momentum of all visible τ decay products. The resolution is estimated with a Gaussian fit of the $(p_F^\tau - p_{\mathrm{vis}}^{\tau-\mathrm{true}})$ distribution in bins of $p_{\mathrm{vis}}^{\tau-\mathrm{true}}$ [89].

events, the description of the detector geometry and the remaining difference between the reconstructed and the true visible energy in simulated τ-jets after applying the simulation-based calibration [89].

The uncertainty of the calorimeter response is determined by the measurements of this response to different particles (single particle response). The corresponding TES uncertainty is then given by the convolution of the uncertainties for each single-particle τ decay product. Single particles are classified in low momentum charged hadrons, high momentum charged hadrons and neutral hadrons. The response to low momentum charged hadrons is given by the average ratio of the energy deposit in the calorimeter to the momentum of the associated track in the inner detector ($\langle E/p \rangle$) in a sample of isolated tracks with momenta up to 20 GeV in $|\eta| < 1.7$ and up to 60 GeV in $1.7 < |\eta| < 2.5$. The response to high momentum charged hadrons is measured using test beam data collected with a full slice of the ATLAS detector corresponding to the pseudo-rapidity region $|\eta| < 0.8$ exposed to a charged pion beam with pion momenta between 20 and 350 GeV. For high momentum charged hadrons in the pseudo-rapidity range of $0.8 < |\eta| < 2.5$ and outside the reach of the $\langle E/p \rangle$ measurement, uncertainties are estimated with simulation using different parton shower models. The energy response for neutral hadrons is estimated from the response of the electromagnetic calorimeter to electrons from observed $Z \to e^+ e^-$ decays.

The impact of other sources of uncertainties is determined from simulation by comparing predictions of different detector descriptions, underlying event models and pile-up conditions.

The total TES uncertainty ranges from 2-3% depending on the τ-jet transverse momentum, pseudo-rapidity and track multiplicity. Fig. 3.2 shows the total TES systematic uncertainty for 1-prong τ-jets in the pseudo-rapidity range of $0.3 < |\eta| < 0.8$, together

with the individual systematic components in dependence of the calibrated τ-jet transverse momentum.
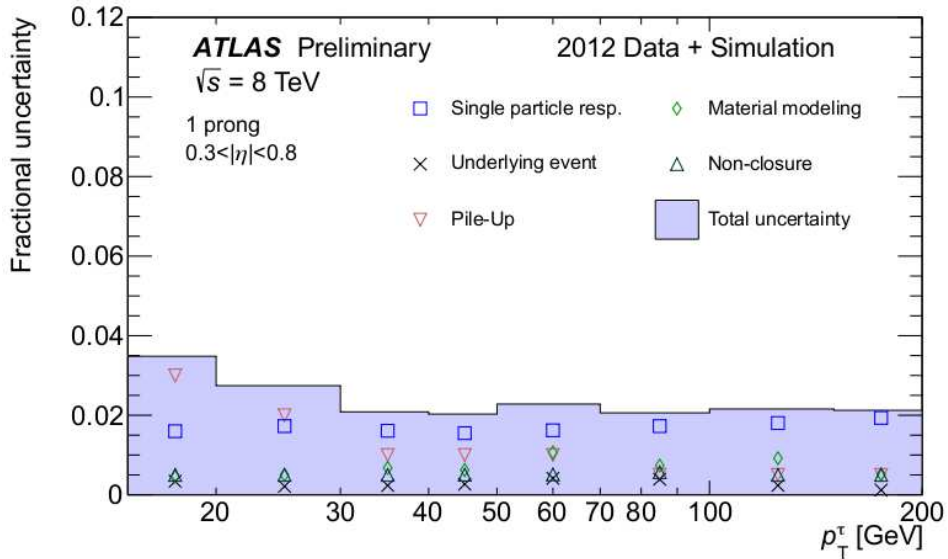
Figure 3.2: Systematic TES uncertainty from the simulation-based calibration in dependence on the calibrated τ-jet transverse momentum for 1-prong τ-jets in the pseudorapidity range $0.3 < |\eta| < 0.8$. The total uncertainty is decomposed into several independent components. "Single particle response" denotes the convolution of the measured uncertainties from each τ-jet constituent. "Underlying event", "Pile-Up" and "Material modeling" denote uncertainties of the τ-jet simulation. "Non-closure" denotes the difference in energy scale between the reconstructed and the true τ-jet energy after applying the simulation-based calibration on simulated τ-jets [89].

An independent measurement of the TES uncertainty is obtained from the data-driven calibration procedure. Given the rather low average transverse momentum of the τ-jets in $Z \to \mu\tau_{\mathrm{jet}} + 3\nu$ events, the uncertainty can be determined only for τ-jets with p_T up to about 70 GeV. This uncertainty is 1-2% and replaces the uncertainty of the single particle response for τ-jets with $p_\mathrm{T} < 70$ GeV.

3.2   The τ-jet Identification

The τ-jet reconstruction provides poor discrimination between τ-jets and QCD jets. A dedicated τ-jet identification algorithm is employed to enhance this discrimination based on the following features:

▷ the calorimeter cells containing energy deposits from a τ-jet are more collimated to each other than those with energy deposits from a QCD jet,

▷ the size of the cone containing all tracks associated to a τ-jet is narrower compared to the corresponding cone for QCD jets,

  $\triangleright$ the fraction of total energy carried by the highest momentum track associated to a
    τ-jet is larger than in the case of QCD jets where the total energy is more evenly
    distributed among all associated tracks,

  $\triangleright$ contrary to QCD jets initiated by light quarks or gluons, the tracks associated to
    a τ-jet can have a significant impact parameter relative to the primary vertex or
    form a secondary vertex in case of multi-prong τ-jets because of the relatively large
    τ lepton decay length and

  $\triangleright$ the number of neutral pions is typically zero or one in τ-jets, while it can be higher
    in QCD jets.

Based on discriminating variables corresponding to these features, in ATLAS two different identification methods have been designed based on multivariate algorithms with the projective likelihood (LLH) [90] and Boosted Decision Trees (BDT) [91]. Fig. 3.3 shows the distributions of several input variables used by the LLH and BDT identification algorithms [92]. Both algorithms combine these and other variables in order to achieve a high τ-jet identification efficiency and simultaneously a strong rejection of QCD jets. Fig. 3.4 shows the rejection power, i.e. the inverse of the QCD jet selection efficiency, as a function of the selection efficiency for true τ-jets as obtained with the two identification algorithms. Similar τ-jet identification performance is achieved. Three increasingly strict identification criteria, labelled *loose*, *medium* and *tight*, are defined corresponding to selection efficiencies for true τ-jets of 70%, 60% and 40% for 1-prong and 65%, 55% and 35% for multi-prong τ-jets, respectively. The corresponding rejection factor for QCD jets ranges from about 10 with the *loose* up to 100 with the *tight* identification criterion. The BDT τ-jet identification algorithm is used for the Higgs boson search performed in this thesis.

Rejection of Fake τ-jets from Electrons and Muons

In addition to QCD jets, also electrons and muons can be reconstructed and mis-identified as τ-jets, calling for additional discriminating variables.

The shape of the particle shower in the electromagnetic calorimeter provides a good discrimination against electrons, since electron-induced showers are in general shorter and narrower than those from τ-jets. Similar to the mentioned τ-jet identification algorithms, several shower-shape variables are combined in a dedicated BDT-based algorithm (electron BDT) optimised for the selection of electrons mis-identified as 1-prong τ-jets [92,93].

A muon can be mis-identified as a τ-jet if it deposits a large energy in the calorimeter or if an energy cluster is wrongly associated to the muon track. To reject such objects a muon veto algorithm is designed based on the fraction of the total energy deposited in the electromagnetic calorimeter and on the fraction of energy carried by the highest-momentum track associated to the τ-jet candidate [93].

3.2.1 Measurement of the τ-jet Identification Efficiency

The τ-jet identification efficiency has to be measured with data in order to correct the possible mis-modellings of the detector response in simulation. This measurement is performed in the frame of this thesis on a data sample enriched with $W \to \tau_{\mathrm{jet}} + 2\nu$ events
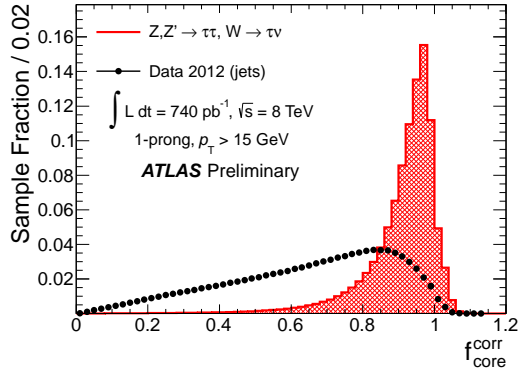
(a) $f_{\mathrm{core}}^{\mathrm{corr}}$: ratio of the energy deposited in the core cone of radius $\Delta R < 0.1$ to the energy in the cone of radius $\Delta R < 0.4$, corrected for pile-up contributions
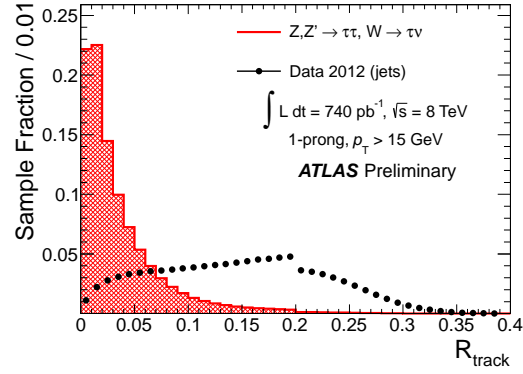
(b) R_{track}: p_{T}-weighted distance of the tracks from the τ-jet axis
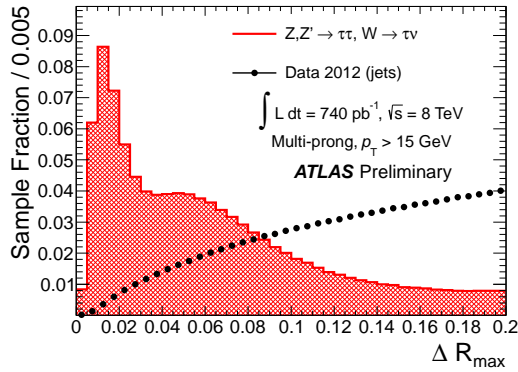
(c) ΔR_{max}: the maximum distance of the tracks from the τ-jet axis
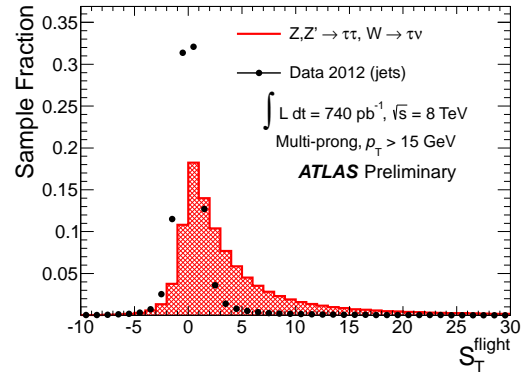
(d) $S_{\mathrm{T}}^{\mathrm{flight}}$: the ratio of the distance between the primary and the secondary vertex and the corresponding resolution

Figure 3.3: Distributions of discriminating variables used for the τ-jet identification. The distributions for true τ-jets (red) are obtained from simulated $Z \to \tau\tau$, $Z' \to \tau\tau$ and $W \to \tau\nu$ samples with hadronically decaying τ leptons, while the distributions for QCD jets (black) are based on jet-enriched data at $\sqrt{s} = 8$ TeV [92].
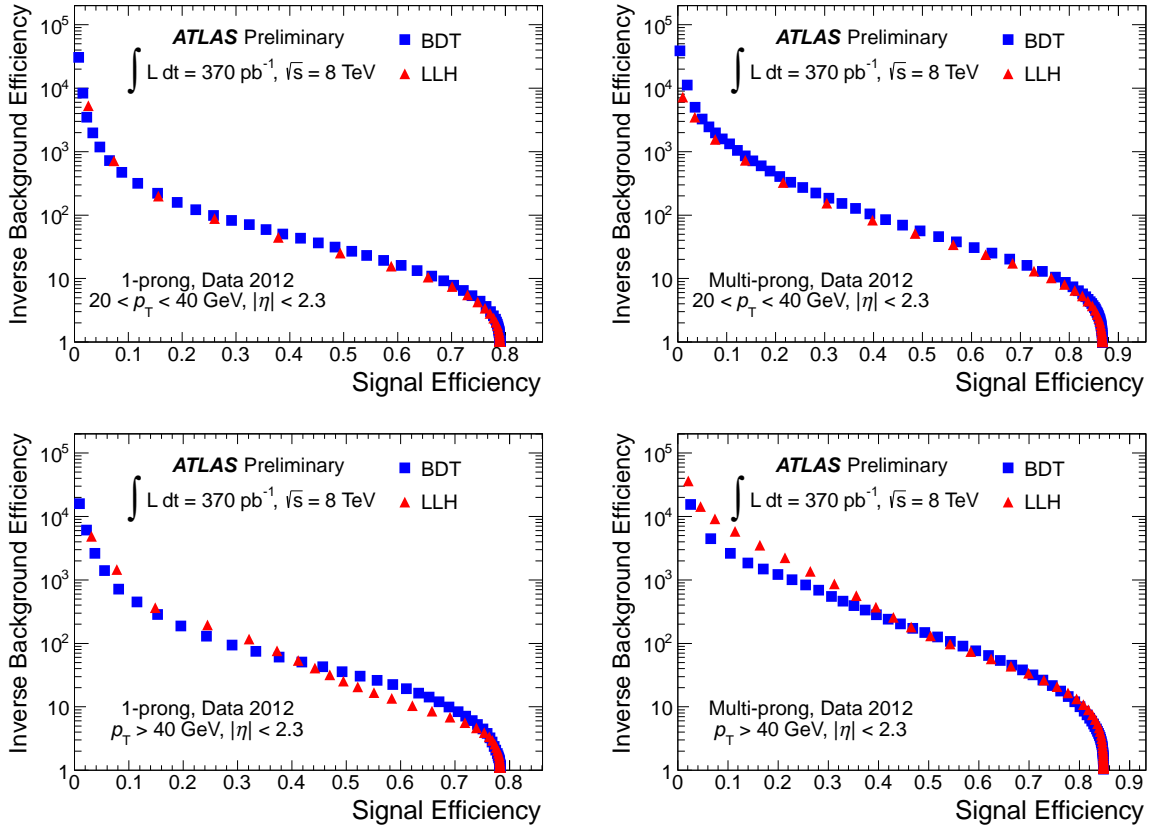
Figure 3.4: QCD jet rejection factor as a function of the τ-jet identification efficiency for 1-prong (left) and multi-prong (right) true low-p_T (top) and high-p_T (bottom) τ-jets, as obtained by the BDT (blue squares) and LLH (red triangles) τ-jet identification algorithms [92].

recorded at $\sqrt{s} = 7$ TeV [90]. The data sample is selected by requiring large $E_{\mathrm{T}}^{\mathrm{miss}}$ values in each event. The sample is contaminated mostly by background processes with QCD jets. A dedicated background estimation method is developed based on the multiplicity of tracks associated to the selected τ-jet candidate. A specially developed p_{T}-correlated track counting algorithm is employed to improve the discrimination power between true τ-jets and QCD jets of the nominal track counting, as described in the following. Subsequently, the efficiency measurement procedure is described and the corresponding results are summarised together with the results obtained after the combination with a complementary τ-jet identification efficiency measurement performed with $Z \to \mu\tau_{\mathrm{jet}} + 3\nu$ events.

The p_{T}-correlated Track Counting Algorithm

The track multiplicity is a powerful variable in discriminating true τ-jets from QCD jets because true τ-jets have typically one or three tracks, while QCD jets have more. However, the track counting algorithm used for the τ-jet reconstruction is not optimal for such discrimination, as the tracks are counted in a too narrow cone of radius $\Delta R = 0.2$ around the seed jet axis. This cone size is motivated by the collimation properties of the true τ-jets and is just large enough to contain all tracks associated to a true τ-jet. On the other hand, a QCD jet will on average have a higher track multiplicity with tracks spread in a wider cone. Ignoring the tracks outside the cone $\Delta R < 0.2$ reduces the potential discrimination power of the track multiplicity against QCD jets.

The alternative p_{T}-correlated track counting improves the discrimination between true τ-jets and QCD jets by counting tracks in a wider cone of radius $\Delta R = 0.6$. In order to avoid tracks from pile-up interactions or underlying events, a p_{T}-correlation similar to the one implemented in the anti-k_{T} jet algorithm [86] is applied. The algorithm runs over all so-called outer tracks with $p_{\mathrm{T}} > 500$ MeV in the region $0.2 < \Delta R < 0.6$ passing the same track quality criteria as those applied for the τ-jet reconstruction and computes for each outer track the distance

$$D = \max \left\{ \frac{p_{\mathrm{T}}^{\mathrm{core}} \Delta R(\mathrm{core}, \mathrm{outer})}{p_{\mathrm{T}}^{\mathrm{outer}}} \right\} \tag{3.1}$$

where the so-called core tracks are those within the cone $\Delta R < 0.2$ associated to the reconstructed τ-jet. $p_{\mathrm{T}}^{\mathrm{outer}}$ and $p_{\mathrm{T}}^{\mathrm{core}}$ are the transverse momenta of the outer and core tracks, respectively. Outer tracks with $D < 4$ are selected and added to the existing core tracks. The threshold for the distance D is chosen such that the track multiplicity of QCD jets becomes significantly higher compared to the one obtained with the nominal track counting, while the track multiplicity of true τ-jets remains almost unchanged.

Fig. 3.5 shows the performance of the p_{T}-correlated track counting for simulated true τ-jets and for QCD jets taken from a jet-enriched data sample. The track multiplicity obtained by the p_{T}-correlated track counting is compared to the one from the nominal track counting procedure in the cones of radius $\Delta R = 0.2$ and $\Delta R = 0.6$. With the nominal track counting in the cone $\Delta R < 0.2$ both true τ-jets and QCD jets have low track multiplicities. The same nominal track counting in the wider cone $\Delta R < 0.6$ increases the track multiplicity of both true τ-jets and QCD jets. The p_{T}-correlated track counting increases the track multiplicity of QCD jets with respect to the nominal track

counting while leaving the track multiplicity of true τ-jets unchanged, hence improving the discrimination between true and fake τ-jets.
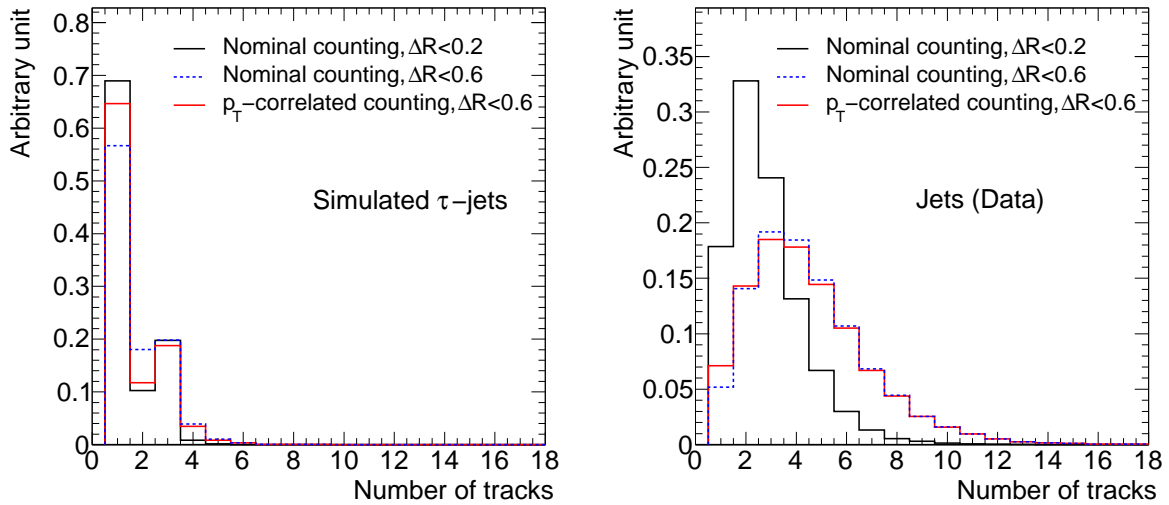
Figure 3.5: Track multiplicity distributions obtained with different track counting algorithms for simulated true τ-jets (left) and QCD jets from a jet-enriched data sample (right). Track multiplicity counted by the nominal procedure in the cone of $\Delta R < 0.2$ (black line) and $\Delta R < 0.6$ (dashed blue line) are shown together with the track multiplicity in cone of $\Delta R < 0.6$ obtained with the p_{T}-correlated track counting procedure (red line).

Efficiency Measurement with $W \to \tau\nu$ Data

The measurement of the τ-jet identification efficiency with a data sample enriched with $W \to \tau_{\mathrm{jet}} + 2\nu$ events has been designed on data collected in early 2011 [90, 94] and then applied to the full 2011 dataset in this thesis [93].

Events are selected by requiring large transverse missing energy, larger than 80 GeV at the trigger level and larger than 100 GeV in the reconstructed event. This requirement selects mostly events with a boosted W boson which recoils from a hard jet. Thus, the W boson decay products, the τ lepton and the neutrino, are emitted within a narrow cone. In order to increase the purity of the $W \to \tau_{\mathrm{jet}} + 2\nu$ data sample, additional requirements are imposed on the angular separation between the τ-jet and the $\mathbf{E}_{\mathrm{T}}^{\mathrm{miss}}$ vector, $\Delta\phi(\tau_{\mathrm{jet}}, \mathbf{E}_{\mathrm{T}}^{\mathrm{miss}}) < 1.5$, and on the transverse mass $m_{\mathrm{T}} = \sqrt{2E_{\mathrm{T}}^{\mathrm{miss}} p_{\mathrm{T}}(\tau_{\mathrm{jet}}) \cos\Delta\phi(\tau_{\mathrm{jet}}, \mathbf{E}_{\mathrm{T}}^{\mathrm{miss}})} < 80$ GeV.

The fractions of true and fake τ-jets in the observed data can be estimated based on the τ-jet track multiplicity distribution obtained with the previously described p_{T}-correlated counting procedure. The fit to the observed data is performed using three track multiplicity templates that model three classes of reconstructed τ-jets: true τ-jets, fake τ-jets from QCD jets and fake τ-jets from electrons or muons. The true τ-jet and lepton templates are based on simulation, while the QCD jet template is obtained from a control data sample enriched with QCD jets selected by requiring the presence of a lepton

in the final state. The fit of these templates to data has only one free parameter, the fraction f_{jet} of QCD jets relative to the total number of τ-jet candidates. The fraction of fake τ-jets from leptons, mostly electrons reconstructed as 1-prong τ-jets, is small and cannot be precisely determined by the fit. For this reason, the lepton fraction f_{lep} is estimated from simulation. The fraction of true τ-jets is defined as $f_{\mathrm{tau}} = 1 - f_{\mathrm{jet}} - f_{\mathrm{lep}}$, where f_{jet} is the only free parameter of the fit. After applying the full set of selection criteria, about 20% of the selected τ-jet candidates are estimated to be true τ-jets from $W \to \tau_{\mathrm{jet}} + 2\nu$ events, while the rest are fake τ-jets mostly from QCD jets in multi-jet events or electrons in $W \to e\nu$ events.

The systematic uncertainties of the fraction f_{tau} originate from uncertainties of the normalisation of the lepton template and of the shapes of the true τ-jet and QCD jet templates. The systematic uncertainty of the lepton template normalisation is dominated by the uncertainties on the rate of mis-identifying an electrons as a τ-jet. The uncertainty from the shape of the true τ-jet template is estimated comparing simulations with different parton shower and underlying event models as well as different detector geometries. The shape uncertainty of the QCD jet template is estimated by varying the requirements applied for the selection of the jet-enriched data sample.

The τ-jet identification efficiency is evaluated by measuring the ratio $\epsilon = f_{\mathrm{tau}}^{\mathrm{after\,ID}} / f_{\mathrm{tau}}^{\mathrm{before\,ID}}$ of the fitted fractions of true τ-jets $f_{\mathrm{tau}}^{\mathrm{before\,ID}}$ and $f_{\mathrm{tau}}^{\mathrm{after\,ID}}$ before and after applying a given τ-jet identification criterion, respectively. Fig. 3.6 shows the observed track multiplicities, together with the fitted templates before and after applying the BDT *tight* identification criterion.
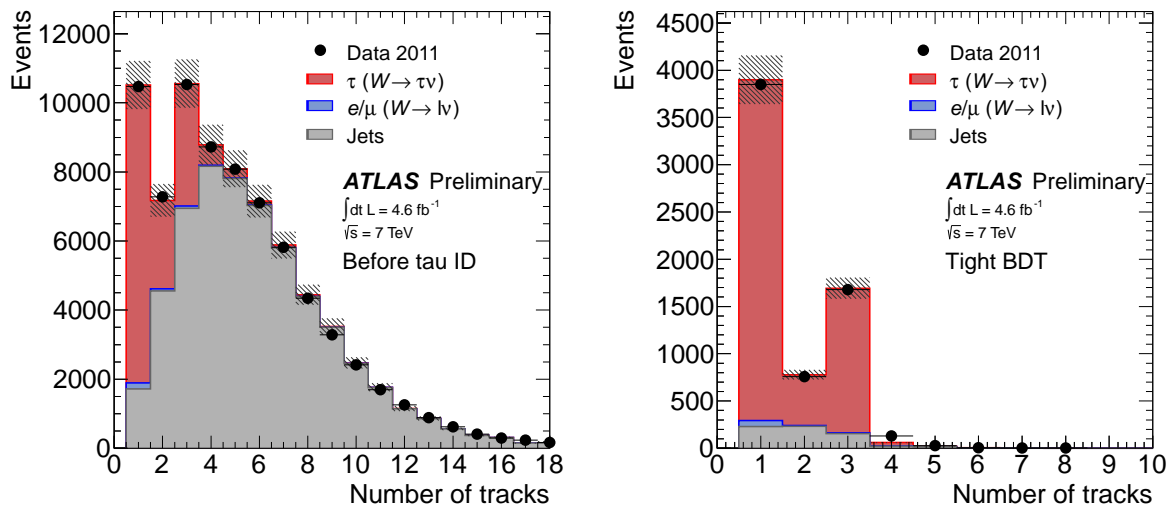
Figure 3.6: Observed and expected p_{T}-correlated track multiplicity distributions before (left) and after (right) applying the BDT *tight* identification criterion on τ-jet candidates with $p_{\mathrm{T}} > 22$ GeV. The yield of the QCD jet events is determined from the fit of the three template distributions to data [93].

The measurement is performed in different ranges of τ-jet transverse momentum above 20 GeV and the measured τ-jet identification efficiencies are compared to those obtained

from simulated $W \to \tau_{\mathrm{jet}} + 2\nu$ events. A good agreement is seen except for the low-p_{T} region between 20 GeV and 22 GeV. In these events the fraction of true τ-jets before identification is small and, given the available number of events, a possible mis-modelling of the background contribution cannot be excluded. For this reason, results are provided only for τ-jets with $p_{\mathrm{T}} > 22$ GeV.

Table 3.2 reports the measured τ-jet identification efficiencies for the *loose*, *medium* and *tight* criteria of the BDT and LLH algorithms. The statistical and systematic uncertainties are estimated from the differences in measured efficiencies when including or excluding the impact of the systematic uncertainties in the track multiplicity fit. The sum of statistical and systematic components is therefore bigger than the total uncertainty.

As a cross-check, the observed distributions of the outputs of the identification algorithms, the BDT score and the log-likelihood ratio, are shown in Fig. 3.7. Good agreement is obtained between the observed and the expected distributions normalised based on the f_{jet} fraction obtained from the track multiplicity fit.
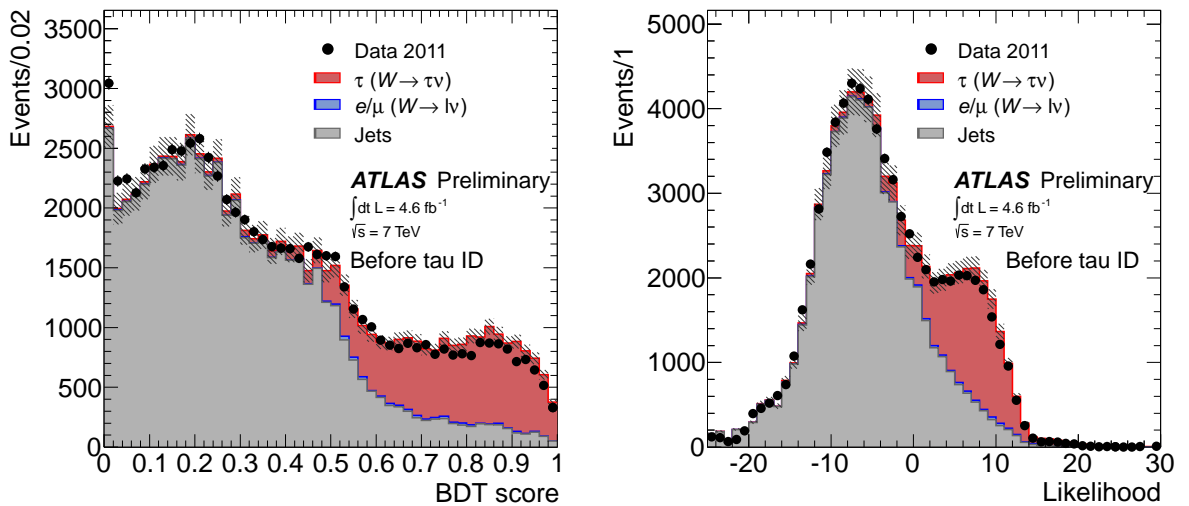
Figure 3.7: Observed and expected distributions of the BDT score (left) and log-likelihood ratio (right) for 1-prong and 3-prong τ-jet candidates with $p_{\mathrm{T}} > 22$ GeV before applying the identification criteria. Distributions are obtained with $W \to \tau\nu$ events in $\sqrt{s} = 7$ TeV data. The normalisations of the true and fake τ-jet contributions are estimated from the fit of the τ-jet track multiplicity templates to the observed data [93].

From this measurement, efficiency correction factors, defined as the ratios of the efficiency measured in data to the efficiency obtained from simulation can be determined. The correction factors are statistically compatible with unity for all levels of τ-jet identification, so that no corrections are applied on simulated τ-jets. The uncertainties of the correction factors are taken into account as a systematic uncertainty of the simulation of the τ-jet identification efficiency.

Table 3.2: τ-jet identification efficiencies and efficiency correction factors for simulated τ-jets with $p_{\mathrm{T}} > 22$ GeV, as measured in $W \to \tau\nu$ events in $\sqrt{s} = 7$ TeV data. Results are shown for three levels of τ-jets identification applied using the BDT and the LLH algorithms [93].

ID	ϵ_{Data}	Data/MC correction factor
BDT *loose*	$0.73 \pm 0.03(\mathrm{stat}) \pm 0.04(\mathrm{syst})$	$0.96 \pm 0.04(\mathrm{stat}) \pm 0.05(\mathrm{syst})$
BDT *medium*	$0.59 \pm 0.02(\mathrm{stat}) \pm 0.03(\mathrm{syst})$	$0.93 \pm 0.04(\mathrm{stat}) \pm 0.04(\mathrm{syst})$
BDT *tight*	$0.37 \pm 0.01(\mathrm{stat}) \pm 0.01(\mathrm{syst})$	$0.99 \pm 0.04(\mathrm{stat}) \pm 0.03(\mathrm{syst})$
LLH *loose*	$0.79 \pm 0.05(\mathrm{stat}) \pm 0.04(\mathrm{syst})$	$0.93 \pm 0.06(\mathrm{stat}) \pm 0.05(\mathrm{syst})$
LLH *medium*	$0.70 \pm 0.03(\mathrm{stat}) \pm 0.03(\mathrm{syst})$	$0.97 \pm 0.04(\mathrm{stat}) \pm 0.05(\mathrm{syst})$
LLH *tight*	$0.46 \pm 0.02(\mathrm{stat}) \pm 0.03(\mathrm{syst})$	$0.96 \pm 0.05(\mathrm{stat}) \pm 0.06(\mathrm{syst})$

Efficiency Correction Factors and Systematic Uncertainties

The results of the τ-jet identification efficiency measurement with $W \to \tau_{\mathrm{jet}} + 2\nu$ events are combined with the results of a complementary measurement performed with $Z \to \mu\tau_{\mathrm{jet}} + 3\nu$ events. The efficiency correction factors in both measurements show no deviations from unity for any track multiplicity and any transverse momentum above $p_{\mathrm{T}} > 22$ GeV. Table 3.3 reports the combined relative uncertainties on the correction factors for the simulated τ-jet identification efficiency obtained with $\sqrt{s} = 7$ TeV data. Similar measurements have been repeated with 8 TeV data.

Table 3.3: Relative uncertainties of the correction factors for the simulated τ-jet identification efficiency obtained for true τ-jets with $p_{\mathrm{T}} > 22$ GeV in $\sqrt{s} = 7$ TeV data [93].

ID	Inclusive [%]	1-prong [%]	3-prong [%]
BDT *loose*	4	4	8
BDT *medium*	4	5	8
BDT *tight*	4	4	7
LLH *loose*	5	4	10
LLH *medium*	4	5	10
LLH *tight*	5	5	11

3.3   The τ-jet Trigger

Dedicated trigger algorithms are used during the data acquisition for the selection of events with hadronically decaying τ leptons [95]. Similar to the previously described identification algorithms, these triggers take advantage of τ-jet features such as the collimation of the calorimeter clusters and the low track multiplicity to reject QCD jets.

At L1, the τ-jet trigger uses low-granularity information from the electromagnetic and hadronic calorimeters and selects events by applying energy thresholds on the calorimeter clusters and on the energy deposited around these clusters (isolation ring). At L2, the selection is performed with full-granularity information from the calorimeters and also the tracking system. At EF, thanks to the longer latency, reconstruction and identification algorithms similar to those used in the τ-jet reconstruction are used.

Due to the poor energy resolution and the lack of tracking information, the bottleneck of the τ-jet trigger selection is the L1 where the rejection of events with QCD jets is small, resulting in a high trigger rate strongly dependent on the peak luminosity. A meticulous optimisation of the L1 τ-jet trigger is needed in order to achieve high selection efficiency for true τ-jets at an affordable trigger rate.

The minimum τ-jet transverse energy (E_T) threshold for a trigger selecting events with at least one τ-jet (single-hadronic-tau trigger) has to be set to 125 GeV to allow for a reasonable trigger rate. This threshold is too high for the selection of fully hadronic $H \to \tau\tau$ decays. The trigger used for this search rather requires the presence of at least two τ-jets (double-hadronic-tau trigger) on which lower E_T thresholds at 20 and 29 GeV can be applied[3]. This trigger has been optimised during the 2011 and 2012 data taking within the framework of this thesis.

The strategy adopted to limit the trigger rate with increasing peak luminosity during the 2011 and 2012 data taking periods was to increase the E_T threshold and to employ isolation requirements. As shown in Fig. 3.8, these two requirements have different impacts on the trigger efficiency for a τ-jet with a given transverse momentum. On one hand, the τ-jet transverse momentum at which the trigger is fully efficient increases with increasing E_T threshold. On the other hand, the trigger efficiency for τ-jets with high transverse momentum decreases with tighter isolation requirements. Given that the bulk of the SM $H \to \tau\tau$ events have τ-jets with relatively low transverse momenta, the tightening of the isolation requirements is not expected to reduce the signal selection efficiency. However, isolation requirements are not as effective as the raising of the E_T threshold in reducing the trigger rate. A trade-off has been adopted by applying both loose isolation requirements and slightly higher E_T thresholds. Table 3.4 details the L1 double-hadronic-tau triggers used for the SM $H \to \tau\tau$ search in the fully hadronic final state. During the data

Table 3.4: L1 double-hadronic-tau triggers used for the SM $H \to \tau\tau$ search in the fully hadronic final state. Lower thresholds on the transverse energies and upper thresholds on the energy in the isolation ring applied on the two τ-jet trigger objects are indicated.

L1	E_T [GeV]	Iso [GeV]	$\sqrt{s}$ [TeV]	Peak Luminosity [cm^{-2}s^{-1}]	Integrated Luminosity [fb^{-1}]
`2TAU8_TAU11`	$> 8, 11$		7	2.3×10^{33}	2.2
`2TAU11_TAU15`	$> 11, 15$		7	3.8×10^{33}	2.4
`2TAU11I_TAU15`	$> 11, 15$	$\leq 4, 4$	8	7.6×10^{33}	20.3

---

[3]The choice of this trigger is the result of optimisation studies presented in the Master Thesis [96].
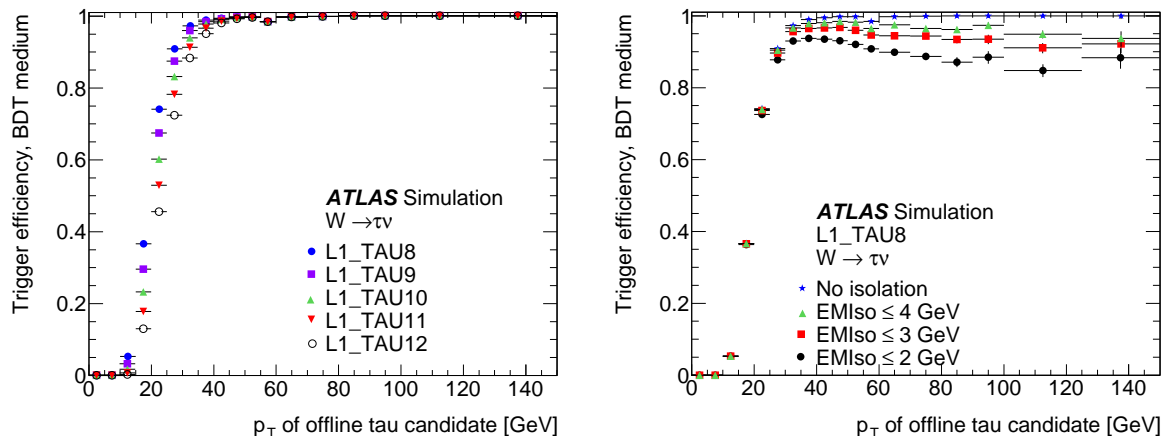
Figure 3.8: Comparison of L1 single-hadronic-tau trigger efficiencies for different E_T thresholds (left) and different isolation requirements (right) as a function of the p_T of the reconstructed τ-jets in $W \to \tau\nu$ sample at 7 TeV. The digit X in "L1_TAUX" indicates the E_T threshold in GeV applied to the calorimeter cluster and the requirement on "EMiso" is the threshold applied on the energy deposited in the ring around the calorimeter cluster [95].

taking period in 2011 the lower threshold on the transverse energy of the two τ-jet trigger objects has been raised and, in addition, at the beginning of the data taking period in 2012 isolation cuts have been applied. In spite of these changes, the L2 and EF triggers remained unchanged since the rates remained low enough.

3.3.1 Measurement of the τ-jet Trigger Efficiency

The τ-jet trigger efficiency for true τ-jets is measured with $Z \to l\tau_{jet} + 3\nu$ and $W \to \tau_{jet} + 2\nu$ data, using the same procedure applied for the measurement of the τ-jet identification efficiency [95]. These events are selected by muon and E_T^{miss} triggers, so that the reconstructed τ-jet is not affected by any trigger requirements and the trigger efficiency can be measured in an unbiased way. In the selected events, the τ-jet trigger efficiency is measured for the τ-jets passing the BDT *tight* τ-jet identification criterion, which are a sample of true τ-jets with a purity of about 95%. The background contribution from fake τ-jets is estimated as described Section 3.2.1 and is subtracted from the observed data. The trigger efficiency is measured as the fraction of selected τ-jets passing the trigger requirements as a function of the reconstructed identified τ-jet p_T. In order to reduce the statistical fluctuations, the p_T dependence of the trigger efficiency is described by the empirical analytic function with four parameters C_i, $i = 0, 1, 2, 3$

$$f(p_T) = C_0 \tan^{-1}\left(\frac{p_T - C_1}{C_2}\right) + C_3 \qquad (3.2)$$

Fig. 3.9 shows the efficiency of the trigger used for the selection of the highest-p_T τ-jet at the beginning of the data taking period in 2011 (EF_tau29_medium1) as measured

with the $W \to \tau_{\mathrm{jet}} + 2\nu$ events and compared to the corresponding efficiency estimated from simulation. The lower panel shows the ratio of the measured and predicted trigger efficiencies in each p_{T} bin together with the ratio of the corresponding fitted analytic efficiency functions. Good agreement is observed as the ratio is close to unity.
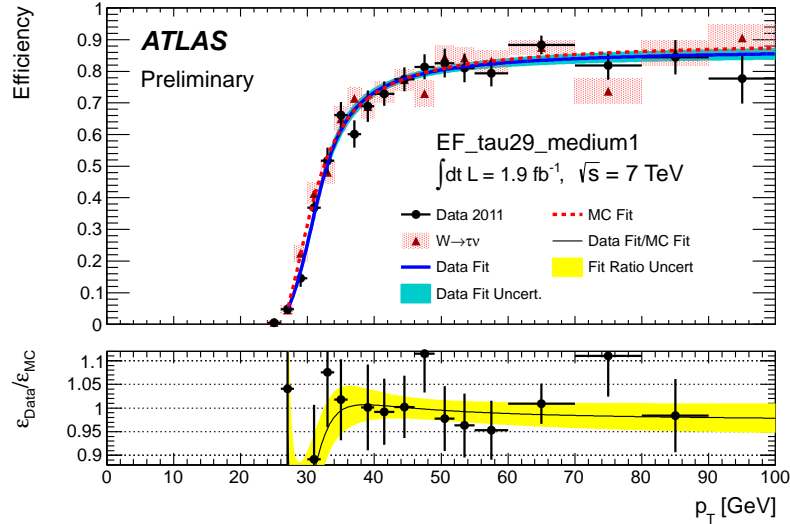
Figure 3.9: Single-hadronic-tau trigger efficiencies for the EF_tau29_medium1 trigger as a function of the reconstructed τ-jet transverse momentum measured in $W \to \tau_{\mathrm{jet}} + 2\nu$ events in $\sqrt{s} = 7$ TeV data [95].

The ratio of the fitted measured and simulated trigger efficiencies is used to correct the τ-jet trigger efficiency predicted by simulation. The uncertainty on this correction is about $2-3\%$ and is dominated by the uncertainty of the background subtraction in the low-p_{T} range and by limited statistics in the high-p_{T} region.

Together with the τ-jet identification efficiency measurement, the τ-jet trigger efficiency has been measured within the framework of this thesis using $W \to \tau_{\mathrm{jet}} + 2\nu$ events in $\sqrt{s} = 7$ TeV data. Similar results have been obtained in $Z \to l\tau_{\mathrm{jet}} + 3\nu$ events and the corresponding correction factors have been combined. Similar measurements have been performed in $\sqrt{s} = 8$ TeV data.

3.4 τ-jet Trigger and Identification Efficiency for QCD jets

Similar to the τ-jet trigger and identification efficiency for true τ-jets, corresponding efficiencies are also measured for QCD jets mis-identified as τ-jets using dedicated data samples. These efficiencies are expected to be poorly modelled in simulation due to uncertainties in the modelling of the detector response to QCD jets.

Generic correction factors valid for any mis-identified QCD jet cannot be derived since the mis-identification rate is expected to be different for quark-initiated and gluon-initiated jets. Quark-initiated jets have typically a smaller track multiplicity than gluon-

initiated jets, resulting in a higher probability to be identified as hadronically decaying τ leptons. Therefore, the QCD jet correction factors depend on the specific flavour composition of QCD jets selected by a given analysis.

The correction factors applied for the $H \to \tau\tau$ search in this thesis are derived from data in $W \to \mu\nu$ [97] and $Z \to \mu\mu$ events with additional jets in the final state and are mainly applied on jets in simulated W+jets and $t\bar{t}$ events. In these events, the jet mis-identified as a τ-jet typically originates from a quark.

The measured combined trigger and identification efficiency is defined as $\epsilon_{\text{ID+Trig}} = N_{\text{Reco+ID+Trig}}^{\text{probe}}/N_{\text{Reco}}^{\text{probe}}$, where $N_{\text{Reco}}^{\text{probe}}$ is the number of τ-jet candidates and $N_{\text{Reco+ID+Trig}}^{\text{probe}}$ is the number of τ-jet candidates passing the τ-jet trigger and identification requirements. Correction factors accounting for the differences between data and simulation are derived from the ratio of the efficiency measured in data to the simulated one. Table 3.5 reports the correction factors on the trigger and identification efficiency $\epsilon_{\text{ID+Trig}}$ for QCD jets passing the BDT *medium* identification criterion and the τ-jet trigger selections applied in the presented $H \to \tau\tau$ search. The uncertainties of the correction factors are dominated by the statistical component.

Table 3.5: Correction factors for the combined τ-jet trigger and identification efficiencies for mis-identified QCD jets in $\sqrt{s} = 7$ TeV data.

p_T [GeV]	BDT *medium*			
	EF_tau20_medium1	EF_tau29_medium1	EF_tau20T_medium1	EF_tau29T_medium1
$25 < p_T \leq 30$	0.34 ± 0.05	0.74 ± 0.43	0.36 ± 0.06	0.70 ± 0.41
$30 < p_T \leq 35$	0.25 ± 0.04	0.29 ± 0.07	0.25 ± 0.04	0.27 ± 0.06
$35 < p_T \leq 40$	0.28 ± 0.06	0.32 ± 0.07	0.28 ± 0.06	0.31 ± 0.07
$p_T > 40$	0.35 ± 0.05	0.36 ± 0.05	0.35 ± 0.05	0.35 ± 0.05

The measurement with $Z \to \mu\mu$ events is performed on data at $\sqrt{s} = 8$ TeV and the corresponding results are shown in Fig. 3.10.
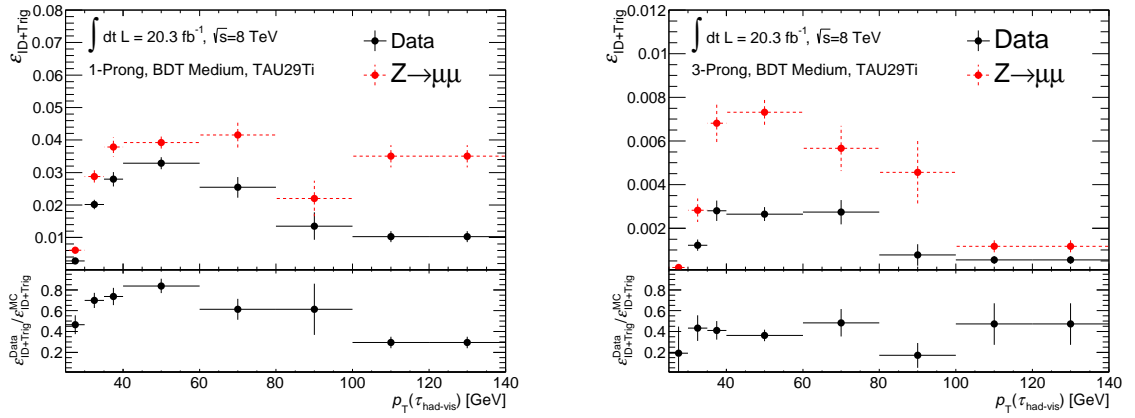
Figure 3.10: Measured combined τ-jet trigger and identification efficiencies (upper panel) and corresponding correction factors (lower panel) for mis-identified QCD jets as a function of the τ-jet transverse momentum for 1-prong (left) and 3-prong (right) τ-jets passing the BDT *medium* identification criterion and the EF_tau29Ti_medium1 trigger as obtained with $Z \rightarrow \mu\mu$ events in $\sqrt{s} = 8$ TeV data and simulation.

Chapter 4

$H \rightarrow \tau^+\tau^-$ Search in the Fully Hadronic Final State

4.1   Introduction

$H \rightarrow \tau^+\tau^-$ is an important decay mode for the SM Higgs boson search. It is the most important channel for the direct observation of the Higgs boson coupling to fermions and is also relevant for the measurement of the Higgs boson CP quantum number. Due to the small branching ratio and high background contributions, the sensitivity of this channel for the Higgs boson observation is not as high as in di-boson decay modes in which the Higgs boson has been discovered [5,6]. The sensitivity is also reduced by the poor $\tau^+\tau^-$ mass resolution because of the presence of at least two neutrinos from the τ lepton decays. The strategy to cope with these problems is to use

▷ event categories selecting specific Higgs boson production modes which allow for improved background rejection such as VBF production or ggF production with a highly boosted Higgs boson opposite to a jet and

▷ dedicated algorithms that estimate the most probable mass of the $\tau^+\tau^-$ resonance based on the kinematic properties of the visible τ decay products and the missing transverse energy $E_\mathrm{T}^\mathrm{miss}$.

The Higgs boson search in the $\tau^+\tau^-$ decay channel is performed in the three decay modes of the τ lepton pair. The visible products of a hadronically decaying τ leptons is referred to as either τ_jet or τ-jet.

▷ The fully leptonic final state $H \rightarrow \tau^\pm\tau^\mp \rightarrow l^\pm l'^\mp + 4\nu$, where both τ leptons decay leptonically into electrons or muons ($l = \mu, e$). In spite of the high detection efficiency of lepton pairs, the search in this channel has the lowest sensitivity because it has the smallest branching ratio of the $\tau^+\tau^-$ decays of 12.4% [25] and the worst mass resolution due to the presence of four neutrinos in the final state.

▷ The semi-leptonic final state $H \rightarrow \tau^\pm\tau^\mp \rightarrow l^\pm\tau_\mathrm{jet}^\mp + 3\nu$, where one τ lepton decays leptonically and the other one hadronically. This is the most sensitive final state because it has the largest branching ratio of 45.6% and requires the reconstruction of only one τ-jet.

▷ The fully hadronic final state $H \to \tau^\pm\tau^\mp \to \tau_{\mathrm{jet}}^\pm \tau_{\mathrm{jet}}^\mp + 2\nu$, where both τ leptons decay hadronically. This is the second most sensitive final state with a branching ratio of 42% and the necessity to reconstruct two τ-jets.

Despite the relatively large branching ratio and the presence of only two neutrinos in the final state, the fully hadronic channel requires powerful rejection of the QCD multi-jet background. The selection criteria used for this purpose at both the trigger and the reconstruction level lead to a significant reduction of the signal selection efficiency. Therefore, the sensitivity of this channel relies on the optimisation of the event selection criteria. Advantages of the fully hadronic channel are the simple background composition, where the individual background contributions can be measured almost completely from control data samples, and the relatively small number of discriminating variables.

The $H \to \tau^+\tau^-$ search is performed by selecting events with a well reconstructed pair of hadronically decaying τ leptons. These events are grouped into categories with different background contributions and signal-to-background ratios improving the sensitivity of the search. Although the combined sensitivity is dominated by the few highly sensitive categories, the remaining ones are also important not just to improve the sensitivity, but also to constrain the estimated background contributions with high statistics.
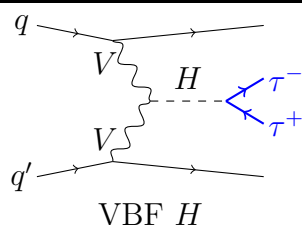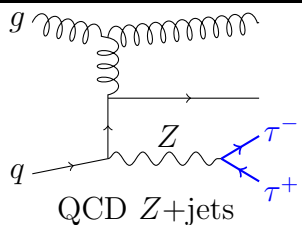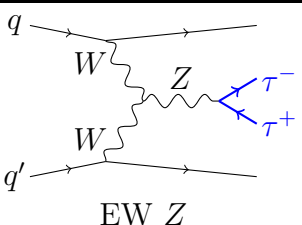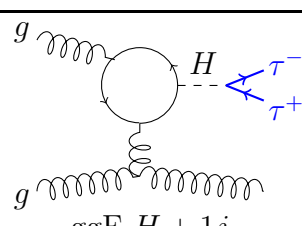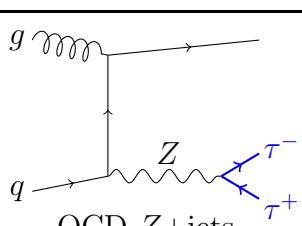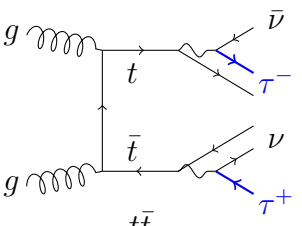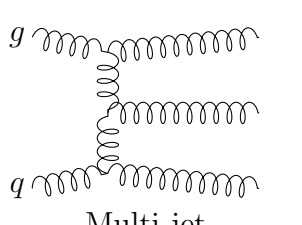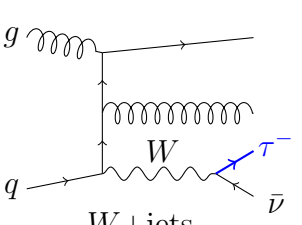
The Higgs boson production modes targeted by the event selection are the vector boson fusion (VBF) and the gluon fusion with the emission of a hard jet (ggF H+1jet) which results in a highly boosted Higgs boson. These two signal processes provide the characteristic event features allowing for the efficient suppression of background events: the presence of a pair of hard jets emitted close to the beam direction and the high transverse momentum of the $\tau^+\tau^-$ system. Fig. 4.1 shows the tree-level Feynman diagrams for the signal and the corresponding main background processes. For both signal processes the two main background sources are the irreducible Z boson production[1] with a resonant pair of τ leptons in the final state and the reducible multi-jet background with two jets mis-identified as τ-jets. For the VBF signal process the Z boson background contribution is composed not only of events produced via QCD, but also via electroweak interactions. Additional small background contributions originate from non-resonant $t\bar{t}$ and W+jets processes with one or two τ leptons in the final state.

The contribution of a given background process after the full event selection can be predicted using methods common to all event categories. The prediction of the Z+jets and multi-jet contributions is "data-driven", i.e. it is based on measurements performed with signal-depleted control data samples enriched with events from the corresponding background process. The rather small contribution of all other background processes is predicted by simulation.

The most powerful variable discriminating between the background and the signal processes is the invariant mass of the two τ leptons, $m_{\tau\tau}$. Due to the presence of two neutrinos and the relatively low energy resolution of the τ-jets, the $m_{\tau\tau}$ mass resolution is rather poor, on the order of 15%. The highest mass resolution is obtained using the Missing Mass Calculator (MMC) algorithm [98] and the distribution of the corresponding

---

[1]The Z boson background processes implicitly contain also events from virtual photons γ^*+jets background, although their contribution is very small at photon energies close to the Z boson mass which are in the focus of this analysis.

Table 4.1: Typical tree-level Feynman diagrams for the two signal processes targeted by the event selection, together with the corresponding main background processes. The final states are characterised by hadronically decaying τ leptons (blue) and quark- or gluon-initiated jets.

$m_{\tau\tau}^{\text{MMC}}$ variable is used for the statistical interpretation of the observed data relative to the predicted signal and background contributions. The $m_{\tau\tau}^{\text{MMC}}$ variable can also be used for the measurement of the mass of the potential Higgs boson signal. The search for a SM Higgs boson is performed for Higgs boson masses in the range of $100 \leq m_H \leq 150$ GeV. The analysis is optimised for the entire mass range without a specific mass dependence of the analysis within this range. This search is also documented in Ref. [99].

This chapter details the analysis of the fully hadronic $\tau^+\tau^-$ final state and is structured as follows. Section 4.2 describes the observed and simulated event samples. Sections 4.3 and 4.4 detail the requirements for the selection of physical objects reconstructed in each event and of the events themselves, respectively. Section 4.5 describes the MMC algorithm for the reconstruction of the invariant $m_{\tau\tau}$ mass. Section 4.6 details the methods for the prediction of the background contributions. Section 4.7 lists all experimental and theoretical systematic uncertainties impacting on the prediction of signal and background contributions. Section 4.8 describes the statistical model used for the interpretation of the observed data and Section 4.9 reports the observed results.

4.2   Data and Simulated Samples

This section briefly describes the analysed data and the corresponding simulated samples used for the modelling of the $H \to \tau^+\tau^-$ signal and SM background processes. The control data samples used for the data-driven predictions of the $Z \to \tau^+\tau^-$ and the multi-jet background contributions are described in Section 4.6.

4.2.1   Data

The analysed data from proton-proton collisions at the center-of-mass energy of 7 TeV collected in 2011 and of 8 TeV in 2012 amount to integrated luminosities of $4.6\,\text{fb}^{-1}$ and $20.3\,\text{fb}^{-1}$, respectively. The two data samples are analysed separately and only the respective final results are combined. Events are analysed only if recorded with all subsystems of the ATLAS detector being operational. The events are required to satisfy a number of quality criteria ensuring that the physical objects in each event are correctly reconstructed and that the event passes the double-hadronic-tau trigger.

The trigger requirements are optimised for events with two hadronically decaying τ leptons with low transverse momenta, as the ones produced in the $H \to \tau^+\tau^-$ decay. The Event Filter trigger level (EF) selects events with two τ-jets with transverse momenta above 29 GeV and 20 GeV and passing the *medium* τ-jet identification criterion. Moreover, both τ-jet trigger objects are required to have at least one and no more than three associated charged particle tracks. During the data taking, in response to the increase of the center-of-mass energy, the peak luminosity and consequently the level-1 (L1) trigger rate, the L1 trigger requirements have been gradually tightened. In the middle of the 2011 data taking period the transverse momentum thresholds have been increased and in 2012 τ-jet isolation requirements have been implemented. The impact of these additional requirements is described in Section 3.3.

4.2.2   Simulated Samples

$H \rightarrow \tau^+\tau^-$ events are simulated with NLO accuracy using POWHEG Monte-Carlo event generator for the ggF [100] and the VBF [101] production modes. The POWHEG ggF samples are generated using the dynamic renormalisation and factorisation scales $\mu = m_T^{\mathrm{Higgs}} = \sqrt{(m_H)^2 + (p_{\mathrm{T}}^{\mathrm{Higgs}})^2}$ and reweighted in such a way that the simulated Higgs boson transverse momentum spectrum matches the one predicted at NNLO+NNLL by the HqT programme [102, 103] as recommended by the LHC Higgs Cross Section Working Group [27]. The advantage of the dynamic compared to a fixed scale is an improved description of the ggF events containing a Higgs boson with high transverse momentum. Following the generation of the hard-scattering process with the POWHEG generator, the parton shower, the hadronization and the underlying event are simulated with the PYTHIA generator [104]. VH events are generated at LO with the PYTHIA generator. Given the small cross section, the $t\bar{t}H$ process is not simulated and is neglected in the analysis.

The generated signal samples are normalised to the cross sections calculated at the highest-order perturbation theory available, as described in Section 1.3. The ggF cross section is computed including NNLO+NNLL QCD and NLO EW corrections. The VBF cross section is computed with NLO QCD+EW and approximated NNLO QCD corrections. The VH cross section is calculated with NNLO QCD and NLO EW corrections. Table 4.2 lists the Higgs boson production cross sections and $H \rightarrow \tau^+\tau^-$ branching ratios for different Higgs boson mass hypotheses in pp collisions at $\sqrt{s} = 7$ TeV and 8 TeV [105]. Additional details of the Higgs boson phenomenology can be found in Section 1.3.

The $W/Z + $ jets background process is simulated with the ALPGEN generator [106] which employs the MLM scheme [107] for the matching between jets from the LO matrix element calculation with up to five jets and from the parton shower. The $t\bar{t}$ and di-boson (WW, WZ, ZZ) processes are produced with the MC@NLO [108] generator with NLO accuracy. Single top quark events are generated with the ACERMC programme [109]. The parton shower and the hadronization in all mentioned background processes are simulated with the HERWIG [110] and the underlying event with the JIMMY programme [111]. The loop-induced $gg \rightarrow WW$ events are generated using the gg2WW package [112]. Table 4.3 lists the production cross sections for each background process as predicted by the corresponding event generator.

Following the above event generation, the TAUOLA [113] and PHOTOS [114] packages are employed to simulate the τ lepton decay and the additional photon radiation in this decay, respectively. The set of parton distribution functions CT10 [115] is used for the generation of the POWHEG and MC@NLO samples, while CTEQ6L1 [116] is used for the PYTHIA and ALPGEN samples. Generated events subsequently undergo the detailed simulation of the ATLAS detector response modelled with the GEANT4 package [117]. The reconstruction of the simulated events is performed with the same software framework used also for the recorded data. An event-by-event reweighting based on the average number of interactions per bunch crossing is applied to all simulated events in order to reproduce pileup conditions equivalent to the observed ones.

Table 4.2: SM Higgs boson production cross sections and $H \rightarrow \tau^+\tau^-$ decay branching ratios in pp collisions at $\sqrt{s} = 7$ TeV and 8 TeV for different Higgs boson mass hypotheses [105]. The last column indicates the expected total signal yield for $H \rightarrow \tau_{\mathrm{jet}}\tau_{\mathrm{jet}} + 2\nu$ assuming $\mathrm{BR}(\tau\tau \rightarrow \tau_{\mathrm{jet}}\tau_{\mathrm{jet}} + 2\nu) = 0.4199$ [25].

m_H [GeV]	ggF [pb]	VBF [pb]	WH [pb]	ZH [pb]	BR [%]	$H \rightarrow \tau_{\mathrm{jet}}\tau_{\mathrm{jet}} + 2\nu$ [fb]
$\sqrt{s} = 7$ TeV						
100	23.64	1.56	1.20	0.66	8.28	940.5
105	21.45	1.48	1.03	0.57	8.17	841.6
110	19.56	1.41	0.88	0.50	7.95	746.1
115	17.89	1.34	0.76	0.43	7.58	650.4
120	16.43	1.28	0.66	0.38	7.04	554.3
125	15.13	1.22	0.58	0.34	6.32	458.3
130	13.98	1.17	0.51	0.30	5.45	364.9
135	12.95	1.12	0.44	0.26	4.49	278.4
140	12.02	1.07	0.39	0.23	3.52	202.7
145	11.24	1.02	0.34	0.21	2.61	140.5
150	10.51	0.98	0.30	0.18	1.78	89.5
$\sqrt{s} = 8$ TeV						
100	29.68	1.99	1.45	0.81	8.28	1179.5
105	27.01	1.90	1.24	0.70	8.17	1058.2
110	24.70	1.81	1.07	0.61	7.95	941.1
115	22.66	1.73	0.93	0.54	7.58	822.8
120	20.86	1.65	0.81	0.47	7.04	703.2
125	19.27	1.58	0.70	0.42	6.32	582.9
130	17.85	1.51	0.62	0.37	5.45	465.5
135	16.57	1.45	0.54	0.33	4.49	356.1
140	15.42	1.39	0.48	0.29	3.52	259.8
145	14.46	1.33	0.42	0.26	2.61	180.6
150	13.55	1.28	0.37	0.23	1.78	115.4

Table 4.3: Cross section times the branching ratio for the relevant background processes. The cross section for the $Z/\gamma^* \to \tau\tau$ + jets process is computed for events with $m_{\tau\tau}$ above 40 GeV and 60 GeV at $\sqrt{s} = 7$ TeV and 8 TeV, respectively.

Process	Cross section [pb]	
	$\sqrt{s} = 7$ TeV	$\sqrt{s} = 8$ TeV
$W \to \tau\nu$ + jets	10.46×10^3	12.22×10^3
$Z/\gamma^* \to \tau\tau$ + jets	1.07×10^3	1.15×10^3
$t\bar{t}$	164.57	238.06
Single top quark, t-channel ($W \to \tau\nu$)	6.93	9.48
Single top quark, s-channel ($W \to \tau\nu$)	0.5	0.61
Single top quark, Wt-channel	15.6	22.4
$qq \to WW \to l\nu l\nu + l\nu\tau\nu + \tau\nu\tau\nu$	4.72	5.66
$gg \to WW \to l\nu l\nu + l\nu\tau\nu + \tau\nu\tau\nu$	0.13	0.16
WZ (not fully hadronic decay)	4.35	5.22
ZZ (not fully hadronic decay)	1.08	1.30

4.3   Selection of Physics Objects

This section describes the requirements, listed in Table 4.4, imposed on the reconstructed physical objects relevant for the event selection: τ-jets, muons, electrons, jets and missing transverse energy. A more detailed description of the reconstruction algorithms can be found in Section 2.2.4.

τ-jets are the most important objects for the presented search and their reconstruction is detailed in Chapter 3. τ-jets are selected if their transverse momentum is higher than 20 GeV, the pseudo-rapidity is in $|\eta| < 2.5$ and the total electric charge of all associated tracks is $|q| = 1$. The constraint on the charge is equivalent to the requirement of having exactly one or three associated tracks within the cone of size $\Delta R < 0.2$ around the τ-jet axis. In order to improve the rejection of QCD jets mis-identified as τ-jets a track-based isolation criterion is applied, requiring that the number of tracks within the cone $\Delta R < 0.6$ is either 1 or 3. If the highest-p_{T} track associated to the τ-jet points to the transition region between the electromagnetic calorimeters ($1.37 < |\eta| < 1.52$), the τ-jet is rejected because of the high rate of electrons mis-identified as τ-jets in that poorly instrumented region of the detector. Following identification criteria discriminating against QCD jets, electrons and muons are applied: τ-jet candidates have to satisfy the BDT *medium* criterion for the τ-jet identification, to fail the electron BDT *loose* criterion and to pass the muon veto. Finally, τ-jets are required to match one of the two τ-jet trigger objects selected by the double-hadronic-tau trigger, ensuring the precise knowledge of the trigger efficiency in each selected event.

All simulated events are reweighted using correction factors for the τ-jet trigger and identification efficiency obtained from calibration data. If the reconstructed τ-jet object originates from a hadronically decaying τ lepton, the mentioned correction factors are

Table 4.4: Requirements imposed on the reconstructed physics objects used in the analysis. The star (*) indicates that a different requirement is used for the selection of control data (see Table 4.9).

τ-jet	$p_{\mathrm{T}} > 20$ GeV
	$|\eta| < 2.5$
	Highest-p_{T} track in $|\eta| < 1.37$ or $1.52 < |\eta| < 2.5$
	$|q| = 1^*$ ($N_{\mathrm{tracks}} = 1$ or 3)
	$N_{\mathrm{tracks},\Delta\mathrm{R}<0.6} = 1$ or 3^*
	BDT *medium* τ-jet identification
	If $N_{\mathrm{tracks}} = 1$: electron veto (electron BDT *loose*)
	Muon veto
	Trigger matching (see text)
Muon	$p_{\mathrm{T}} > 10$ GeV
	$|\eta| < 2.5$
Electron	$p_{\mathrm{T}} > 15$ GeV
	$|\eta| < 1.37$ or $1.52 < |\eta| < 2.47$
	Medium identification and quality requirements
Jet	At $\sqrt{s} = 7$ TeV: At $\sqrt{s} = 8$ TeV:
	$p_{\mathrm{T}} > 25$ GeV $p_{\mathrm{T}} > 30$ GeV
	$|\eta| < 4.5$ $|\eta| < 4.5$
	If $p_{\mathrm{T}} < 50$ GeV and $|\eta| < 2.4$: JVF > 0.5
	If $p_{\mathrm{T}} < 50$ GeV and $|\eta| > 2.4$: $p_{\mathrm{T}} > 35$

obtained from $Z \to \tau\tau$ and $W \to \tau\nu$ data (Sections 3.2.1 and 3.3.1). If, on the other hand, the reconstructed τ-jet originates from a mis-identified QCD jet, correction factors derived from $Z \to \mu\mu$ and $W \to \mu\nu$ data (Section 3.4) are applied.

The contribution of events in which both reconstructed τ-jets originate from mis-identified QCD jets is predicted using a dedicated control data sample with modified τ-jet selection requirements. In contrast to the nominal τ-jet selection mentioned above, there is no requirement imposed on the electric charge and all τ-jets with a track multiplicity equal or greater than one are accepted (see Section 4.6.2).

Muons are selected if they have either two matching tracks reconstructed both in the muon spectrometer and the inner detector or only a inner detector track. These muons are required to have $p_T > 10$ GeV and $|\eta| < 2.5$.

Electrons are selected if they have $p_T > 15$ GeV and $|\eta| < 2.47$. Electrons reconstructed in the transition region $1.37 < |\eta| < 1.52$ between the electromagnetic calorimeters in the barrel and end-cap detector regions are rejected in order to reduce the rate of QCD jets mis-identified as electrons. *Medium* identification and quality criteria are applied, as described in Ref. [84].

Jets are selected if they have $|\eta| < 4.5$ and transverse momentum $p_T > 20$ GeV and 25 GeV in $\sqrt{s} = 7$ TeV and 8 TeV data, respectively. The $\sqrt{s} = 8$ TeV dataset is characterised by a large number of pileup interactions, resulting in a high rate of reconstructed jets with p_T above the required threshold. To reject such jets, a cut is applied on the jet vertex fraction (JVF). The fraction of transverse momentum carried by the associated tracks originating from the primary vertex relative to the transverse momentum of all associated tracks is required to be JVF > 0.5. This requirement is applied only to reconstructed jets with $p_T < 50$ GeV and $|\eta| < 2.4$. For jets with $p_T > 50$ GeV this requirement is not needed as there is only a small impact of pileup interactions to jets above this p_T threshold. Jets outside the acceptance of the inner detector ($|\eta| > 2.4$) do not have associated tracks and in this case a higher p_T threshold of 35 GeV is used.

Missing transverse energy, E_T^{miss}, is reconstructed as described in Section 2.2.4. Different thresholds are applied on this variable depending on the event category, as described in the next section.

4.4   Event Selection and Categorisation

This section describes the event selection and categorisation used for the Higgs boson search as sketched in Fig. 4.1 and summarised in Tables 4.5 and 4.6.

4.4.1   Event Preselection

The first stage of the event selection, the so-called preselection, introduces a number of requirements to select signal-like events, as listed in the following and summarised in Table 4.5.
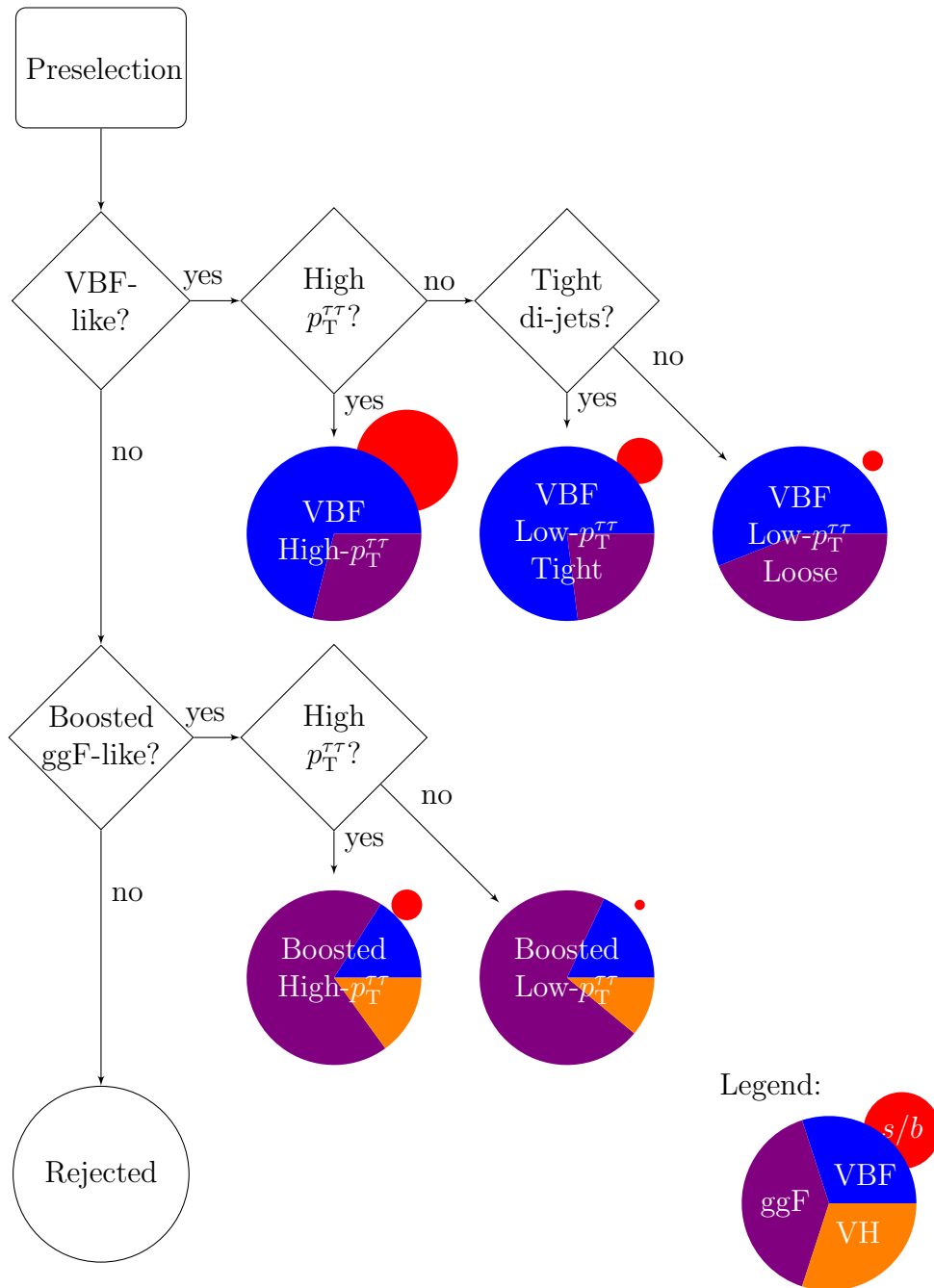
Figure 4.1: Flow-chart of the event selection and the resulting event categories used for the Higgs boson search. The pie-charts represent the expected signal fractions from ggF, VBF and VH production modes for the SM Higgs boson with $m_H = 125$ GeV. The size of the red circles indicates qualitatively the relative signal-to-background ratios in different event categories. Background contributions are predicted from simulation and auxiliary measurements with control data samples.

Table 4.5: Summary of the event preselection requirements. The star (*) indicates that a different requirement is used for the selection of the control data sample (see Table 4.9).

Preselection
▷ Double-hadronic-tau trigger
▷ Primary vertex with at least four associated tracks
▷ Exactly two τ-jets from the same primary vertex
▷ $q(\tau_{\mathrm{jet},1}) \cdot q(\tau_{\mathrm{jet},2}) = -1$*
▷ No electrons or muons in the final state
▷ ≥ 1 τ-jet passing the BDT *tight* τ-jet identification criterion
▷ $p_{\mathrm{T}}(\tau_{\mathrm{jet},1}) > 35$ GeV and $p_{\mathrm{T}}(\tau_{\mathrm{jet},2}) > 25$ GeV
▷ $0.6 < \Delta R_{\tau\tau} < 2.5$
▷ $|\Delta\eta_{\tau\tau}| < 1.5$
▷ $\mathbf{E}_{\mathrm{T}}^{\mathrm{miss}}$ in between the τ-jet vectors or $\min\left\{\Delta\phi(\mathbf{E}_{\mathrm{T}}^{\mathrm{miss}}, \tau_{\mathrm{jet},1}), \Delta\phi(\mathbf{E}_{\mathrm{T}}^{\mathrm{miss}}, \tau_{\mathrm{jet},2})\right\} < 0.2\pi$
▷ $E_{\mathrm{T}}^{\mathrm{miss}} > 20$ GeV
▷ Physical solution for $m_{\tau\tau}^{\mathrm{MMC}}$

▷ Each event is required to pass the double-hadronic-tau trigger.

▷ The presence of one primary vertex with at least 4 associated tracks with $p_{\mathrm{T}}^{\mathrm{track}} >$ 500 MeV is required in order to reject background contributions from non-colliding processes, for example from cosmic rays.

▷ Exactly two identified τ-jets associated to the same primary vertex and matched with the corresponding trigger objects are required.

▷ The two τ-jets are required to have opposite electric charges, $q(\tau_{\mathrm{jet},1}) \cdot q(\tau_{\mathrm{jet},2)} = -1$. For the selection of the control data sample used for the measurement of the multi-jet background contribution a different requirement on the charge product is applied, as described in Section 4.6.2.

▷ In order to improve the background rejection, no electrons or muons are allowed to be present in the final state and at least one of the two identified τ-jets has to satisfy the stricter BDT *tight* τ-jet identification criterion.

▷ The lower thresholds on the transverse momenta of the τ-jets are defined by the trigger acceptance and are set to the lowest values at which the trigger selection efficiency is almost p_{T}-independent. These thresholds are 35 GeV for the leading (highest-p_{T}) and 25 GeV for the subleading τ-jet.

▷ The constraint on the angular separation of the two τ-jets is imposed, $0.6 < \Delta R_{\tau\tau} <$ 2.5, to reject mainly QCD di-jet events where the two τ-jets are emitted back-to-back or events with two overlapping and poorly reconstructed τ-jets.

▷ The difference in pseudo-rapidity of the two τ-jets is required to be $|\Delta\eta_{\tau\tau}| < 1.5$ in order to reject non-resonant background events.

▷ Requirements on the vector of the missing transverse energy $\mathbf{E}_{\mathrm{T}}^{\mathrm{miss}}$ are imposed to

select events with two neutrinos emitted close to the two corresponding τ-jets. This selection is motivated by the strong boost of the two τ leptons from the Higgs boson decay which leads to the collinearity of the neutrinos and the visible products of the τ lepton decays. Consequently, the $\mathbf{E}_\mathrm{T}^\mathrm{miss}$ vector in the transverse plane is expected to be reconstructed in between the two τ-jets. Due to the limited resolution of the $\mathbf{E}_\mathrm{T}^\mathrm{miss}$ reconstruction, also events in which the $\mathbf{E}_\mathrm{T}^\mathrm{miss}$ vector is not in between but close to one of the two τ-jets are accepted. This requirement is implemented by selecting events with $E_\mathrm{T}^\mathrm{miss} > 20$ GeV and either

(a) $\max\left\{\Delta\phi(\mathbf{E}_\mathrm{T}^\mathrm{miss}, \tau_{\mathrm{jet},1}), \Delta\phi(\mathbf{E}_\mathrm{T}^\mathrm{miss}, \tau_{\mathrm{jet},2})\right\} \leq \Delta\phi(\tau_{\mathrm{jet},1}, \tau_{\mathrm{jet},2})$ and
  $\Delta\phi(\mathbf{E}_\mathrm{T}^\mathrm{miss}, \tau_{\mathrm{jet},1}) + \Delta\phi(\mathbf{E}_\mathrm{T}^\mathrm{miss}, \tau_{\mathrm{jet},2}) \leq \pi$ or

(b) $\min\left\{\Delta\phi(\mathbf{E}_\mathrm{T}^\mathrm{miss}, \tau_{\mathrm{jet},1}), \Delta\phi(\mathbf{E}_\mathrm{T}^\mathrm{miss}, \tau_{\mathrm{jet},2})\right\} < 0.2\pi$,

where $\Delta\phi$ is defined in the range $[0, \pi)$.

▷ The Missing Mass Calculator (MMC) algorithm is required to provide a physical solution for the invariant mass $m_{\tau\tau}$ of the two τ leptons (see Section 4.5).

The sample of events selected at this preselection stage is expected to contain a small fraction of Higgs boson signal events ($< 1\%$) compared to the background contribution and can therefore be used for the validation of methods for the background prediction from control data samples (see Section 4.6.3).

4.4.2   Event Categorisation

All events passing the preselection stage are sorted into several categories defined by selection criteria optimised to select signal events in which the Higgs boson is produced either via VBF or ggF production (see Fig. 4.1 and Table 4.6). The resulting so-called VBF and Boosted event categories are further subdivided into categories with different background compositions and signal-to-background ratios.

All event categories are inclusive in jet multiplicity, having no upper threshold on the number of jets in the final state. This avoids theoretical uncertainties on the signal yield predictions arising from the suppression of real emissions from high-order QCD corrections. The inclusive jet selection determines the ordering of the event selection criteria. First, events with two or more jets are selected targeting the VBF Higgs boson production. Then, the remaining events have to satisfy selection criteria targeting the ggF production of a high momentum Higgs boson recoiling from a hard jet.

VBF Event Categories

The selection of VBF-like events relies on the characteristic signature of the VBF Higgs boson production mode with two hard jets in the final state emitted in the forward detector regions. Common selection criteria require the presence of at least one jet in the final state with $p_\mathrm{T} > 50$ GeV and a second one with $p_\mathrm{T} > 30$ GeV. These two, labelled as "tagging" jets, have to be well separated in pseudo-rapidity with $|\Delta\eta_{jj}| > 2.6$ and have a high invariant mass of $M_{jj} > 250$ GeV. The tagging jets are required to be emitted in opposite directions relative to the Higgs boson decay products, such that the two τ-jets should be emitted within the region defined by the pseudo-rapidities of the two tagging jets, i.e.

Table 4.6: Summary of the event selection and categorisation criteria.

VBF		
$\triangleright \geq 2$ jets with $p_{\mathrm{T}} > 50, 30$ GeV		
$\triangleright |\Delta\eta_{jj}| > 2.6$ and $M_{\mathrm{jj}} > 250$ GeV		
$\triangleright \min(\eta_{\mathrm{jet},1}, \eta_{\mathrm{jet},2}) < \eta_{\tau,1}, \eta_{\tau,2} < \max(\eta_{\mathrm{jet},1}, \eta_{\mathrm{jet},2})$		
High-$p_{\mathrm{T}}^{\tau\tau}$	**Low-$p_{\mathrm{T}}^{\tau\tau}$**	
$\triangleright \Delta R_{\tau\tau} < 1.5$	$\triangleright \Delta R_{\tau\tau} > 1.5$ or $p_{\mathrm{T}}^{\tau\tau} < 140$ GeV	
$\triangleright p_{\mathrm{T}}^{\tau\tau} > 140$ GeV	**Tight**	**Loose**
	$\triangleright M_{jj}[\text{ GeV}] > -250|\Delta\eta_{jj}| + 1550$	$\triangleright M_{jj}[\text{ GeV}] < -250|\Delta\eta_{jj}| + 1550$
Boosted		
$\triangleright$ Preselected events not accepted in any VBF category above		
$\triangleright \mathbf{E}_{\mathrm{T}}^{\mathrm{miss}}$ in between the τ-jet vectors or $\min\left\{\Delta\phi(\mathbf{E}_{\mathrm{T}}^{\mathrm{miss}}, \tau_{\mathrm{jet},1}), \Delta\phi(\mathbf{E}_{\mathrm{T}}^{\mathrm{miss}}, \tau_{\mathrm{jet},2})\right\} < 0.1\pi$		
$\triangleright p_{\mathrm{T}}^{\tau\tau} > 80$ GeV		
High-$p_{\mathrm{T}}^{\tau\tau}$	**Low-$p_{\mathrm{T}}^{\tau\tau}$**	
$\triangleright \Delta R_{\tau\tau} < 1.5$	$\triangleright \Delta R_{\tau\tau} > 1.5$ or $p_{\mathrm{T}}^{\tau\tau} < 140$ GeV	
$\triangleright p_{\mathrm{T}}^{\tau\tau} > 140$ GeV		

the τ-jet pseudo-rapidity η_τ should satisfy $\min(\eta_{\mathrm{jet},1}, \eta_{\mathrm{jet},2}) < \eta_{\tau,1}, \eta_{\tau,2} < \max(\eta_{\mathrm{jet},1}, \eta_{\mathrm{jet},2})$. In Fig. 4.2 a display of a VBF Higgs boson candidate event recorded at $\sqrt{s} = 8$ TeV is given showing a pair of τ-jets with high transverse momenta emitted in the central part of the detector and two tagging jets with high invariant mass emitted in forward detector regions. The $\mathbf{E}_{\mathrm{T}}^{\mathrm{miss}}$ vector points close to the direction of the two τ-jets.

The VBF-like events selected by the above criteria are subdivided into three categories.

VBF High-$p_{\mathrm{T}}^{\tau\tau}$ Category: The most sensitive VBF category is defined by requirements on the transverse momentum of the Higgs boson candidate $p_{\mathrm{T}}^{\tau\tau} = |\mathbf{p}_{\mathrm{T}}^{\tau-\mathrm{jet},1} + \mathbf{p}_{\mathrm{T}}^{\tau-\mathrm{jet},2} + \mathbf{E}_{\mathrm{T}}^{\mathrm{miss}}| > 140$ GeV and on the angular separation of the two τ-jets $\Delta R_{\tau\tau} < 1.5$. As shown in Fig. 4.3, these requirements reject most of the multi-jet background and select events in which the $m_{\tau\tau}^{\mathrm{MMC}}$ invariant mass is reconstructed with higher resolution. The $p_{\mathrm{T}}^{\tau\tau}$ and $\Delta R_{\tau\tau}$ variables are strongly correlated (Fig. 4.3c), such that events with high $p_{\mathrm{T}}^{\tau\tau}$ values will in general have small $\Delta R_{\tau\tau}$ values. Nonetheless, explicit requirements on both variables are needed to further suppress background events with high $\Delta R_{\tau\tau}$ affecting the sensitive $m_{\tau\tau}$ invariant mass range where signal events are expected. This is illustrated in Fig. 4.4 which shows the $\Delta R_{\tau\tau}$ and $m_{\tau\tau}^{\mathrm{MMC}}$ distributions for preselected events with $p_{\mathrm{T}}^{\tau\tau} > 140$ GeV, before and after applying the requirement of $\Delta R_{\tau\tau} < 1.5$. The latter selection criterion helps rejecting background events produced by non-resonant processes or events in which the $m_{\tau\tau}^{\mathrm{MMC}}$ is not correctly reconstructed. 70% of all signal events in the VBF High-$p_{\mathrm{T}}^{\tau\tau}$ category are expected to be produced via VBF and the remaining signal predominately via ggF production mode. The expected background originates mostly from the Z+jets process. The expected signal-to-background ratio in
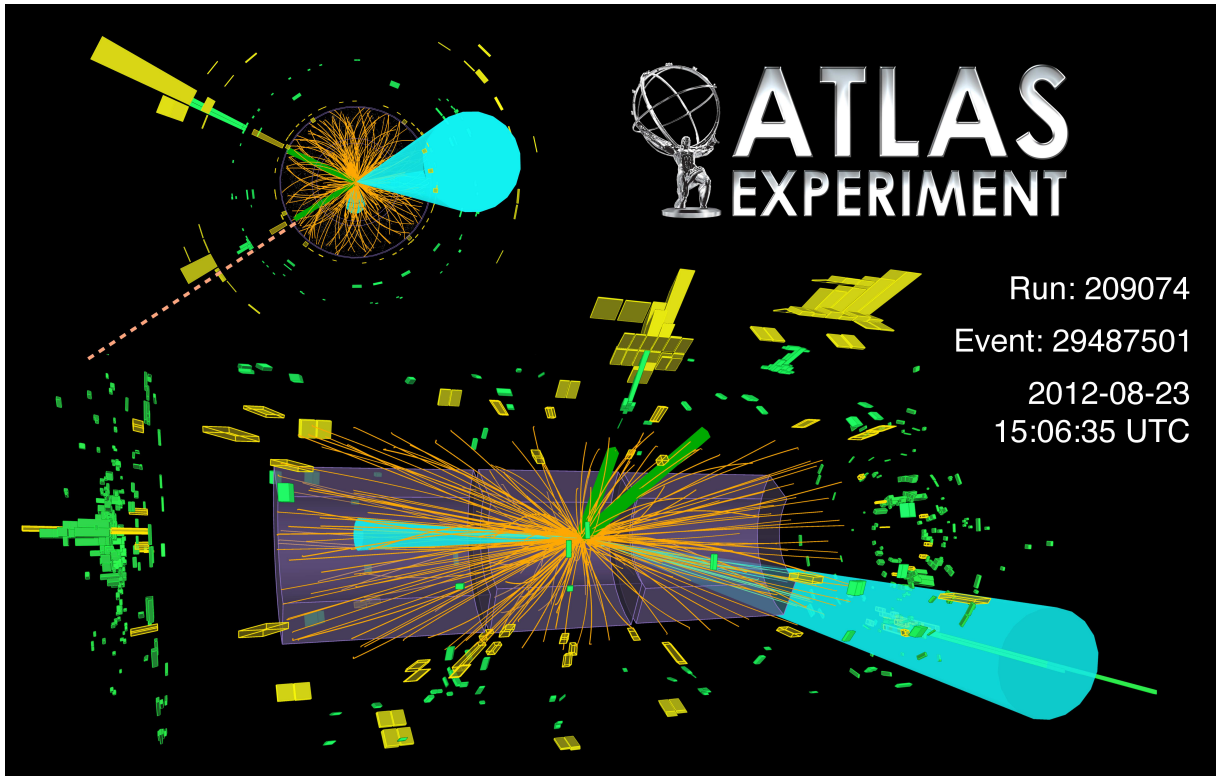
Figure 4.2: Display of a VBF Higgs boson candidate event selected in the $H \to \tau_{\text{jet}}\tau_{\text{jet}} + 2\nu$ search at $\sqrt{s} = 8$ TeV. The two τ-jets are indicated by green tracks, the dashed line in the upper left quadrant of the R-ϕ view represents the direction of the $\mathbf{E}_{\text{T}}^{\text{miss}}$ and the two tagging jets are marked with turquoise cones. The leading τ-jet p_{T} is 122 GeV, the sub-leading τ-jet p_{T} is 67 GeV, $E_{\text{T}}^{\text{miss}}$=72 GeV, M_{jj}=1.02 TeV and $m_{\tau\tau}^{\text{MMC}}$=126 GeV [67].
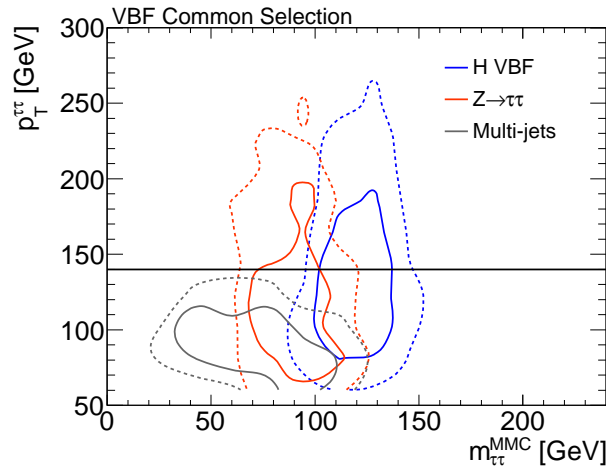
(a) $m_{\tau\tau}^{\mathrm{MMC}}$ vs. $p_{\mathrm{T}}^{\tau\tau}$
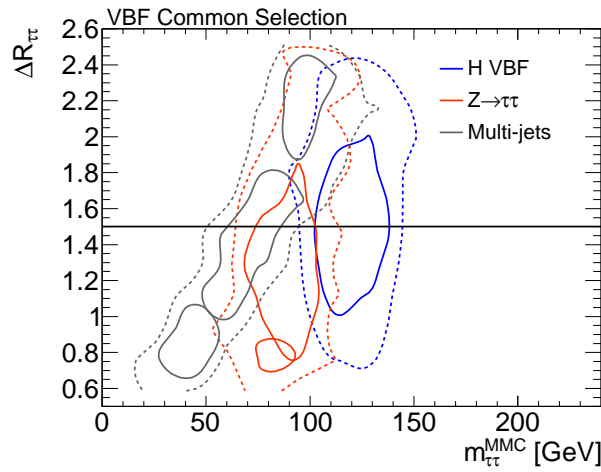
(b) $m_{\tau\tau}^{\mathrm{MMC}}$ vs. $\Delta R_{\tau\tau}$
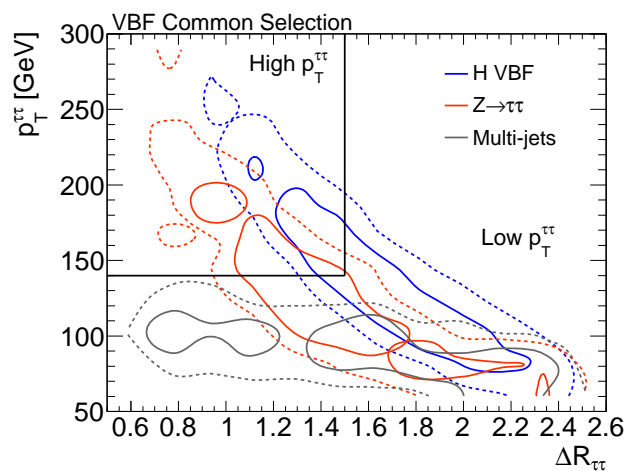
(c) $\Delta R_{\tau\tau}$ vs. $p_{\mathrm{T}}^{\tau\tau}$

Figure 4.3: Correlations among $p_{\mathrm{T}}^{\tau\tau}$, $\Delta R_{\tau\tau}$ and $m_{\tau\tau}^{\mathrm{MMC}}$ variables for VBF-like signal and background events. Contour plots indicate the regions containing 50% and 80% of the event yields. The black solid lines indicate requirements defining different VBF event categories. "H VBF" are simulated Higgs boson events with $m_H = 125$ GeV produced via VBF. The contribution of $Z \to \tau\tau$ and multi-jet events is predicted from control data samples (see Section 4.6).

the search range with $m_{\tau\tau}$ between 100 and 150 GeV is about 0.5.

VBF Low-$p_T^{\tau\tau}$ Categories: The remaining VBF-like events with small $p_T^{\tau\tau}$ or high $\Delta R_{\tau\tau}$ values enter the VBF Low-$p_T^{\tau\tau}$ category and are further categorised according to the kinematic properties of the tagging jets as shown in Fig. 4.5. Events with higher invariant mass M_{jj} or distance in pseudo-rapidity $|\Delta\eta_{jj}|$ of the tagging jets are more likely to originate from signal processes. Thus, the tightening of the requirements on these two variables with a diagonal cut in the $(|\Delta\eta_{jj}|, M_{jj})$ plane is used to define two VBF Low-$p_T^{\tau\tau}$ categories with different signal-to-background ratios:

> ▷ the **Tight** di-jet event category for events with $M_{jj}[\,\mathrm{GeV}] > -250|\Delta\eta_{jj}|+1550$ and the

> ▷ **Loose** di-jet event category for events with $M_{jj}[\,\mathrm{GeV}] < -250|\Delta\eta_{jj}| + 1550$

with signal-to-background ratios of 0.2 and 0.07, respectively. The dominant source of background events is the multi-jet production. Given the relatively small number of events, in the $\sqrt{s} = 7$ TeV dataset the Tight and Loose VBF Low-$p_T^{\tau\tau}$ categories are merged.

Boosted Event Categories

Events not accepted by any of the above VBF event categories are required to pass alternative selection criteria designed to select ggF signal events in which the Higgs boson has a high transverse momentum due to its recoil from a hard jet. Such "Boosted-like" events are selected by requiring $p_T^{\tau\tau} > 80$ GeV without imposing any specific requirement on jets. In addition, a tighter requirement is imposed on the $\mathbf{E}_T^{\mathrm{miss}}$ vector compared to the preselection stage, namely $\min\left\{\Delta\phi(E_T^{\mathrm{miss}}, \tau_{\mathrm{jet},1}), \Delta\phi(E_T^{\mathrm{miss}}, \tau_{\mathrm{jet},2})\right\} < 0.1\pi$ if the $\mathbf{E}_T^{\mathrm{miss}}$ vector is not in between the two τ-jets. This requirement is motivated by the fact that the decay products of the boosted Higgs boson are more collimated to each other and the vector of the missing transverse energy from the two neutrinos is even closer to the visible τ decay products.

The Boosted-like events are further categorised in two event categories based on selection criteria applied on $p_T^{\tau\tau}$ and $\Delta R_{\tau\tau}$ as in the VBF category, since there is a similar correlation among the two variables also for the Boosted-like events (see Fig. 4.6).

Boosted High-$p_T^{\tau\tau}$ Category: The most sensitive Boosted category is defined by requiring $p_T^{\tau\tau} > 140$ GeV and $\Delta R_{\tau\tau} < 1.5$. As already mentioned for the VBF High-$p_T^{\tau\tau}$ category, both requirements are needed in order to improve the rejection of background events in the sensitive $m_{\tau\tau}$ mass range. The signal-to-background ratio in this category is about 0.13. The expected background contribution originates mostly from Z+jets events.

Boosted Low-$p_T^{\tau\tau}$ Category: Boosted-like events with small $p_T^{\tau\tau}$ or high $\Delta R_{\tau\tau}$ values constitute the least sensitive event category. This category is characterised by a significant fraction of multi-jet events with signal-to-background ratio of about 0.04.
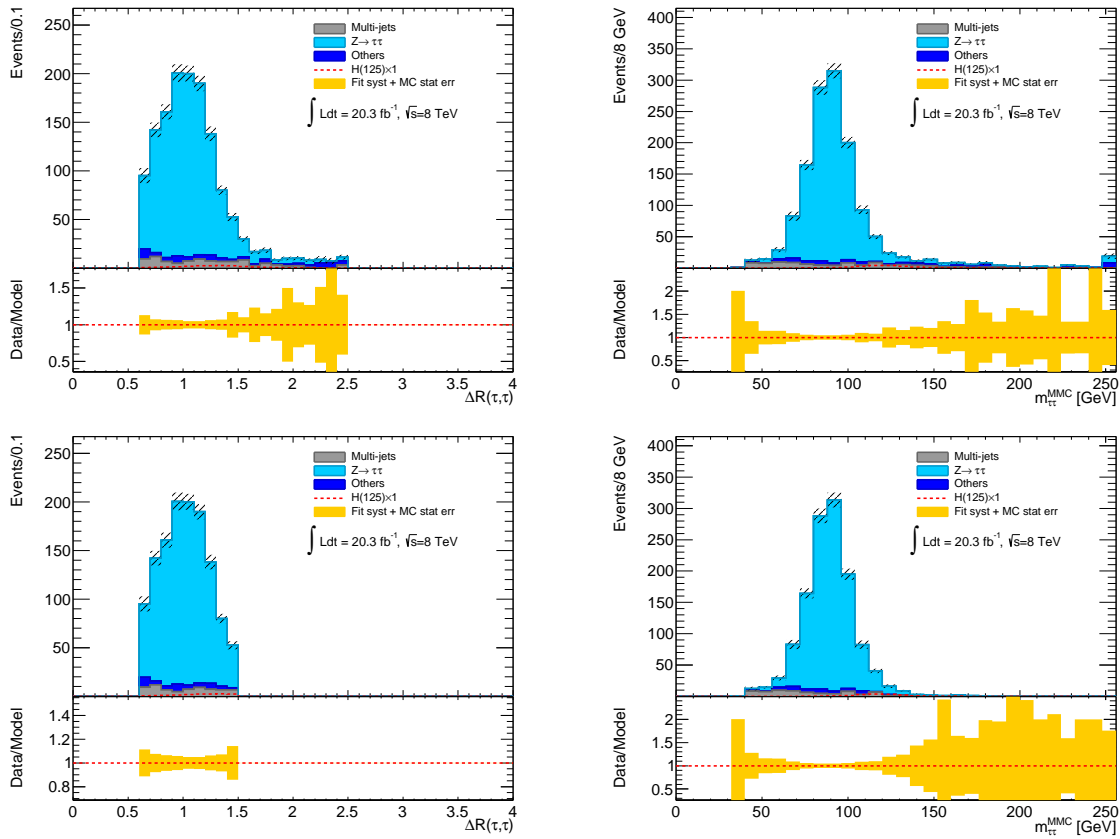
Figure 4.4: Comparison of the expected $\Delta R_{\tau\tau}$ (left) and $m_{\tau\tau}^{\mathrm{MMC}}$ (right) distributions for preselected signal and background events with $p_{\mathrm{T}}^{\tau\tau} > 140$ GeV and without (top) or with (bottom) the $\Delta R_{\tau\tau} < 1.5$ requirement applied.
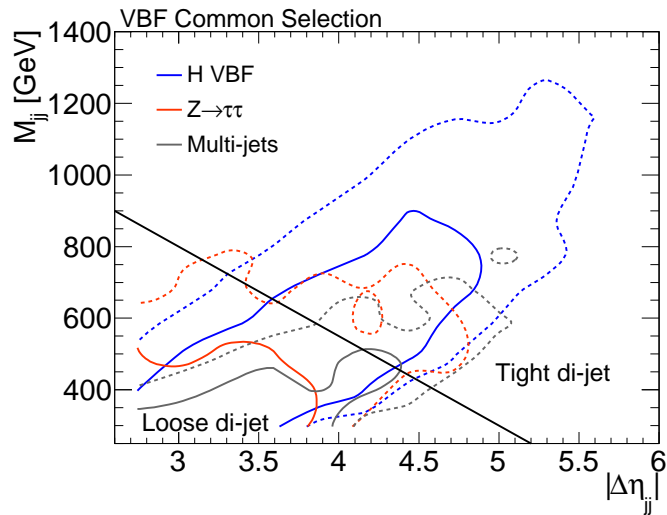
Figure 4.5: Fraction of VBF-like signal and background events in dependence on $|\Delta\eta_{jj}|$ and M_{jj}. Contour plots define the regions containing 50% and 80% of the event yields. The black solid line indicates the requirement defining the Tight and Loose VBF Low-$p_{\mathrm{T}}^{\tau\tau}$ event categories. "H VBF" indicates simulated Higgs boson events with $m_H = 125$ GeV produced via VBF. The contribution of $Z \to \tau\tau$ and multi-jet events is predicted from control data samples (see Section 4.6).

The Higgs boson production in association with a weak boson (VH) is not specifically targeted by this search. Nonetheless, non-negligible contributions (11-15%) from this production mode are expected in the Boosted categories and they are treated as additional signal.

Table 4.7 lists the event yields at each stage of the event selection, separately for the $\sqrt{s} = 7$ TeV and 8 TeV datasets. Table 4.8 reports the final event yields after the full selection in each event category. Only statistical uncertainties are reported. The yields indicated for the $Z \to \tau\tau$ and the multi-jet background processes are preliminary estimates derived from data, as described in Section 4.6. The final expected background yields are obtained after the fit of the invariant $m_{\tau\tau}$ mass distribution to data, as described in Section 4.8.
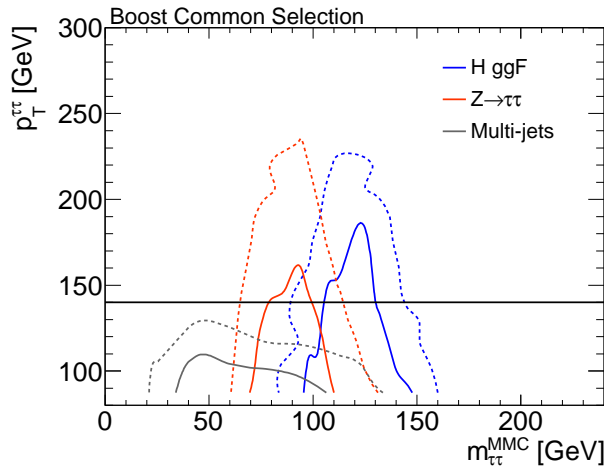
(a) $m_{\tau\tau}^{\mathrm{MMC}}$ vs. $p_{\mathrm{T}}^{\tau\tau}$
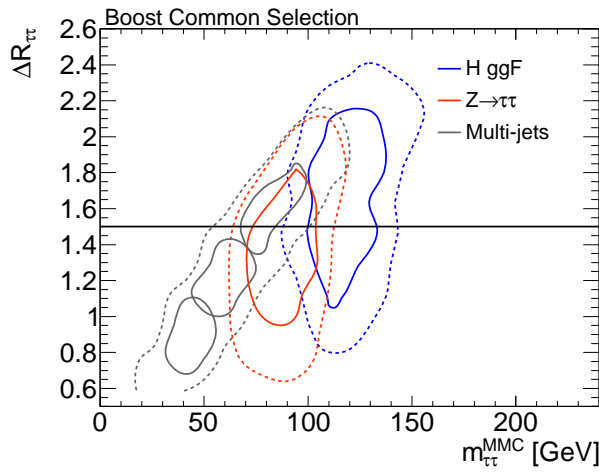
(b) $m_{\tau\tau}^{\mathrm{MMC}}$ vs. $\Delta R_{\tau\tau}$
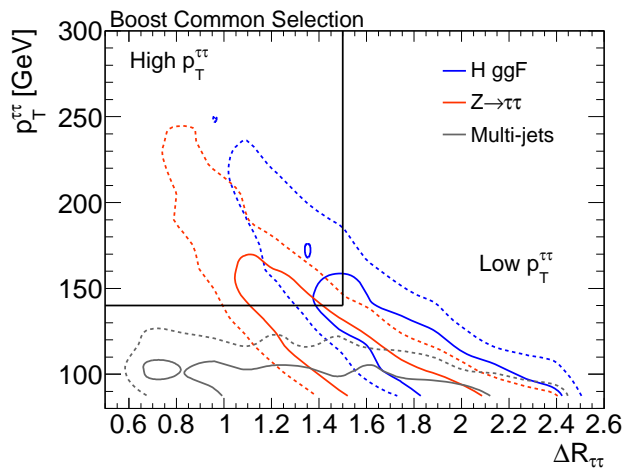
(c) $\Delta R_{\tau\tau}$ vs. $p_{\mathrm{T}}^{\tau\tau}$

Figure 4.6: Correlations among $p_{\mathrm{T}}^{\tau\tau}$, $\Delta R_{\tau\tau}$ and $m_{\tau\tau}^{\mathrm{MMC}}$ variables for Boosted-like signal and background events. Contour plots indicate the regions containing 50% and 80% of the event yields. The black solid lines indicate requirements defining different VBF event categories. "H ggF" are simulated Higgs boson events with $m_H = 125$ GeV produced via VBF. The contribution of $Z \to \tau\tau$ and multi-jet events is predicted from control data samples (see Section 4.6).

Table 4.7: Observed and expected event yields at each stage of the event selection in the $\sqrt{s} = 7$ and 8 TeV datasets. The rows in bolded font indicate the final event categories used for the search. The errors include only statistical uncertainties. The $Z \to \tau\tau$ and the multi-jet background event yields are estimated from control data samples, while the expected contributions of the remaining background processes are obtained from simulation.

	Data	$H(m_H{=}125$ GeV$)$	ggF	VBF	VH	Z	Multi-jet	$W+j$	top	VV'
					$\sqrt{s} = 7$ TeV, $\int$ Ldt $= 4.6\,\mathrm{fb}^{-1}$					
Preselection	1994	9.6 ± 0.3	6.1 ± 0.3	2.5 ± 0.1	1.1 ± 0.1	950 ± 10	1300 ± 20	22 ± 2	11 ± 1	7.7 ± 0.8
≥ 2 jets	469	4.3 ± 0.2	2.2 ± 0.2	1.5 ± 0.1	0.68 ± 0.04	290 ± 8	160 ± 7	7 ± 1	8.4 ± 0.6	4.9 ± 0.6
$\Delta\eta_{jj} > 2.6 + M_{jj} > 250$ GeV	60	1.7 ± 0.1	0.53 ± 0.08	1.1 ± 0.04	0.026 ± 0.007	42 ± 3	28 ± 3	1.6 ± 0.7	0.9 ± 0.2	0.2 ± 0.1
$\eta_{\mathrm{min,j}} < \tau, \tau < \eta_{\mathrm{max,j}}$	40	1.5 ± 0.1	0.42 ± 0.07	1.1 ± 0.04	0.02 ± 0.01	25 ± 2	21 ± 2	1.5 ± 0.7	0.68 ± 0.2	0.15 ± 0.09
VBF High-$p_\mathrm{T}^{\tau\tau}$	11	0.71 ± 0.06	0.25 ± 0.05	0.45 ± 0.03	0.012 ± 0.005	9 ± 1	0.9 ± 0.6	0.2 ± 0.2	0.1 ± 0.1	0.02 ± 0.02
VBF Low-$p_\mathrm{T}^{\tau\tau}$	29	0.8 ± 0.1	0.16 ± 0.04	0.63 ± 0.03	0.008 ± 0.004	16 ± 2	20 ± 2	1.3 ± 0.7	0.6 ± 0.1	0.14 ± 0.09
$\Delta\phi_{\mathrm{min}}(\tau, \mathrm{E_T^{miss}}) < 0.1\pi$	1647	7.5 ± 0.3	5.3 ± 0.3	1.3 ± 0.1	0.97 ± 0.05	820 ± 10	790 ± 10	15 ± 2	8.3 ± 0.6	7.2 ± 0.7
$\mathrm{p_T^{Higgs}} > 80$ GeV	942	5.8 ± 0.2	3.9 ± 0.2	1.1 ± 0.04	0.83 ± 0.04	540 ± 10	350 ± 10	11 ± 2	7.2 ± 0.6	6.2 ± 0.7
Boosted High-$p_\mathrm{T}^{\tau\tau}$	182	2 ± 0.1	1.3 ± 0.1	0.33 ± 0.02	0.39 ± 0.03	200 ± 7	11 ± 2	2.9 ± 0.9	0.7 ± 0.2	3 ± 0.5
Boosted Low-$p_\mathrm{T}^{\tau\tau}$	760	3.8 ± 0.2	2.6 ± 0.2	0.73 ± 0.03	0.44 ± 0.03	330 ± 8	340 ± 9	8.5 ± 2	6.5 ± 0.5	3.2 ± 0.5
					$\sqrt{s} = 8$ TeV, $\int$ Ldt $= 20.3\,\mathrm{fb}^{-1}$					
Preselection	10433	72 ± 1	45 ± 1	17 ± 0.1	6.7 ± 0.1	4900 ± 30	5200 ± 60	210 ± 30	120 ± 7	14 ± 0.4
≥ 2 jets	3271	36 ± 0.4	18 ± 0.4	11 ± 0.1	4.8 ± 0.1	1700 ± 20	1300 ± 30	71 ± 20	100 ± 7	8.1 ± 0.3
$\Delta\eta_{jj} > 2.6 + M_{jj} > 250$ GeV	478	13 ± 0.2	4 ± 0.2	8.1 ± 0.1	0.11 ± 0.01	230 ± 8	240 ± 10	5 ± 3	9.7 ± 2	0.7 ± 0.1
$\eta_{\mathrm{min,j}} < \tau, \tau < \eta_{\mathrm{max,j}}$	335	11 ± 0.2	3 ± 0.2	7.6 ± 0.1	0.07 ± 0.01	140 ± 6	170 ± 10	1.8 ± 1	6.5 ± 2	0.52 ± 0.08
VBF High-$p_\mathrm{T}^{\tau\tau}$	74	4.4 ± 0.1	1.0 ± 0.1	3.1 ± 0.1	0.028 ± 0.005	55 ± 4	4.6 ± 2	0.7 ± 0.7	1.6 ± 0.9	0.14 ± 0.04
VBF Low-$p_\mathrm{T}^{\tau\tau}$ Tight	97	3.8 ± 0.1	0.86 ± 0.09	2.9 ± 0.04	0.008 ± 0.003	26 ± 2	60 ± 5	0 ± 0	0.6 ± 0.6	0.11 ± 0.04
VBF Low-$p_\mathrm{T}^{\tau\tau}$ Loose	164	2.8 ± 0.1	1.2 ± 0.1	1.5 ± 0.03	0.034 ± 0.006	61 ± 4	100 ± 8	1 ± 1	4.4 ± 1	0.27 ± 0.06
$\Delta\phi_{\mathrm{min}}(\tau, \mathrm{E_T^{miss}}) < 0.1\pi$	8227	54 ± 1	36 ± 1	8.8 ± 0.1	6 ± 0.1	4100 ± 30	3800 ± 50	160 ± 20	87 ± 6	12 ± 0.4
$\mathrm{p_T^{Higgs}} > 80$ GeV	5029	41 ± 0.5	26 ± 0.5	7.1 ± 0.1	5.3 ± 0.1	2800 ± 30	1800 ± 30	110 ± 20	77 ± 6	9.6 ± 0.4
Boosted High-$p_\mathrm{T}^{\tau\tau}$	1293	16 ± 0.3	8.6 ± 0.3	2.5 ± 0.04	2.4 ± 0.05	1100 ± 20	56 ± 8	24 ± 8	8.4 ± 2	3.9 ± 0.2
Boosted Low-$p_\mathrm{T}^{\tau\tau}$	3736	25 ± 0.4	17 ± 0.4	4.5 ± 0.05	2.9 ± 0.05	1800 ± 20	1800 ± 30	84 ± 20	69 ± 5	5.7 ± 0.3

4.5   Invariant Mass Reconstruction

The invariant mass $m_{\tau\tau}$ of the two τ leptons is the final discriminant for the statistical interpretation of the observed data in terms of the measured signal cross section. It also allows for the measurement of the mass of the potentially observed signal resonance.

Given the presence of two ν_τ neutrinos in the final state, the $m_{\tau\tau}$ invariant mass cannot be directly computed since the contributions of these neutrinos to the four-momenta of the τ leptons are not directly measured. Nonetheless, a rather good approximation of the $m_{\tau\tau}$ mass can be obtained for Higgs boson decays into τ leptons due to the strong boost of the two τ leptons which are much lighter than the Higgs boson. Accordingly, in each τ lepton decay the corresponding neutrino and τ-jet are emitted in a narrow cone around the direction of the mother τ lepton.

In the simplest implementation, the $m_{\tau\tau}$ is reconstructed in the collinear approximation [118] assuming that each neutrino is emitted in the same direction as the corresponding τ-jet, $\mathbf{p}_{\nu,\mathrm{i}}/|\mathbf{p}_{\nu,\mathrm{i}}| = \mathbf{p}_{\tau-\mathrm{jet,i}}/|\mathbf{p}_{\tau-\mathrm{jet,i}}|$. It is also assumed that only the two neutrinos from the τ lepton decays contribute to the total missing transverse energy reconstructed in the event, $\mathbf{E}_\mathrm{T}^\mathrm{miss} = \mathbf{p}_{\nu,1} + \mathbf{p}_{\nu,2}$. Under these assumptions the energies $E_{\nu,\mathrm{i}}$ of the two neutrinos can be computed from the measurement of the $\mathbf{E}_\mathrm{T}^\mathrm{miss}$ components using

$$
\begin{aligned}
E_\mathrm{x}^\mathrm{miss} &= p_{\nu,1,\mathrm{x}} + p_{\nu,2,\mathrm{x}} = E_{\nu,1}\frac{p_{\tau-\mathrm{jet,1,x}}}{|\mathbf{p}_{\tau-\mathrm{jet,1}}|} + E_{\nu,2}\frac{p_{\tau-\mathrm{jet,2,x}}}{|\mathbf{p}_{\tau-\mathrm{jet,2}}|} \\[2mm]
E_\mathrm{y}^\mathrm{miss} &= p_{\nu,1,\mathrm{y}} + p_{\nu,2,\mathrm{y}} = E_{\nu,1}\frac{p_{\tau-\mathrm{jet,1,y}}}{|\mathbf{p}_{\tau-\mathrm{jet,1}}|} + E_{\nu,2}\frac{p_{\tau-\mathrm{jet,2,y}}}{|\mathbf{p}_{\tau-\mathrm{jet,2}}|}
\end{aligned}
\tag{4.1}
$$

where $p_{\nu,\mathrm{i,x(y)}}$ is the momentum components of the neutrinos in the x(y)-direction and $\mathbf{p}_{\tau-\mathrm{jet,i}}$ is the spatial momentum of the τ-jet. This system of equations can be solved only if the two τ-jets are not emitted back-to-back in the transverse plane since otherwise the contributions of the two neutrinos to the missing transverse energy cancel out. Once the energies $E_{\nu,1}$ and $E_{\nu,2}$ are determined, the four-momentum of each τ lepton is defined as $p_{\tau,\mathrm{i}} = (E_{\nu,\mathrm{i}} + E_{\tau-\mathrm{jet,i}})\left(1, \frac{\mathbf{p}_{\tau-\mathrm{jet,i}}}{|\mathbf{p}_{\tau-\mathrm{jet,i}}|}\right)$ where $E_{\tau-\mathrm{jet,i}}$ is the energy of the τ-jet. The invariant $\tau\tau$ mass is then determined as $m_{\tau\tau}^2 = (p_{\tau,1} + p_{\tau,2})^2$.

The reconstructed $m_{\tau\tau}$ mass obtained with the collinear approximation strongly depends on the momentum of the Higgs boson and on the measured $E_\mathrm{T}^\mathrm{miss}$. The mass resolution worsens the lower is the momentum of the $\tau\tau$ system and in events with small $E_\mathrm{T}^\mathrm{miss}$ values the probability for unphysical solutions of Eq. 4.1 is higher. Due to these shortcomings, the collinear approximation is a sub-optimal mass reconstruction algorithm applicable to only a fraction of events (see Fig. 4.8).

An improved algorithm implemented in the Missing Mass Calculator (MMC) programme [98] can be applied to any event topology. Relaxing the assumptions of the collinear approximation, the neutrino four-momenta are defined by six unknown variables, three for each neutrino: $(p_x, p_y, p_z)_{\nu,1}$ and $(p_x, p_y, p_z)_{\nu,2}$. These can be only partially constrained by 4 external parameters, namely the measured $E_\mathrm{x}^\mathrm{miss}$ and $E_\mathrm{y}^\mathrm{miss}$ components and the masses of the two τ leptons $m_{\tau,1}$ and $m_{\tau,2}$ [25]. The most likely among the infinite solutions is determined based on the kinematic properties of the τ lepton decay products, i.e. by comparing for each solution the three-dimensional angle $\Delta\theta_\mathrm{3D} = \cos^{-1}(\hat{\mathbf{p}}_{\tau-\mathrm{jet}} \cdot \hat{\mathbf{p}}_\nu)$

Table 4.8: Summary of the observed and expected signal and background event yields in each event category in the $\sqrt{s} = 7$ TeV and 8 TeV datasets. The errors include only statistical uncertainties. The $Z \rightarrow \tau\tau$ and the multi-jet event yields are estimated from control data samples, while the expected contributions of the remaining background processes are obtained from simulation.

$\sqrt{s} = 7$ TeV	VBF		Boosted	
$\int \mathrm{Ldt} = 4.6\,\mathrm{fb}^{-1}$	High-$p_\mathrm{T}^{\tau\tau}$	Low-$p_\mathrm{T}^{\tau\tau}$	High-$p_\mathrm{T}^{\tau\tau}$	Low-$p_\mathrm{T}^{\tau\tau}$
H($m_H = 125$ GeV)	0.75 ± 0.06	0.88 ± 0.06	2 ± 0.1	3.8 ± 0.2
$Z \rightarrow \tau\tau$	9 ± 1	16 ± 2	202 ± 7	335 ± 8
Multi-jet	0.9 ± 0.6	20 ± 2	11 ± 2	340 ± 9
$W, t, t\bar{t}, VV'$	0.3 ± 0.2	2.2 ± 0.7	7.2 ± 1	21 ± 2
Total Bkg.	10 ± 2	39 ± 3	220 ± 7	691 ± 12
Data	11	29	182	760

$\sqrt{s} = 8$ TeV	VBF			Boosted	
$\int \mathrm{Ldt} = 20.3\,\mathrm{fb}^{-1}$	High-$p_\mathrm{T}^{\tau\tau}$	Low-$p_\mathrm{T}^{\tau\tau}$ Tight	Low-$p_\mathrm{T}^{\tau\tau}$ Loose	High-$p_\mathrm{T}^{\tau\tau}$	Low-$p_\mathrm{T}^{\tau\tau}$
H($m_H = 125$ GeV)	4.1 ± 0.1	3.8 ± 0.1	2.8 ± 0.1	14 ± 0.3	25 ± 0.4
$Z \rightarrow \tau\tau$	55 ± 4	26 ± 2	61 ± 4	1060 ± 18	1780 ± 21
Multi-jet	5 ± 2	60 ± 5	103 ± 8	56 ± 8	1790 ± 33
$W, t, t\bar{t}, VV'$	2 ± 1	0.7 ± 0.7	7 ± 2	40 ± 9	180 ± 20
Total Bkg.	62 ± 5	86 ± 6	171 ± 9	1160 ± 22	3750 ± 44
Data	74	97	164	1293	3736

between the normalised spatial vectors of the τ-jet and the neutrino with the corresponding probability density function. These probability density functions $\mathcal{P}(\Delta\theta_{3D}, p_\tau)$ (see Fig. 4.7) are obtained from the simulated $Z/\gamma^* \to \tau\tau$ process in dependence on the τ lepton momentum p_τ, separately for the 1-prong or 3-prong τ decays.
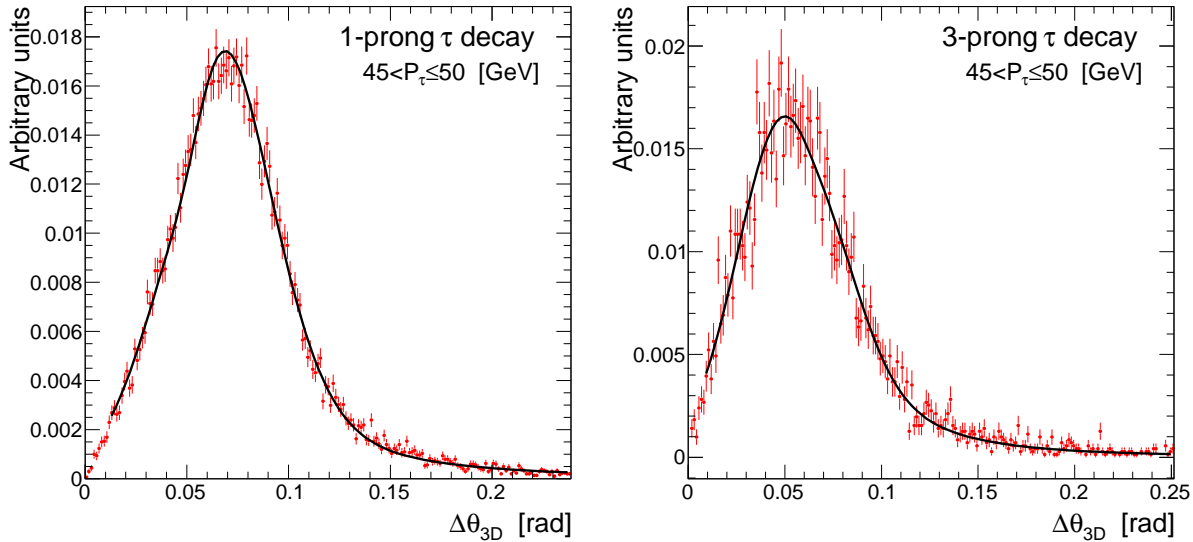

Figure 4.7: Probability density functions $\mathcal{P}(\Delta\theta_{3D}, p_\tau)$ obtained from simulated $Z/\gamma^* \to \tau\tau$ samples, shown separately for the 1-prong (left) and 3-prong (right) τ decays for τ leptons with momenta in the range $45 < p_\tau \leq 50$ GeV. $\mathcal{P}(\Delta\theta_{3D}, p_\tau)$ are fitted with a linear combination of a Gaussian and a Landau function [98].

In its simplest implementation the MMC algorithm computes the solutions for the neutrino momenta for any point in the $(\phi_{\nu,1}, \phi_{\nu,2})$ plane and then calculates the corresponding $\Delta\theta_{3D,1}$ and $\Delta\theta_{3D,2}$ angles and the invariant mass $m_{\tau\tau}$ of the two τ leptons. For each point on the plane the computed $m_{\tau\tau}$ is weighted by the event probability $\mathcal{P}_{\text{event}} = \mathcal{P}(\Delta\theta_{3D,1}, p_{\tau,1}) \times \mathcal{P}(\Delta\theta_{3D,2}, p_{\tau,2})$. The best estimate of the $\tau\tau$ invariant mass $m_{\tau\tau}^{\text{MMC}}$ is defined as the mass corresponding to the maximum of the weighted $m_{\tau\tau}$ distribution.

To improve the $m_{\tau\tau}$ resolution, the MMC algorithm can also take into account the limited resolution of the $E_{\text{T}}^{\text{miss}}$ measurement introducing as additional parameters the true values of the $\mathbf{E}_{\text{T}}^{\text{miss}}$ components $E_{\text{x(y)}}^{\text{miss,truth}}$. The parameter space $(\phi_{\nu,1}, \phi_{\nu,2}, E_{\text{x}}^{\text{miss,truth}}, E_{\text{y}}^{\text{miss,truth}})$ has then four dimensions and the probability density function is given as $\mathcal{P}_{\text{event}} = \mathcal{P}(\Delta\theta_{3D,1}, p_{\tau,1}) \times \mathcal{P}(\Delta\theta_{3D,2}, p_{\tau,2}) \times \mathcal{P}(\Delta E_{\text{x}}^{\text{miss}}) \times \mathcal{P}(\Delta E_{\text{y}}^{\text{miss}})$. The resolution function $\mathcal{P}(\Delta E_{\text{x(y)}}^{\text{miss}})$ is defined as

$$\mathcal{P}(\Delta E_{\text{x(y)}}^{\text{miss}}) = \exp\left\{-\frac{(\Delta E_{\text{x(y)}}^{\text{miss}})^2}{2\sigma^2}\right\} \tag{4.2}$$

where σ is the $E_{\text{T}}^{\text{miss}}$ resolution determined from dedicated calibration data samples [119] and $\Delta E_{\text{x(y)}}^{\text{miss}}$ is the difference between the measured and the true $E_{\text{T}}^{\text{miss}}$ value at a given point of the $(\phi_{\nu,1}, \phi_{\nu,2}, E_{\text{x}}^{\text{miss,truth}}, E_{\text{y}}^{\text{miss,truth}})$ parameter space. Given the limited processing

time and computing resources, the MMC algorithm is applied only on a grid of discrete points of the parameter space. The fraction of events for which the MMC algorithm fails to find the maximum of the weighted $m_{\tau\tau}$ distribution is small, on the order of few percent for signal events.

Fig. 4.8 compares the invariant mass distributions obtained with the collinear approximation and with the MMC algorithm, separately for $Z \to \tau\tau$ and $H \to \tau\tau$ events with $m_H = 125$ GeV. The MMC algorithm has a lower rate of unphysical solutions (first bin
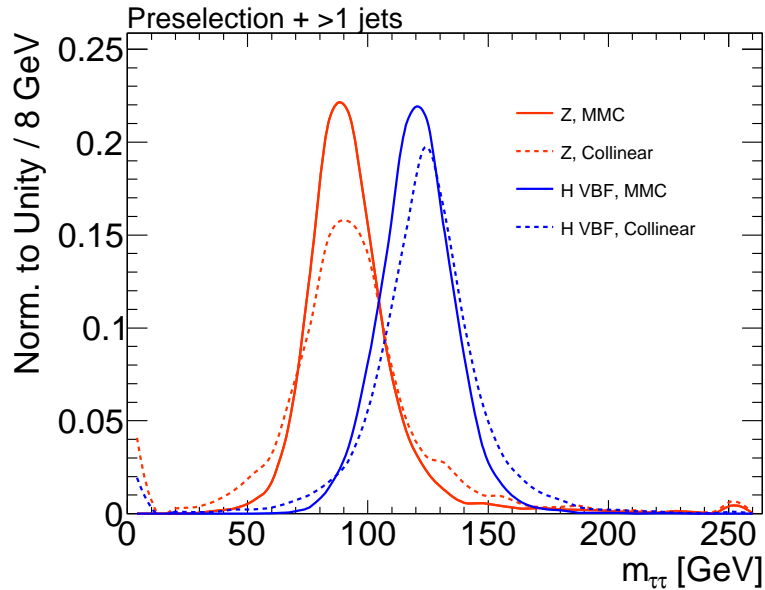

Figure 4.8: $m_{\tau\tau}$ invariant mass distributions computed with MMC algorithm (solid line) and with the collinear approximation (dashed line) in fully hadronic $Z \to \tau\tau$ (red) and $H \to \tau\tau$ events with $m_H = 125$ GeV (blue) produced at $\sqrt{s} = 8$ TeV. The distributions are obtained with preselected events with at least two jets in the final state. The first bin contains the events with an unphysical solution for the invariant mass.

in Fig. 4.8) and a better mass resolution. Particularly relevant for the sensitivity of this search is the suppression of the high-mass tail of the Z boson distribution with the MMC algorithm.

In Fig. 4.9, the $m_{\tau\tau}^{\mathrm{MMC}}$ distributions obtained for Higgs boson events with different masses m_H are compared to each other and to the Z boson distribution. The mean value and resolution of the $m_{\tau\tau}^{\mathrm{MMC}}$ distribution estimated with a Gaussian fit are shown in Fig. 4.10 as a function of the Higgs boson mass m_H. The reconstructed $m_{\tau\tau}^{\mathrm{MMC}}$ mean value is on average few percent smaller than the true m_H because the probability density functions used by the MMC algorithm are obtained from simulated $Z/\gamma^* \to \tau\tau$ events and differ from the corresponding probability density functions for Higgs boson events. Even though the mass measurement is not exact, the linear correspondence between the reconstructed mean $m_{\tau\tau}^{\mathrm{MMC}}$ value and m_H can be used to determine the mass m_H of the potentially observed Higgs boson signal. For this purpose, the $m_{\tau\tau}^{\mathrm{MMC}}$ template distributions obtained for different mass hypotheses are compared to the ones observed in data separately for
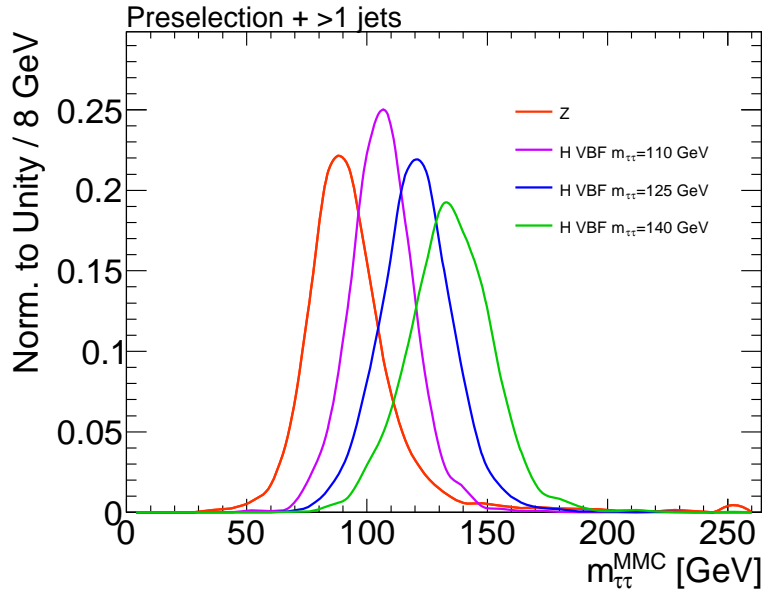
Figure 4.9: $m_{\tau\tau}^{\mathrm{MMC}}$ distributions in fully hadronic $Z \to \tau\tau$ and $H \to \tau\tau$ events with different m_H at $\sqrt{s} = 8$ TeV. The distributions are shown for the preselected events with at least two jets in the final state.
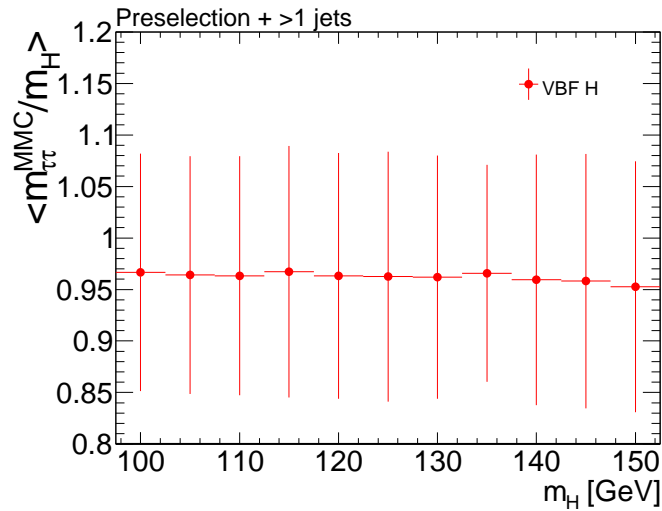
Figure 4.10: $m_{\tau\tau}^{\mathrm{MMC}}$ mass resolution of preselected VBF $H \to \tau\tau$ events with at least two jets in the final states in the $\sqrt{s} = 8$ TeV simulation. The mean value of $m_{\tau\tau}^{\mathrm{MMC}}/m_H$ and its variance (error bars) obtained from the Gaussian fit to the $m_{\tau\tau}^{\mathrm{MMC}}$ distributions are shown as a function of m_H.

each event category. The measured Higgs boson mass is then given as the hypothesised m_H value for which the $m_{\tau\tau}^{\mathrm{MMC}}$ distribution provides the best statistical agreement with data.

4.6   Measurement of the Background Contributions

An accurate prediction of the background contributions is essential for the statistical interpretation of the observed data. This prediction includes the event yields in each event category together with the modelling of the discriminating variables, including those used for the event selection and the $m_{\tau\tau}^{\mathrm{MMC}}$ variable for the final result. Both the event yield and the modelling of the dominant $Z \to \tau\tau$ and multi-jet background processes are determined using data-driven techniques. The small contribution from other background processes with at least one hadronically decaying τ lepton, including mostly W+jets and $t\bar{t}$ events, is predicted based on simulation.

The modelling of the $Z \to \tau\tau$ and multi-jet backgrounds is determined using control data samples signal-depleted and enriched with events from the respective background process, as described next in Sections 4.6.1 and 4.6.2.

The yields of the $Z \to \tau\tau$ and multi-jet contributions are measured in data fitting the normalisation of the expected $m_{\tau\tau}^{\mathrm{MMC}}$ distributions to the data observed after the full event selection simultaneously in all event categories. The yield of the multi-jet events has to be measured in this way, since the multi-jet control data samples do not provide a reliable information on the background yield (see Section 4.6.2). In addition, even though the $Z \to \tau\tau$ event yield can be obtained directly from the corresponding control data sample, the mentioned fit gives more precise results as the systematic uncertainties on the Z boson production cross section and on the event selection efficiency can be avoided. The fit is performed as part of the statistical procedure for the interpretation of the observed data. The expected $m_{\tau\tau}^{\mathrm{MMC}}$ distributions from signal and background processes are summed together and are fitted to the corresponding observed data in each event category with the normalisations of the $Z \to \tau\tau$ and multi-jet background distributions set as free parameters of the fit. The normalisation factor for a given background process is assumed to be common to all event categories. The discrimination power between the $Z \to \tau\tau$ and multi-jet contributions is provided by the corresponding $m_{\tau\tau}^{\mathrm{MMC}}$ distributions since the low tail of this distribution is dominated by multi-jet events (see Fig. 4.15a). This is sufficient to separately determine the contributions of the two background processes, together with the contribution of the potential signal.

The normalisation factors are determined separately for the $\sqrt{s} = 7$ and 8 TeV datasets. The relative background contributions in each event category are determined by the respective background models, as described in the following. Systematic uncertainties due to the migration of events across categories are taken into account (see Section 4.7.1).

4.6.1 Modelling of the $Z \to \tau\tau$ Background Process

$Z \to \tau\tau$ is the irreducible source of background events with two τ-jets originating from hadronically decaying τ leptons produced in a resonant decay. Due to the same final state topology, this process can be effectively distinguished from signal events only by the different $m_{\tau\tau}$ invariant masses. Since signal events are expected to contribute to the high tail of the $m_{\tau\tau}$ mass distribution above the Z boson peak, it is crucial to precisely model the kinematic event properties which impact the shape of the Z mass distribution, such as the transverse momentum of the Z boson and the number of jets in the event.

In order to reduce the systematic uncertainties originating from the simulation of the detector response, the contribution of this background is determined in a hybrid approach by combining simulation with measurements in a control data sample. This control data sample contains $Z \to \mu\mu$ events which are selected with high efficiency and purity and can be assumed to be free of signal contributions. The two muons from the Z boson decay in the control data sample are replaced by two simulated hadronically decaying τ leptons while ensuring that the kinematic properties of the Z boson and its decay products remain preserved. In such a so-called "embedded" event the decays of the τ leptons from the Z boson are described by simulation while all other event properties, such as the Z boson transverse momentum, the jets produced in the hard-scattering process as well as the underlying event and pile-up interactions are directly given by data. The described hybrid approach is needed since it is not possible to select a pure control sample of $Z \to \tau\tau$ events free of signal contributions. Since the $Z \to \mu\mu$ events are selected inclusively without any jet requirement, all Z boson production modes contribute to this control data sample with unbiased relative cross sections. In particular, the irreducible background from the electroweak Z boson production in which the Z boson is produced via vector boson fusion is accounted for.

In the following, the embedding procedure [120] is described in more detail. The observed $Z \to \mu\mu$ events are selected by single and di-muon triggers and by requiring the presence of exactly two oppositely charged well reconstructed and isolated muons (Fig. 4.11a). Both muons should have a transverse momentum of $p_\mathrm{T}^\mu > 15$ GeV. The p_T-threshold for the highest-p_T muon in 8 TeV data is raised to $p_\mathrm{T}^\mu > 20$ GeV to account for the higher trigger threshold. The invariant $\mu^+\mu^-$ mass is required to be $m_{\mu\mu} > 40$ GeV. The Z boson decay into two hadronically decaying τ leptons is simulated based on the reconstructed four-momentum of the $\mu^+\mu^-$ system. The τ lepton decays are simulated with the TAUOLA and PHOTOS packages and the decay products are processed through the detailed simulation of the ATLAS detector (Fig. 4.11b). The two reconstructed muons from the observed $Z \to \mu\mu$ event as well as all associated inner detector tracks and the expected muon energy deposits in the calorimeter are replaced by the simulated $Z \to \tau\tau$ decay products. The resulting embedded event (Fig. 4.11c) is fully reconstructed and all physics objects, including $E_\mathrm{T}^\mathrm{miss}$, are re-reconstructed based on the modified energy deposits and tracks. In order to increase the size of the embedded sample, only hadronic τ lepton decays are simulated. Small background contribution from the semi-leptonic $Z \to \tau\tau \to \tau_\mathrm{jet} l + 3\nu$ decay is predicted from simulation using the ALPGEN generator. The described embedding procedure is validated by applying the same procedure to simulated instead of observed $Z \to \mu\mu$ events and comparing the resulting distributions of

(a) $pp \to Z \to \mu\mu + X$ data

(b) $Z \to \tau\tau$ simulation

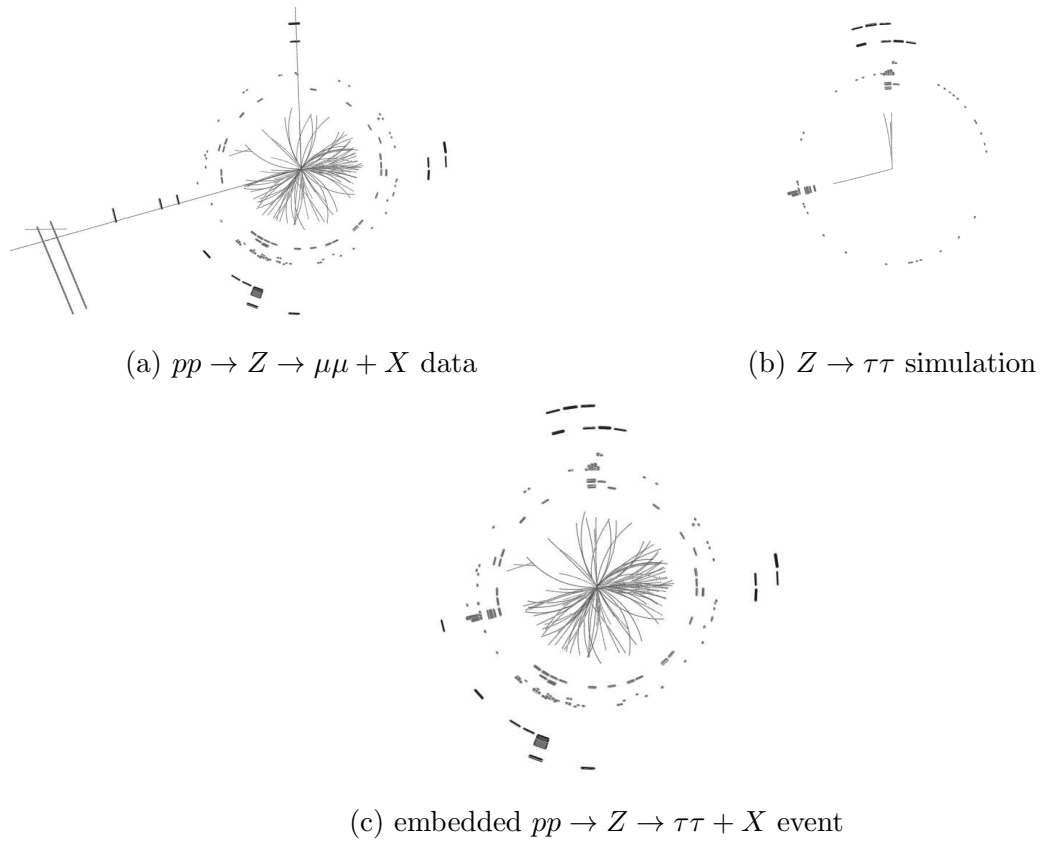(c) embedded $pp \to Z \to \tau\tau + X$ event

Figure 4.11: Event displays in the transverse plane illustrating the embedding procedure.

the discriminating variables to those obtained with simulated $Z \to \tau\tau$ events.

The sample of $Z \to \tau\tau$ events obtained with the embedding procedure has to be corrected for several effects in order to provide a good description of the real $Z \to \tau\tau$ data.

▷ The number of selected $Z \to \mu\mu$ events is affected by the muon trigger and reconstruction inefficiencies, which are not affecting the observed $Z \to \tau\tau$ data.

▷ The selection efficiency of the $Z \to \tau\tau$ data will be affected by the efficiencies of the double-hadronic-tau trigger and of the τ-jet identification, which are not affecting the selection of $Z \to \mu\mu$ data.

▷ The polarisation of the τ leptons is taken into account, but impacts the τ decay properties.

The inefficiencies of the $Z \to \mu\mu$ trigger selection and muon reconstruction have to be corrected in order to obtain a sample of Z boson events unbiased by the detector acceptance effects. This is done by reweighting each $Z \to \mu\mu$ event with the reciprocal of the muon trigger and reconstruction efficiency for each τ-jet. The weights are binned in muon p_{T} and η and measured with $Z \to \mu\mu$ calibration data [85].

The observed fully hadronic $Z \to \tau\tau$ decays are triggered by the double-hadronic-tau trigger but this trigger is not applied in the embedding procedure. Therefore, the embedded events have to be weighted by the double-hadronic-tau trigger efficiency measured with $Z \to \mu\tau_{\mathrm{jet}} + 3\nu$ and $W \to \tau_{\mathrm{jet}} + 2\nu$ data (see Section 3.3.1). The τ-jet identification efficiency is accounted for in the simulation of the τ lepton decays, but like in the other simulated samples, correction factors are applied as event weight to account for the difference of the τ-jet identification efficiency in simulation and calibration data.

A correction accounting for the τ lepton polarisation effects is needed since the TAUOLA package cannot predict the correct relative fractions of embedded $Z \to \tau\tau$ events with two left-handed (LL) and two right-handed (RR) τ leptons. While the spin correlation in the $Z \to \tau\tau$ decay is correctly taken into account, i.e. only LL and RR events are produced, the probabilities for these two event configurations are not correctly defined. These probabilities depend on the scattering angle of the Z boson decay and on the center-of-mass energy for the Z boson production. At LHC the average polarisation is known to be $(RR - LL)/(RR + LL) \approx -0.15$ [121]. The TAUOLA package can properly account for the polarisation effects if the properties of the initial partons entering the collision are known, but this is not the case for the embedded events. On the other hand, the proper modelling of the τ polarisation is important since it affects kinematic properties of the event and, hence, the event selection efficiency and invariant $m_{\tau\tau}$ mass distributions, as illustrated in Fig. 4.12. In LL events the fraction of τ momentum carried by the visible decay products is on average smaller than in RR events and this has a sizeable impact on the invariant mass of the visible τ decay products (visible mass). Since LL events are produced more frequently than RR events, the visible mass in events with properly simulated tau polarisation effects will on average be lower than if no polarisation effects are considered. If case of the invariant $m_{\tau\tau}^{\mathrm{MMC}}$ mass, which accounts also for the neutrino momenta, the impact of the polarisation effects is reduced but still present, as shown in Fig. 4.13. This source of bias can be corrected by the TAUSPINNER [122] algorithm which
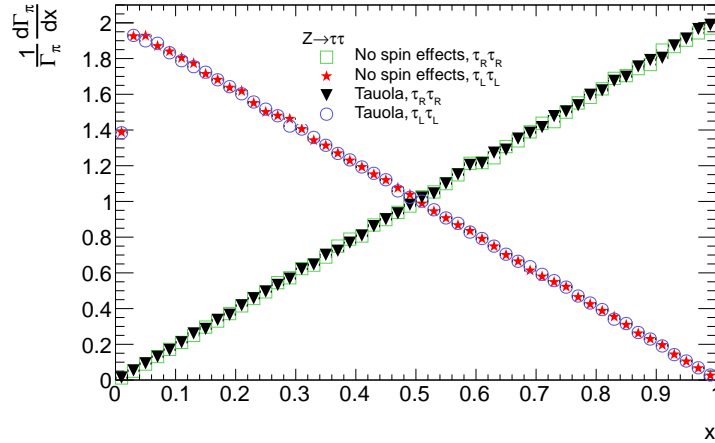
(a) fraction of τ momentum carried by the hadron (π, K) for right-handed (τ_R) and left-handed (τ_L) τ leptons
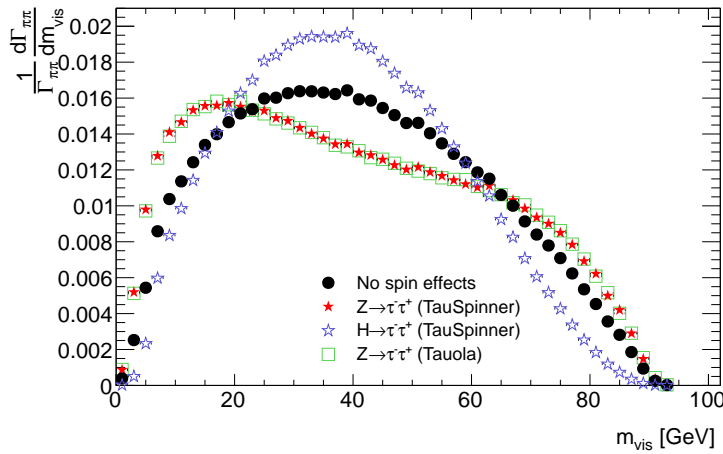
(b) invariant mass of the two hadrons from the fully hadronic $Z \to \tau\tau$ decay

Figure 4.12: τ spin correlation and polarisation effects in simulated $Z \to \tau\tau$ events with $\tau \to \pi\nu, K\nu$ decays [122]. "No spin effects" (black points) denotes the $Z \to \tau\tau$ events without spin correlation, i.e. with all configurations of τ polarisations (LL, LR, RL and RR), and without polarisation effects, i.e. with all configurations having equal probability. "Tauola" (open squares) indicates the $Z \to \tau\tau$ events simulated with the TAUOLA package, where spin correlations and τ polarisations are properly taken into account. "TauSpinner" (filled stars) indicates events with proper spin correlations and tau polarisations obtained by applying TAUSPINNER weights to the events with no spin and polarisation effects. The comparison with TAUOLA prediction is used as a validation of the TauSpinner algorithm. As a comparison (open stars), the TAUSPINNER package is also applied on the $Z \to \tau\tau$ events without spin effects to obtain the visible mass distribution expected for SM $H \to \tau\tau$ events with $m_H = m_Z$, in which LR and RL events are produced with equal probability. No detector simulation is included.
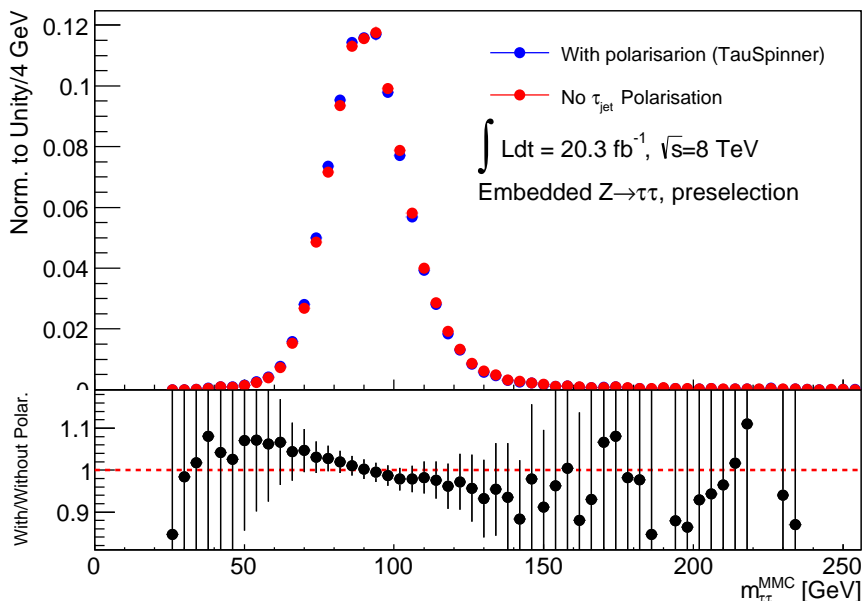
Figure 4.13: Comparison of the $m_{\tau\tau}^{\mathrm{MMC}}$ distributions of preselected embedded $Z \to \tau\tau$ events without (red) and with (blue) the polarisation effects as estimated by TauSpinner.

computes event weights for a given polarisation configuration starting from events with no polarisation effects based on a stochastically selected initial state parton configuration determined from the Z boson kinematic properties and the PDF set. In Fig. 4.12b the visible mass distribution obtained with the TauSpinner algorithm is compared to the TAUOLA prediction, which uses the exact polarisation configuration in each $Z \to \tau\tau$ event. The agreement between the two simulations validates the TauSpinner algorithm, which can therefore be used to model the polarisation effects missing in the embedded events.

4.6.2   Modelling of the Multi-jet Background Process

Multi-jet processes are the reducible source of background events in which both reconstructed τ-jets are mis-identified QCD jets from a non-resonant production of quarks and gluons. The contribution of this background process can only be predicted from data due to very large cross sections as well as the relatively poor modelling of the detector performance for QCD jets mis-identified as τ-jets.

Two control data samples enriched with multi-jet events are used: the so-called "not-oppositely-signed" (nOS) sample is used for the nominal multi-jet background prediction and the alternative so-called "same-sign" (SS) sample is used for the estimation of systematic uncertainties. These two samples are defined by selection criteria orthogonal to the nominal selection as summarised in Table 4.9. While the kinematic requirements defining the event categories remain the same, modified criteria on the τ-jet and $\tau\tau$ pair selection are introduced. In the nOS sample the τ-jet selection is loosened such that recon-

structed τ-jets with any charge and at least one associated track are accepted. All other requirements on τ-jets listed in Table 4.4 are still applied. The $\tau\tau$ pair selection imposes an inverted requirement on the charge product selecting events with a not oppositely charged τ-jet pair, i.e. $q_1 q_2 \neq -1$.

Table 4.9: Event selection criteria defining the nOS and SS control data samples enriched with multi-jet events. Only selection criteria differing from the nominal selection are indicated.

Selection	Nominal	nOS	SS
τ-jet	$\lvert q\rvert = 1$ ($N_{\mathrm{tracks}}= 1$ or 3) $N_{\mathrm{tracks},\Delta R<0.6}= 1$ or 3	Any charge $N_{\mathrm{tracks}} \geq 1$	$\lvert q\rvert = 1$ ($N_{\mathrm{tracks}}= 1$ or 3) $N_{\mathrm{tracks},\Delta R<0.6}= 1$ or 3
$\tau\tau$ pair	$q_1 q_2 = -1$	$q_1 q_2 \neq -1$	$q_1 q_2 = 1$

The nOS data sample is enriched with events containing two QCD jets mis-identified as τ-jets. In order to avoid double-counting, this control data sample is used for the modelling of all background processes producing two mis-identified τ-jets in the final state. In the nominal event selection simulated events in which both τ-jets are mis-identified QCD jets are rejected, while in the nOS sample the estimated contamination from processes containing at least one real hadronically decaying τ lepton, like $Z \rightarrow \tau\tau$, $W \rightarrow \tau\nu$ and $t\bar{t}$, are subtracted using predictions from simulation (see Table 4.10). The signal contamination in the nOS sample is expected to be negligible.

In order to evaluate the systematic uncertainty of the multi-jet background prediction based on the nOS data, the alternative SS control data sample is introduced, containing only events in which both τ-jets have the same electric charge, i.e. $q_1 q_2 = +1$, as listed in Table 4.9. The SS sample constitutes about 10-20% of the nOS sample, the exact amount being dependent on the event category. Although correlated, the nOS and the SS data samples are expected to contain significantly different fractions of quark-initiated and gluon-initiated jets. The gluon-initiated jets are more likely to be reconstructed with zero charge, thus entering the nOS but not the SS sample. Because of that, the difference between the prediction of the multi-jet background contribution with the nOS and SS data sample provides an estimate of the corresponding systematic uncertainty.

Fig. 4.14 compares the $m_{\tau\tau}^{\mathrm{MMC}}$ distributions obtained with the nOS and SS data samples in the Boosted Low-$p_{\mathrm{T}}^{\tau\tau}$ event category, in which the multi-jet contribution is significant and its prediction is not dominated by statistical uncertainty. Contributions from contaminating background processes with at least one hadronically decaying τ lepton are subtracted (see Table 4.10). Distributions are compatible with each other, but a slight shift to lower $m_{\tau\tau}^{\mathrm{MMC}}$ values is observed for the nOS sample and this is taken into account as systematic uncertainty on the modelling of the multi-jet process in this event category. Differences in other event categories are not statistically significant because the measured $m_{\tau\tau}^{\mathrm{MMC}}$ distributions are affected by large statistical uncertainties.

Table 4.10 indicates the observed numbers of events in the nOS and SS control data samples in $\sqrt{s} = 8$ TeV data together with the predicted contributions of signal and contaminating background processes with at least one hadronically decaying τ lepton.
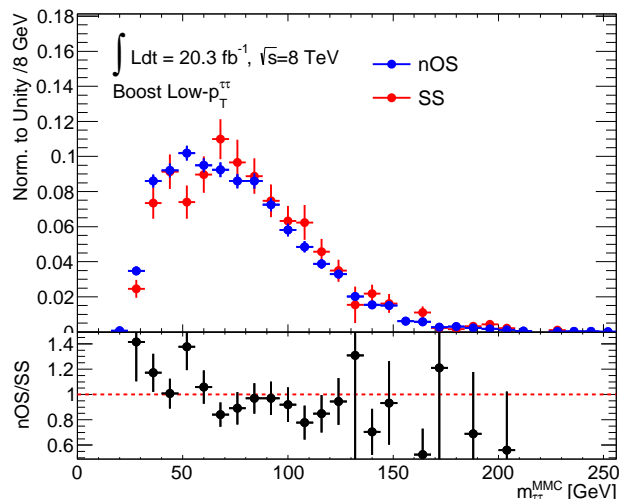
Figure 4.14: $m_{\tau\tau}^{\mathrm{MMC}}$ distributions obtained with the nOS and SS control data samples in the Boosted Low-$p_{\mathrm{T}}^{\tau\tau}$ event category in $\sqrt{s} = 8$ TeV data. Contributions from background processes with at least one hadronically decaying τ lepton are subtracted.

The latter two are estimated from simulation and subtracted from the observed data in order to derive the multi-jet contribution. The subtracted contaminating background constitutes a significant fraction of the total control data in event categories with high $p_{\mathrm{T}}^{\tau\tau}$. However, the possible corresponding systematic uncertainty will not affect the final result in these event categories since the multi-jet contribution relative to the total background is very small (Table 4.8). In the remaining event categories the total fraction of subtracted contaminating backgrounds is smaller than 10%. It is worth noticing that in the SS sample there is typically a smaller contamination with $Z \to \tau\tau$ events compared to the nOS sample, while the contamination from other sources of background processes, like $t\bar{t}$ or W+jets, is higher. This can be explained by the selection of the different classes of events in the two samples as mentioned previously. The signal contamination is negligibly small compared to the statistical uncertainty of the prediction of the multi-jet contribution and can, therefore, be neglected. The relative multi-jet contributions to the $m_{\tau\tau}^{\mathrm{MMC}}$ spectrum in each event category are measured with the nOS and SS control data samples. The obtained differences are taken into account as systematic uncertainties of the prediction of the multi-jet background contribution (see Section 4.7.1).

4.6.3   Validation of the Background Modelling

The modelling of the $Z \to \tau\tau$ and multi-jet background contributions is validated using a sample of events at the preselection stage which is characterised by a signal-like topology and at the same time by a negligible amount of signal events. Two different validation regions (VR) are defined in order to validate the background modelling separately for the VBF and Boosted event categories. The Boosted VR includes the preselected events with at least one jet with $p_{\mathrm{T}} > 50$ GeV, while the VBF VR is a subset of the Boosted VR containing events with at least two jets with p_{T} above 50 GeV and 30 GeV.

Table 4.10: Observed and expected event yields in the nOS and SS control data samples in $\sqrt{s} = 8$ TeV data, shown separately for each event category. The expected yields are determined from simulation. The expected Higgs boson contribution is shown for $m_H = 125$ GeV.

	VBF						Boosted			
$\sqrt{s} = 8$ TeV	High-$p_\mathrm{T}^{\tau\tau}$		Low-$p_\mathrm{T}^{\tau\tau}$ Tight		Low-$p_\mathrm{T}^{\tau\tau}$ Loose		High-$p_\mathrm{T}^{\tau\tau}$		Low-$p_\mathrm{T}^{\tau\tau}$	
$\int \mathrm{Ldt} = 20.3\,\mathrm{fb}^{-1}$	nOS	SS	nOS	SS	nOS	SS	nOS	SS	nOS	SS
Data	27	4	212	25	375	52	407	41	6695	1012
$Z \to \tau\tau$	10	0.4	4.0	0.0	4.0	0.5	180	11	335	20
$W, t, t\bar{t}, VV'$	1.4	0.5	2.0	0.0	7.0	4.0	33	0.8	200	51
$H \to \tau\tau$	0.4	0.02	0.4	0.2	0.1	0.05	1.6	0.1	3.3	0.2

Before validating the background model, the $Z \to \tau\tau$ and multi-jet contributions are normalised by fitting the corresponding distributions of the $m_{\tau\tau}^{\mathrm{MMC}}$ invariant mass to the observed data (see Fig. 4.15a), similar to the normalisation procedure in the signal region (see Section 4.8). A good convergence of the fit is assured mostly by the events in the lower tail of the $m_{\tau\tau}^{\mathrm{MMC}}$ distribution which originate predominately from multi-jet processes. Scaled by the resulting normalisation factors, the background distributions obtained from control data are compared to the observed distributions of the most interesting discriminating variables.

Fig. 4.15 shows the expected and observed distributions of the discriminating variables in the Boosted VR. The expected signal contribution for a Higgs boson mass $m_H = 125$ GeV is overlaid assuming a cross section ten times larger than the one predicted by SM. In all plots, the shapes of the $Z \to \tau\tau$ and multi-jet distributions, measured by the previously described data-driven methods, are in good agreement with the observed distributions. In particular, the expected distribution of the total $p_\mathrm{T}^{\tau\tau}$, to which the $Z \to \tau\tau$ and multi-jet processes contribute with significantly different spectra (see overflow bin in Fig. 4.15c), is in agreement with the observed data. A similar agreement for all discriminating variables is seen also in the VBF VR. Detailed comparisons of observed and expected distributions for all remaining discriminating variables used for the event selection are shown in Appendix A and in Ref. [99].

The presented validation results demonstrate the reliability of the background modelling based on the data-driven techniques which can, therefore, be successfully applied to all event categories defining the signal regions.

4.7 Systematic Uncertainties

In order to evaluate the compatibility of the observed data with the predicted signal and background contributions, several sources of systematic uncertainties have to be taken into account. These are categorised into experimental (Section 4.7.1) and theoretical (Section 4.7.2) uncertainties. The experimental uncertainties are related to the simulation

(a) $m_{\tau\tau}^{\mathrm{MMC}}$ invariant mass

(b) $\Delta R_{\tau\tau}$

(c) $p_{\mathrm{T}}^{\tau\tau}$

(d) $m_{\tau\tau}^{\mathrm{vis}}$ visible invariant mass
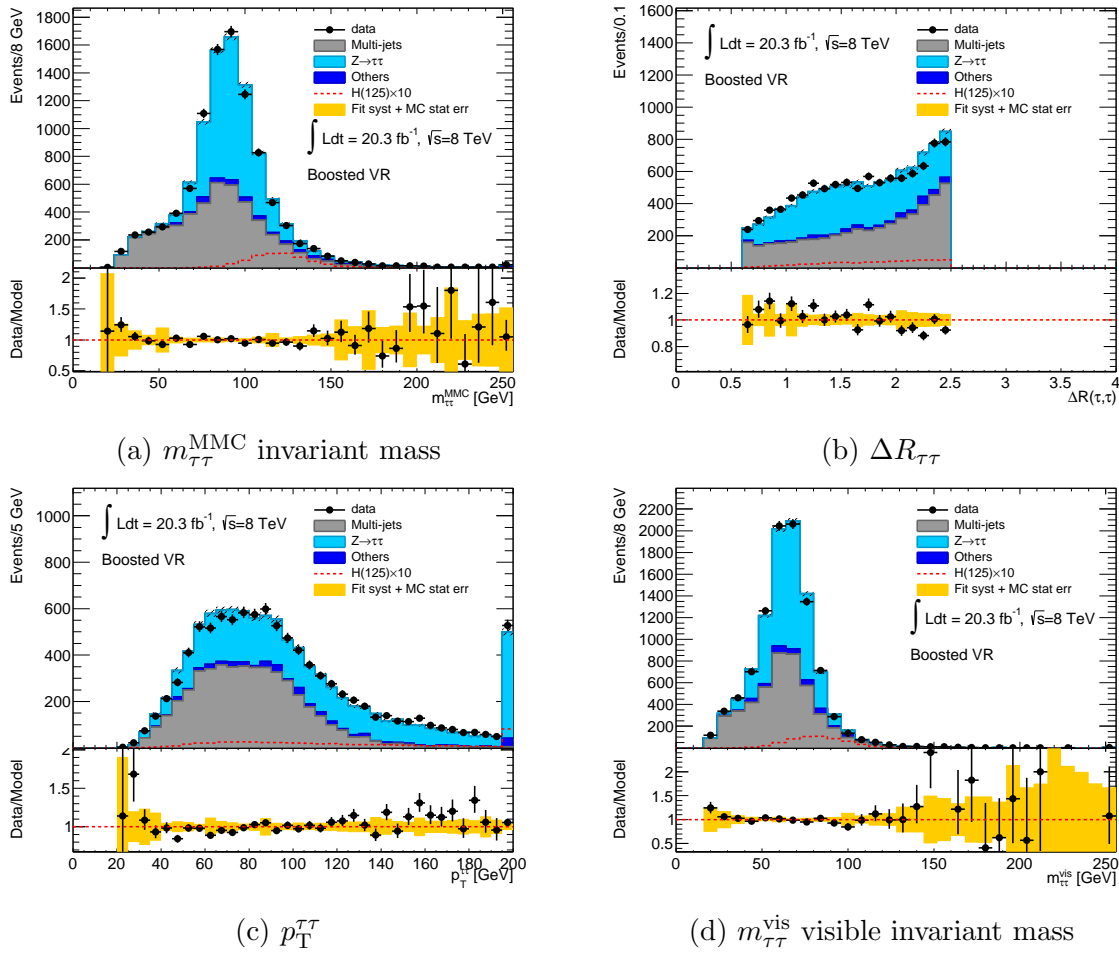
Figure 4.15: Observed and expected distributions of several discriminating variables in the Boosted VR in $\sqrt{s}$ = 8 TeV data. The hashed area on each plot indicates the statistical uncertainty of the background model. The lower panel shows the ratio of the observed data to the total expected background. The yellow band includes the statistical uncertainty of the background model, the multi-jet shape uncertainty and the multi-jet and $Z \to \tau\tau$ event yield uncertainties. The first and the last bin include the under- and over-flow events, respectively. $Z \to \tau\tau$ and multi-jet background distributions are obtained from the corresponding control data sample.

of the detector response and to the measurement of background contributions from control data, while the theoretical uncertainties are related to the cross section predictions and the event modelling with Monte-Carlo event generators.

The impact of different sources of systematic uncertainties is expressed in terms of relative changes of the expected event yields and in the versus of the varying shapes of the $m_{\tau\tau}^{\mathrm{MMC}}$ distributions in each event category. Each uncertainty is obtained by varying a given experimental or theoretical quantity by ± 1 standard deviation around the nominal value.

4.7.1   Experimental Uncertainties

The experimental uncertainties are grouped into four categories:

▷ uncertainty of the measurement of the integrated luminosity,
▷ uncertainties of the measurement of the τ-jet trigger and identification efficiencies and of the τ-jet energy calibration (TES),
▷ uncertainties of the jet energy calibration (JES) and resolution (JER) and
▷ uncertainties of the data driven background measurements.

Whenever the impact of varying the energy scale of a physics object in each event is evaluated, also the corresponding variation of the $E_{\mathrm{T}}^{\mathrm{miss}}$ value is considered. In this way, only the uncertainty of the soft $E_{\mathrm{T}}^{\mathrm{miss}}$ term which accounts for energy deposits not associated to any physics object is not taken into account yet. However, this uncertainty is found to be negligible and is neglected.

Integrated Luminosity

The uncertainty of the measurement of the integrated luminosity is $\pm 1.8\%$ in the $\sqrt{s} = 7$ TeV dataset and $\pm 2.8\%$ in the 8 TeV dataset, as detailed in Ref. [83].

τ-jet Trigger, Reconstruction and Identification

The uncertainties of the τ-jet trigger and identification efficiency for simulated hadronically decaying τ leptons are derived from measurements with $W \to \tau_{\mathrm{jet}} + 2\nu$ and $Z \to \tau_{\mathrm{jet}} l + 3\nu$ data (see Sections 3.2.1 and 3.3.1). Three independent components are evaluated separately: the systematic uncertainties of the measurements of the identification efficiency (τ-jet ID Eff) and of the trigger efficiency (τ-jet Trig Eff) and the quadratic sum of the statistical uncertainties of these two measurements (τ-jet Stat Eff). The impact of these uncertainties on the signal and background event yields are reported in Table 4.11. The impact of the statistical uncertainty on the background processes is not explicitly considered as it is added in quadrature to the statistical uncertainty of the expected background contribution. The impact of the above uncertainties on background can be smaller than on signal contributions since the former partially contains also events with just one hadronically decaying τ leptons.

The equivalent uncertainties for the simulated mis-identified jets are derived from measurements with $W \to \mu\nu$ and $Z \to \mu\mu$ data with additional jets in the final state (see

Table 4.11: Impact of the uncertainties of the τ-jet trigger and identification efficiencies on simulated signal and background processes with hadronically decay τ leptons at $\sqrt{s} = 8$ TeV. Relative changes in the event yield with respect to corresponding nominal values are shown averaged over all event categories.

τ-jet Eff [%]	$H \to \tau\tau$	$W, t, t\bar{t}, VV'$
ID	6	4-6
Trig	1	< 1
Stat	6	-

Section 3.4). Here, the τ-jet trigger and identification efficiencies are measured simultaneously and their uncertainty of about 20% is dominated by the statistical component. This uncertainty is also included as part of the statistical uncertainty of the expected background contribution.

These τ-jet uncertainties are not relevant for the prediction of the $Z \to \tau\tau$ background, even though τ-jets in the embedded $Z \to \tau\tau$ sample are simulated, since the normalisation of this background process is determined directly from the fit to the observed $m_{\tau\tau}^{\mathrm{MMC}}$ distribution.

The systematic uncertainty of the τ-jet energy calibration is composed of several independent components affecting different ranges in τ-jet transverse momentum as described in Section 3.1.1.

▷ *in situ*: uncertainty based on the TES measurement with $Z \to \tau\tau$ calibration data and assigned to τ-jets with $p_{\mathrm{T}} < 70$ GeV.
▷ Single Particle Response: TES uncertainty obtained from single particle response measurements and assigned to τ-jets with $p_{\mathrm{T}} > 70$ GeV.
▷ Modelling: TES uncertainty for the modelling of τ-jet properties with simulation, including uncertainties due to pileup contributions, underlying event description and detector geometry and response. This uncertainty component does not strongly depend on the τ-jet transverse momentum and is applied to all τ-jets.

The TES uncertainties affect both the predicted event yields and the shape of the $m_{\tau\tau}^{\mathrm{MMC}}$ distributions in each event category. Table 4.12 summarises the resulting uncertainties of the signal and $Z \to \tau\tau$ event yields in different event categories. The three components have similar impacts to the event yields with relative variations between 1-2%. The uncertainties of the shape of the $m_{\tau\tau}^{\mathrm{MMC}}$ distribution for the signal and $Z \to \tau\tau$ processes are shown in Fig. 4.16 for the Boosted High-$p_{\mathrm{T}}^{\tau\tau}$ event category in the $\sqrt{s} = 8$ TeV dataset. The full set of distributions for all event categories is given in Appendix B. In the VBF High-$p_{\mathrm{T}}^{\tau\tau}$ event category, where the simulated signal and $Z \to \tau\tau$ $m_{\tau\tau}^{\mathrm{MMC}}$ distributions are dominated by statistical uncertainties, the TES shape uncertainties are neglected. The impact of the TES systematic uncertainties on the remaining background processes with hadronically decaying τ leptons is smaller and the impact on the expected $m_{\tau\tau}^{\mathrm{MMC}}$ distributions is negligible compared to the large statistical uncertainty. Thus, only event

Table 4.12: Impact of the different components of the TES systematic uncertainties on the expected signal and $Z \to \tau\tau$ event yields in each event category in the $\sqrt{s} = 8$ TeV dataset. Relative changes in the event yield with respect to corresponding nominal values are shown.

$\sqrt{s} = 8$ TeV	VBF [%]			Boosted [%]	
TES [%]	High-$p_{\mathrm{T}}^{\tau\tau}$	Low-$p_{\mathrm{T}}^{\tau\tau}$ Tight	Low-$p_{\mathrm{T}}^{\tau\tau}$ Loose	High-$p_{\mathrm{T}}^{\tau\tau}$	Low-$p_{\mathrm{T}}^{\tau\tau}$
	$H \to \tau\tau$				
in situ	< 1	1	1	1	1
Single Particle Resp.	2	1	2	2	2
Modelling	1	< 1	< 1	1	< 1
	$Z \to \tau\tau$				
in situ	1	3	1	1	< 1
Single Particle Resp.	1	1	1	< 1	< 1
Modelling	1	2	1	< 1	< 1

yield uncertainties are taken into account in this case.

The same TES calibration and corresponding systematic uncertainties are assumed to be valid for both hadronically decaying τ leptons and mis-identified jets, even though no dedicated calibration measurement with data containing fake τ-jets is available. Given that the energy calibration of true and fake τ-jets can in general be different due to different jet composition, the corresponding two energy scale uncertainties are assumed to be uncorrelated.

Jet Reconstruction

The calibration of the energy scale of simulated jets (JES) is affected by several different systematic uncertainties:

▷ uncertainty of the energy scale measurement with Z+jet and di-jets data (*in situ*),
▷ uncertainty of the relative calibration across different pseudo-rapidity regions of the detector (η inter-calibration),
▷ uncertainty due to the difference in the calorimeter response to quark-initiated and gluon-initiated jets as well as due to the quark-gluon composition (jet flavour) and
▷ uncertainty on the correction procedure applied to subtract the contribution of pileup interactions to the jet energy (pileup effects).

In addition to the systematic uncertainty on the jet energy scale, also the systematic uncertainty on the jet energy resolution (JER) obtained from Z+jet and di-jets data is taken into account. Table 4.13 summarises the impact of these systematic uncertainties on the expected signal yield at $\sqrt{s} = 8$ TeV separately for each event category. Most affected are the predictions of the signal yields in VBF event categories due to the requirement of two tagging jets in the final state. Even though no jets are explicitly required in Boosted event categories, the requirement of a high $p_{\mathrm{T}}^{\tau\tau}$ value indirectly selects events with a
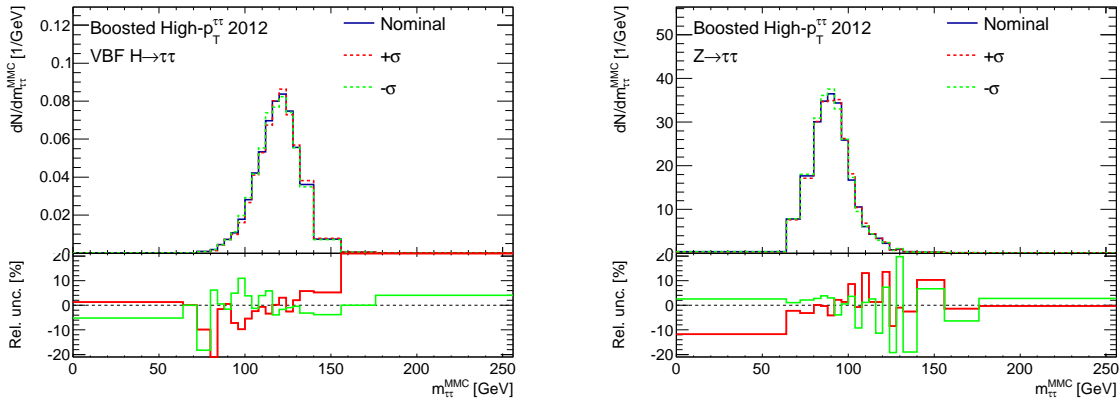
(a) *in situ* component
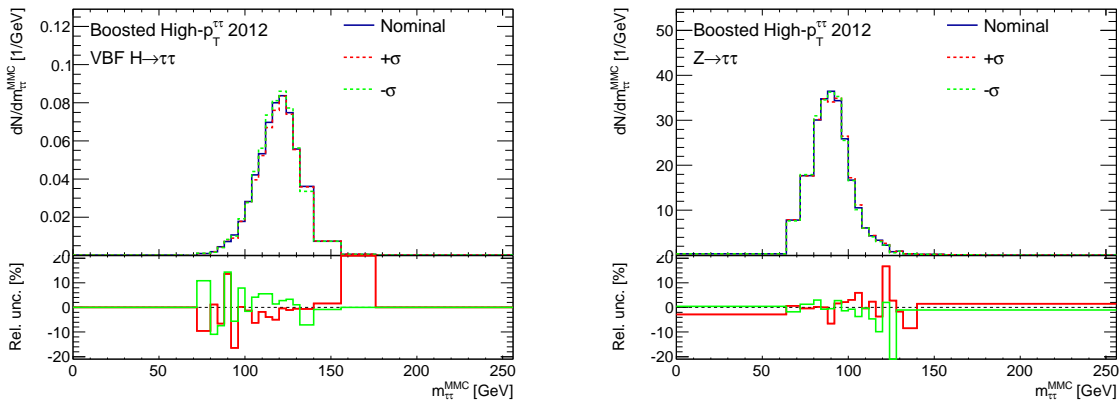
(b) single particle response component
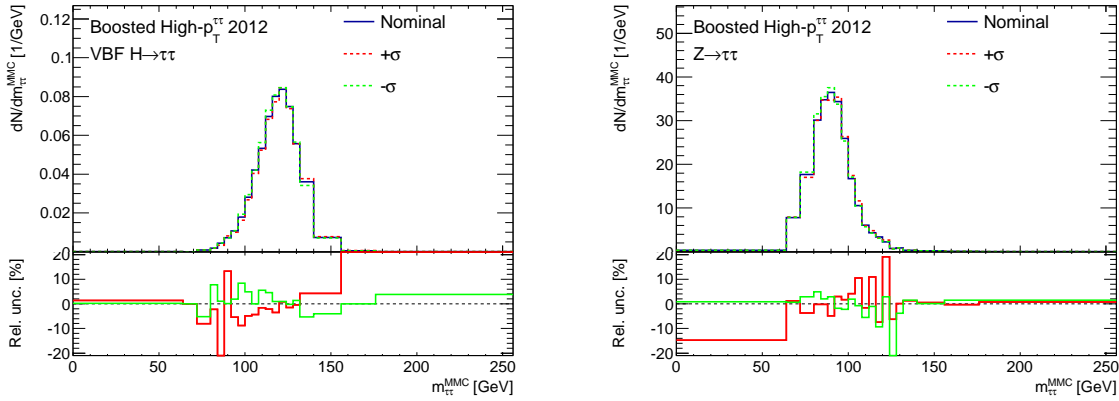
(c) modelling component

Figure 4.16: Impact of different components of the TES systematic uncertainty on the $m_{\tau\tau}^{\mathrm{MMC}}$ distributions in the Boosted High-$p_{\mathrm{T}}^{\tau\tau}$ event category for VBF Higgs boson (left) and embedded $Z \to \tau\tau$ (right) process in simulated $\sqrt{s} = 8$ TeV data. The red (green) dashed line shows the $m_{\tau\tau}^{\mathrm{MMC}}$ distribution resulting from the upward (downward) variation of the TES by one standard deviation for a given component of the systematic uncertainty. The relative changes with respect to the nominal TES are shown in the lower panel.

hard jet whose energy scale uncertainty can affect the E_T^{miss} calculation. $Z \rightarrow \tau\tau$ and multi-jet background contributions are not affected by the systematic uncertainties of the jet reconstruction as the reconstructed jets in these events are taken from data. Other remaining background processes are affected, but the impact is negligible compared to statistical uncertainties and to systematic uncertainties on the simulated τ-jets. Similar uncertainties are estimated for the simulated samples at $\sqrt{s} = 7$ TeV.

Table 4.13: Relative change of the nominal signal event yield in each event category due to jet systematic uncertainties at $\sqrt{s} = 8$ TeV. Same values are obtained for both Boosted Low-$p_T^{\tau\tau}$ and High-$p_T^{\tau\tau}$ categories.

$\sqrt{s} = 8$ TeV	VBF			Boosted
[%]	High-$p_T^{\tau\tau}$	Low-$p_T^{\tau\tau}$ Tight	Low-$p_T^{\tau\tau}$ Loose	
JES				
in situ measurement	3	5	4	3
η inter-calibration	5	10	5	3
Jet Flavour	3	5	5	5
Pileup Effects	1	2	2	1
JER	1	2	1	< 1

In addition to event yields, the jet energy scale uncertainties also affect the shape of the $m_{\tau\tau}^{\mathrm{MMC}}$ distributions, as illustrated in Fig. 4.17. Other related distributions are given in Appendix B.

Data-driven Background Measurement

Systematic uncertainties of the $Z \rightarrow \tau\tau$ embedding procedure described in the previous section account for the uncertainties of the $Z \rightarrow \mu\mu$ event selection efficiency and of the replacement of the selected muons and their associated calorimeter energy deposits by the simulated τ leptons. The former uncertainty is estimated by tightening or removing the muon isolation criterion, while the latter uncertainty is obtained by varying the expected deposited muon energy in the calorimeter by 30%. These uncertainties affect both the $Z \rightarrow \tau\tau$ event yields and the shape of the $m_{\tau\tau}^{\mathrm{MMC}}$ distribution in each event category, as shown in Table 4.14 and Fig. 4.18 for the Boosted High-$p_T^{\tau\tau}$ event category. Impacts on the $m_{\tau\tau}^{\mathrm{MMC}}$ distribution in other categories can be seen in Appendix B.

The systematic uncertainty on the modelling of the multi-jet background is described in Section 4.6. It is based on the difference of event yields and of shapes of the $m_{\tau\tau}^{\mathrm{MMC}}$ distributions observed with the nOS and the SS control data samples.

Table 4.15 reports the observed differences of event yields in each event category relative to the total multi-jet event yield. Large uncertainties in the two High-$p_T^{\tau\tau}$ and VBF categories are due to very small expected number of multi-jet events and the small size of the corresponding control data samples.
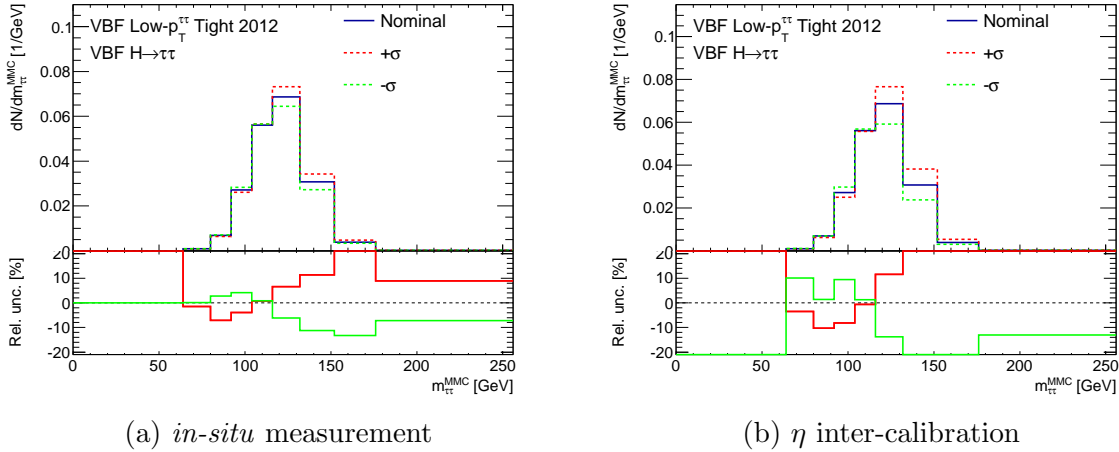
(a) *in-situ* measurement (b) η inter-calibration

Figure 4.17: Impact of different components of the JES systematic uncertainties on the shape of the $m_{\tau\tau}^{\mathrm{MMC}}$ distributions for the VBF Higgs boson signal at $\sqrt{s} = 8$ TeV in the VBF Low-$p_{\mathrm{T}}^{\tau\tau}$ Tight event category. The red (green) dashed line shows the $m_{\tau\tau}^{\mathrm{MMC}}$ distribution resulting from the upward (downward) variation of the TES by one standard deviation for a given component of the systematic uncertainty. Relative changes with respect to the nominal distribution are shown in the lower panel.

Table 4.14: Relative change of the nominal $Z \to \tau\tau$ event yield in each event category due to systematic uncertainties of the $Z \to \tau\tau$ embedding procedure for the $\sqrt{s} = 8$ TeV sample. Same values are obtained for both Boosted Low-$p_{\mathrm{T}}^{\tau\tau}$ and High-$p_{\mathrm{T}}^{\tau\tau}$ categories.

$\sqrt{s} = 8$ TeV [%]	VBF			Boosted
	High-$p_{\mathrm{T}}^{\tau\tau}$	Low-$p_{\mathrm{T}}^{\tau\tau}$ Tight	Low-$p_{\mathrm{T}}^{\tau\tau}$ Loose	
muon replacement	< 1	2	1	< 1
$Z \to \mu\mu$ selection efficiency	2	2	< 1	< 1

Table 4.15: Relative change of the multi-jet event yields in each event category due to the systematic uncertainty of the data-driven background measurement.

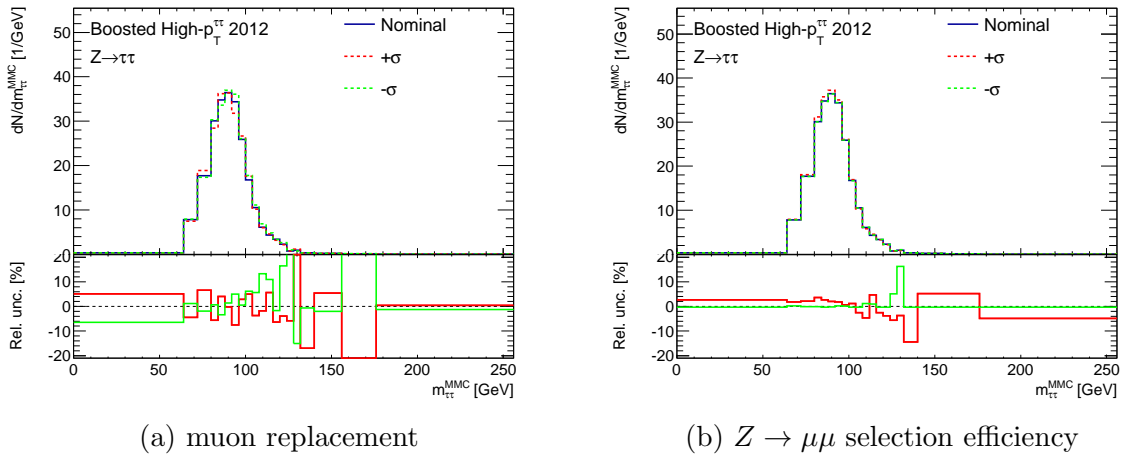| Category	$\delta(\sigma_{\mathrm{category}}/\sigma_{\mathrm{tot}})$ [%]
VBF High-$p_{\mathrm{T}}^{\tau\tau}$	∓ 50
VBF Low-$p_{\mathrm{T}}^{\tau\tau}$ Tight	± 35
VBF Low-$p_{\mathrm{T}}^{\tau\tau}$ Loose	± 25
Boosted High-$p_{\mathrm{T}}^{\tau\tau}$	∓ 50
Boosted Low-$p_{\mathrm{T}}^{\tau\tau}$	∓ 1

(a) muon replacement

(b) $Z \to \mu\mu$ selection efficiency

Figure 4.18: Impact of the systematic uncertainties of the $Z \to \tau\tau$ embedding procedure on the shape of the $Z \to \tau\tau$ $m_{\tau\tau}^{\mathrm{MMC}}$ distribution in the Boosted High-$p_{\mathrm{T}}^{\tau\tau}$ event category.

The systematic uncertainty of the shape of the $m_{\tau\tau}^{\mathrm{MMC}}$ distribution is taken into account only for the Boosted Low-$p_{\mathrm{T}}^{\tau\tau}$ category, for which the difference between the shapes obtained with the two control data samples is statistically significant (compare Fig. 4.14).

4.7.2 Theoretical Uncertainties

Theoretical systematic uncertainties affect the prediction of the cross sections and of the modelling of the simulated signal and background processes. These uncertainties can be divided into several components:

▷ uncertainties of the parton distribution functions (PDF) and of the value of the strong coupling constant (α_{S}),

▷ uncertainty of the branching ratio (BR) of the $H \to \tau\tau$ decay,

▷ uncertainty of the modelling of the underlying event (UE),

▷ uncertainty of the modelling of the Higgs boson transverse momentum $p_{\mathrm{T}}^{\mathrm{Higgs}}$ and

▷ uncertainties of the matrix elements' calculations (QCD scale).

Systematic uncertainties related to the QCD scale and the modelling of the Higgs boson transverse momentum $p_{\mathrm{T}}^{\mathrm{Higgs}}$ in signal events are the dominant theoretical uncertainties since the most sensitive categories contain events with a highly boosted Higgs boson and, respectively or, with two or more jets. In case of the ggF Higgs boson production these two event features can only be described when taking higher-order QCD corrections into account. Thus, the simulation of these events is affected by large uncertainties.

The theoretical systematic uncertainties are estimated following the recommendations of the LHC Higgs Combination Working Group [123] and of the LHC Higgs Cross Section Working Group [105]. The estimates of the uncertainties of the $p_{\mathrm{T}}^{\mathrm{Higgs}}$-modelling and of the QCD scale might still be conservative, as activities are currently ongoing in improving these calculations.

The uncertainties on the production cross sections due to PDF and α_S uncertainties are estimated to be 8% for the ggF Higgs boson production mode and 4% for the qq-initiated Higgs boson production modes, VBF and VH. The largest uncertainty arises from the uncertainty of the gluon PDF.

The uncertainty on the branching ratio of the $H \rightarrow \tau\tau$ decay depends on the Higgs boson mass and ranges from 6.8% at $m_H = 100$ GeV to 3% at $m_H = 150$ GeV [105].

The systematic uncertainty of the modelling of the underlaying event affects mostly the signal selection efficiency in VBF event categories. This is due to the multiple parton interactions (MPI) which are part of the underlaying event and tend to produce jets in the forward detector regions, mimicking the tagging jet properties. The size of these uncertainties is evaluated comparing the event yields in VBF categories for signal samples simulated with and without MPI [73, 124]. The estimated uncertainties amount to 10% for the expected ggF and 3% for the VBF signal yield. Corresponding uncertainties in the Boosted event categories are estimated to be negligible.

As described in Section 1.3, the modelling of the Higgs boson transverse momentum is sensitive to the masses of heavy quarks contributing to the loop in the $gg \rightarrow H$ production process. In particular, in the high-$p_T^{\tau\tau}$ range relevant for this search, the large-m_t approximation does not hold and the correct treatment of the top quark mass running in the ggF loop is required. The ggF events simulated with the POWHEG generator are produced at NLO with the full treatment of the heavy quark masses and with a tuning to match the NNLL+NNLO p_T^{Higgs} spectrum predicted by the HqT programme in the large-m_t approximation. The same NLO matrix elements with massive heavy quarks are implemented also in the MC@NLO generator, which is used to cross-check the POWHEG predictions. An observed significant difference of the signal yields obtained with the two generators indicates a different impact of the heavy quark masses. This difference, although not fully understood, is likely to come from the different matching schemes used in POWHEG and MC@NLO between the jets from the NLO matrix element and the parton shower. Since the predictions of the two generators are equally realistic, this difference is taken as systematic uncertainty on the modelling of the Higgs boson transverse momentum in ggF events. It amounts to 29% for the ggF yields in the VBF High-$p_T^{\tau\tau}$ and in both Boosted categories and to 18% for the ggF yields in the VBF Low-$p_T^{\tau\tau}$ category.

Systematic uncertainties related to missing higher-order terms in the matrix element calculations are estimated with the MCFM programme [125, 126]. VBF and ggF Higgs boson production modes in association with one and two jets are simulated at NLO. The uncertainty of the expected signal yield in each event category is estimated by simultaneously decreasing or increasing the renormalisation and factorisation scales by a factor of two relative to the nominal value m_H, i.e. $m_H/2$ or $2m_H$. For the VBF Higgs boson production, the estimated uncertainties amount to $\pm 2\%$ in all event categories. For the ggF Higgs boson production, the uncertainties in the VBF event categories are estimated based on the NLO calculation of the inclusive ggF $H + 2$jet production ($\Delta\sigma_{\geq 2j}$), while the uncertainties in the Boosted event categories also consider the NLO calculation of the

inclusive ggF $H + 1$jet process ($\Delta\sigma_{\geq 1j}$). The uncertainties of these two inclusive production processes are considered as uncorrelated, as recommended in Ref. [105]. Table 4.16 reports the size of the ggF cross section uncertainties in each event category.

Table 4.16: Impact of systematic uncertainties of the perturbative calculation on the ggF signal yields shown relatively to the corresponding nominal event yields.

[%]	\multicolumn{3}{c}{VBF}	\multicolumn{2}{c}{Boosted}
	High-$p_{\mathrm{T}}^{\tau\tau}$	Low-$p_{\mathrm{T}}^{\tau\tau}$ Tight	Low-$p_{\mathrm{T}}^{\tau\tau}$ Loose	High-$p_{\mathrm{T}}^{\tau\tau}$	Low-$p_{\mathrm{T}}^{\tau\tau}$
$\Delta\sigma_{\geq 2j}$	± 90	± 40	± 30	∓ 5	∓ 5
$\Delta\sigma_{\geq 1j}$				± 24	± 24

A relatively large uncertainty $\Delta\sigma_{\geq 2j}$ of 30-40% is obtained in the VBF Low-$p_{\mathrm{T}}^{\tau\tau}$ event categories since the second tagging jet in ggF events generated at NLO originates from the parton shower and not from the matrix element. In the VBF High-$p_{\mathrm{T}}^{\tau\tau}$ event category a conservative signal yield uncertainty of 90% is assigned since in this region the modelling of both the $p_{\mathrm{T}}^{\mathrm{Higgs}}$ and the second leading high-energy jet are relevant. The assigned uncertainty on the ggF production ensures that the accuracy in the prediction of the total signal yield in this event category is driven only by the VBF Higgs boson signal, which is targeted by this category. The selection criteria for Boosted event categories have a sizeable overlap with the selection of the VBF categories. Because of that, correlations between the expected ggF signal yields in the two classes of event categories are taken into account by applying the $\Delta\sigma_{\geq 2j}$ uncertainty also to the Boosted event categories, as indicated in Table 4.16.

The uncertainty on the VH sample is taken from Ref. [127].

All theoretical systematic uncertainties of the expected Higgs boson event yields in the different event categories are summarised in Table 4.17.

Table 4.17: Summary of all theoretical uncertainties on the expected SM Higgs boson event yields, shown only in the relevant event categories, separately for three Higgs boson production modes.

$\delta\sigma$ [%]	PDF + α_{S}	BR	UE	\multicolumn{2}{c}{p_{T}^{Higgs}}	\multicolumn{2}{c}{QCD Scale}
Category	All	All	VBF	VBF	Boosted	VBF	Boosted
VBF	4	6	3			2	2
ggF	8	6	10	29-18	29	90-30	24
VH	4	6					4

Theoretical uncertainties of the background cross sections affect only the simulated processes with minor contributions, such as $W + j$, $t\bar{t}$ and di-boson events, and are listed in Table 4.18 following the calculations in Ref. [127].

Table 4.18: Theoretical systematic cross section uncertainties for relevant background processes.

$\delta\sigma$ [%]	PDF + α_{S}	QCD scale
W	± 4	± 1
$t\bar{t}$	± 8	± 6
di-boson	± 4	± 5

4.8 The Statistical Model

The results of the search are derived by testing the level of agreement of the observed data with either the "background-plus-signal" or the "background-only" hypotheses. To perform such tests it is necessary to build a proper statistical model of the selected data which encodes all expected signal and background contributions, together with the data-driven background measurements and all systematic uncertainties. The statistical model depends on several input parameters. The parameter of interest (PoI) is the Higgs boson production cross section times $H \to \tau^+\tau^-$ decay branching ratio, expressed in units of the corresponding value predicted by the SM and referred to as the signal strength μ. In addition, the model includes a number of nuisance parameters (NP) related to statistical and systematic uncertainties and to the normalisation of background contributions measured in control data samples.

The statistical model is built with the `HistFactory` software package [128] in the `RooFit/RooStat` framework. It is determined by the probability density functions (pdf) of the binned $m_{\tau\tau}^{\mathrm{MMC}}$ distributions expected in each event category, as described in the following. The $m_{\tau\tau}^{\mathrm{MMC}}$ variable is chosen as it offers a good discrimination power between the signal and the background processes and also as it is sensitive to the Higgs boson mass.

The histograms of the expected $m_{\tau\tau}^{\mathrm{MMC}}$ distributions are composed of the individual contributions, here referred to as samples, from the VBF, ggF and VH signal process and the $Z \to \tau\tau$, multi-jet and other remaining background processes. The modelling of these contributions together with corresponding relative uncertainties is described in previous sections.

The total event yields and, thus, the normalisation of the $Z \to \tau\tau$ and multi-jet processes are obtained directly from the fit to the observed $m_{\tau\tau}^{\mathrm{MMC}}$ distributions. The relative yield in each event category is given by the respective models, the embedded $Z \to \tau\tau$ and the nOS control data sample, together with their associated systematic uncertainties. The event yields of the signal and of the remaining background processes are obtained from the calculated cross sections and the measured integrated luminosity relying on simulation for the relative yields in each event category.

The set $\boldsymbol{\alpha} = \{\mu, \alpha_p, \phi_p, \gamma_p\}$ of parameters of the statistical model includes the parameter of interest μ and the nuisance parameters. The latter are classified in parameters related to systematic uncertainties $\mathbb{S} = \{\alpha_p\}$, normalisation factors $\mathbb{N} = \{\phi_p\}$ determined directly from the fit and statistical uncertainties $\boldsymbol{\Gamma} = \{\gamma_{\mathbf{p}}\}$ of the expected number of

events in each histogram bin.

The statistical model is defined as

$$\mathcal{P}(n_{cb}, a_p | \mu, \phi_p, \alpha_p, \gamma_p) = \prod_{c\in\text{channels}} \prod_{b\in\text{bins}} \text{Pois}(n_{cb}|\nu_{cb}) \cdot \text{G}(L_0|\lambda, \Delta_L) \cdot \prod_{p\in\mathbb{S}} f_p(a_p|\alpha_p) \quad (4.3)$$

where c are the event categories, in the following referred to as channels, b are the bins of the binned $m_{\tau\tau}^{\text{MMC}}$ distribution in channel c and

▷ $\text{Pois}(n_{cb}|\nu_{cb})$ is the Poisson probability of observing n_{cb} events in the bin b of channel c given the expected number of events

$$\nu_{cb}(\mu, \phi_p, \alpha_p, \gamma_p) = \Sigma_{s\in\text{samples}}\lambda_{cs}\gamma_{cb}\phi_{cs}(\boldsymbol{\alpha})\eta_{cs}(\boldsymbol{\alpha})\sigma_{csb}(\boldsymbol{\alpha}), \quad (4.4)$$

where λ_{cs} is the true integrated luminosity needed to normalise the event yields in simulated samples, γ_{cb} is the statistical uncertainty assigned to the bin b, ϕ_{cs} is the normalisation parameter introduced if the normalisation of a given sample s is to be directly determined from the fit ($Z \to \tau\tau$ of multi-jet events), $\eta_{cs}(\boldsymbol{\alpha})$ is the uncertainty of the total event yield of sample s and $\sigma_{csb}(\boldsymbol{\alpha})$ is the content of bin b of the for the sample s obtained either from simulation or a control data sample and dependent on shape systematic uncertainties,

▷ $\text{G}(L_0|\lambda, \Delta_L)$ is the Gaussian probability for measuring integrated luminosity L_0, given the true integrated luminosity λ and the measurement uncertainty Δ_L and

▷ $f_p(a_p|\alpha_p)$ is the pdf that determines the constraint on the nuisance parameter α_p, based on auxiliary measurements or theoretical calculations a_p, as detailed next.

The choice of the pdf f_p for the nuisance parameters depends on the type of the corresponding associated uncertainty. Following the recommendations of the LHC Higgs Combination Group [123], three different classes of pdfs are used distinguishing between experimental systematic and theoretical systematic uncertainties as well as the statistical uncertainties.

The experimental systematic uncertainties (Section 4.7.1) are modelled with a Gaussian distribution

$$G(a_p|\alpha_p, \sigma_p) = \frac{1}{\sqrt{2\pi\sigma_p^2}} \exp\left[-\frac{(a_p - \alpha_p)^2}{2\sigma_p^2}\right], \quad (4.5)$$

where $a_p = 1$ corresponds to the nominal value of the parameter obtained from the auxiliary calibration measurement and applied in the nominal event reconstruction, α_p is the true value of the parameter relative to the nominal value of a_p and σ_p is the uncertainty of the auxiliary measurement. α_p can assume both positive or negative values.

Theoretical systematics uncertainties (Section 4.7.2), typically assuming only positive values like cross sections, are modelled with log-normal distributions

$$P_{LN}(a_p|\alpha_p, k_p) = \frac{1}{\sqrt{2\pi} \ln k_p} \frac{1}{a_p} \exp\left[-\frac{\ln(a_p/\alpha_p)^2}{2(\ln k_p)^2}\right], \quad (4.6)$$

which can be related to the Gaussian function of Eq. 4.5 by $k_p = \exp(\sigma_p)$. This pdf avoids the numerical instabilities that would occur with a Gaussian function truncated

at negative values. For small uncertainties σ_p, the Gaussian and the log-normal function are similar.

The statistical uncertainties are modelled by gamma distributions, i.e. a generalisation of the Poisson distribution allowing for non-integer input parameters. The corresponding Poisson distribution is given as

$$\text{Pois}(n_p|\tau_p\gamma_p) = \frac{(\tau_p\gamma_p)^{n_p}e^{-\tau_p\gamma_p}}{n_p!} \tag{4.7}$$

where the number of expected events $\tau_p\gamma_p$ is decomposed into the nominal event yield τ_p and the relative parameter γ_p in order to deal only with relative uncertainties.

Further details on the implementation of the statistical model are given in Appendix B.

The statistical description of the observed data is given by the maximum likelihood estimator (MLE), defined as the set of parameter values $\hat{\boldsymbol{\alpha}}$ that maximises the likelihood function

$$L(\boldsymbol{\alpha}) = \mathcal{P}(n_{cb}, a_p|\mu, \phi_p, \alpha_p, \gamma_p), \tag{4.8}$$

where n_{cb} is the observed number of events in the bin b of the channel c. The maximisation is performed with the `Minuit` programme [129] which minimises the value of $-\log L(\boldsymbol{\alpha})$.

The observed and expected $m_{\tau\tau}^{\text{MMC}}$ distributions resulting from the MLE, i.e. the best fit of the statistical model to the observed data are shown in Figs. 4.19 and 4.20, separately for each event category. Table 4.19 reports the observed and expected event yields in each category in the mass range $110 < m_{\tau\tau}^{\text{MMC}} < 140$ GeV, where significant signal contributions are expected. The fitted signal strength for $m_H = 125$ GeV is $\mu = 1.2^{+0.7}_{-0.6}$. The uncertainties of the fitted normalisation factors for the total $Z \to \tau\tau$ and multi-jet contributions in the $\sqrt{s} = 7(8)$ TeV data are 7(4)% and 6(4)%, respectively. The fitted normalisation factor for the $Z \to \tau\tau$ background contribution is in agreement with the prediction from the ALPGEN event generator as the ratio of the two is $0.84\pm0.07(\text{stat})\pm0.09(\text{syst})$ in the $\sqrt{s} = 8$ TeV dataset.

The $m_{\tau\tau}^{\text{MMC}}$ distribution in Fig. 4.21 is obtained by combining distributions from all event categories, each weighted by the corresponding $s/\sqrt{b}$ value where s and b are the expected numbers of signal and background events, respectively, computed in the mass range $110 < m_{\tau\tau}^{\text{MMC}} < 140$ GeV. In the lower panel the observed data is shown after the subtraction of the predicted background contribution and compared with the expected signal distribution for the a Higgs boson with mass $m_H = 125$ GeV and signal strength $\mu = 1.2$. A slight excess of data above the expected background can be seen and can be interpreted as a background fluctuation or the presence of Higgs boson signal. The result of the statistical interpretation is given in the next section.

A series of tests has been performed in order to assess the quality of the described fit, as detailed in Appendix B. Firstly, the stability of the fit is tested. The dependence of the likelihood function on the most important nuisance parameters is studied to probe for the presence of unexpected double-minima or non-parabolic behaviour. This test is sensitive to instabilities of the fit results and unexpected biases or correlations between nuisance parameters. As a second test, the fitted values of each nuisance parameter are compared to their input values, probing for statistically significant unexpected differences.
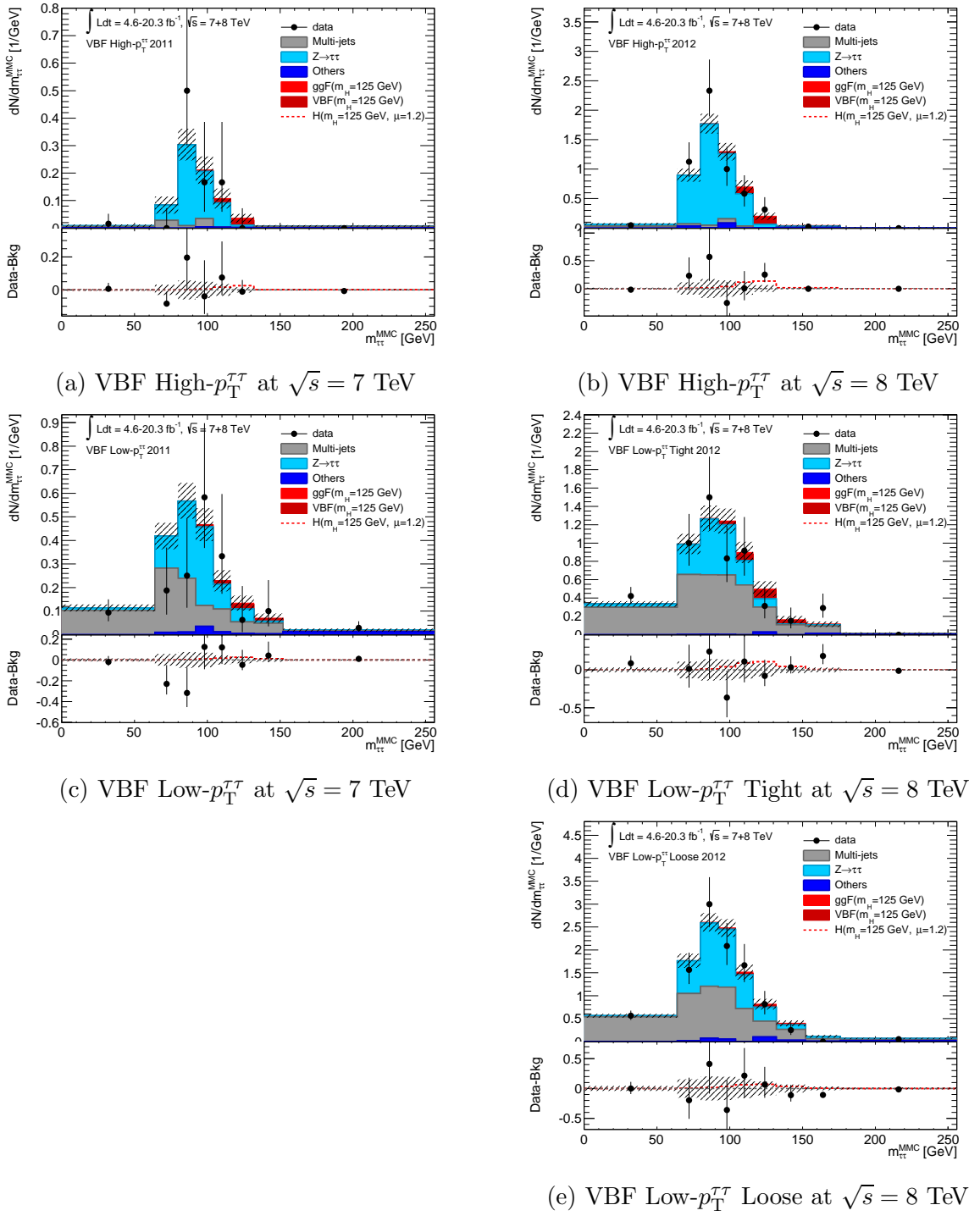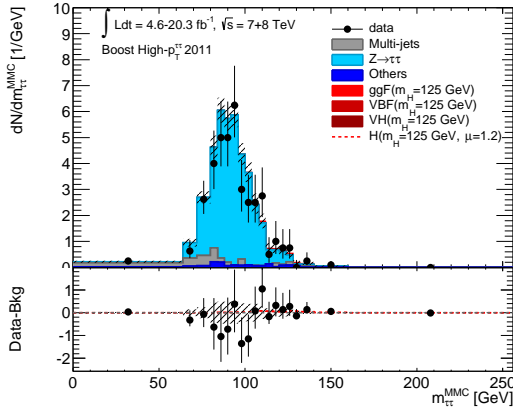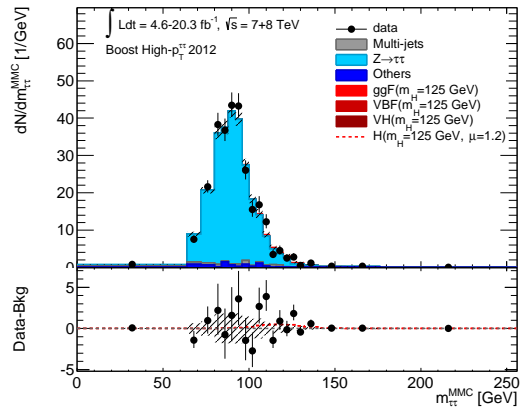
(a) VBF High-$p_{\mathrm{T}}^{\tau\tau}$ at $\sqrt{s} = 7$ TeV

(b) VBF High-$p_{\mathrm{T}}^{\tau\tau}$ at $\sqrt{s} = 8$ TeV

(c) VBF Low-$p_{\mathrm{T}}^{\tau\tau}$ at $\sqrt{s} = 7$ TeV

(d) VBF Low-$p_{\mathrm{T}}^{\tau\tau}$ Tight at $\sqrt{s} = 8$ TeV

(e) VBF Low-$p_{\mathrm{T}}^{\tau\tau}$ Loose at $\sqrt{s} = 8$ TeV

Figure 4.19: Observed and expected $m_{\tau\tau}^{\mathrm{MMC}}$ distributions in the VBF event categories. The lower panel compares the background-subtracted data with the expected signal distribution for the Higgs boson with $m_H = 125$ GeV and with a production cross section 1.2 times higher than the SM prediction. The dashed area represents the expected total background uncertainty.
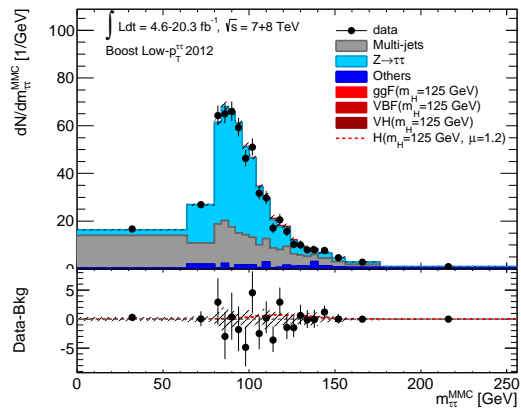
(a) Boosted High-$p_T^{\tau\tau}$ at $\sqrt{s} = 7$ TeV

(b) Boosted High-$p_T^{\tau\tau}$ at $\sqrt{s} = 8$ TeV
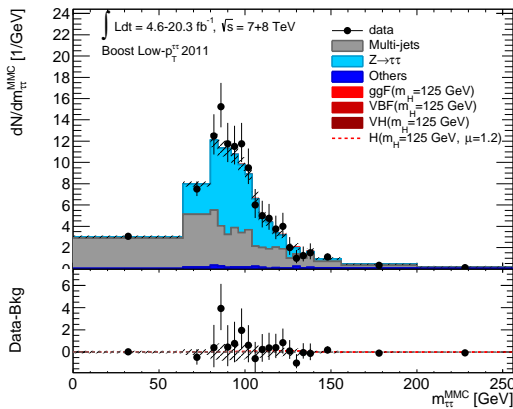
(c) Boosted Low-$p_T^{\tau\tau}$ at $\sqrt{s} = 7$ TeV

(d) Boosted Low-$p_T^{\tau\tau}$ at $\sqrt{s} = 8$ TeV

Figure 4.20: Observed and expected $m_{\tau\tau}^{\mathrm{MMC}}$ distributions in the Boosted event categories. The lower panel compares the background-subtracted data with the expected signal distribution for the Higgs boson with $m_H = 125$ GeV and with a production cross section 1.2 times higher than the SM prediction. The dashed area represents the expected total background uncertainty.

Table 4.19: Observed and predicted event yields in the mass range of $110 < m_{\tau\tau}^{\mathrm{MMC}} <$ 140 GeV in each event category in the $\sqrt{s} = 7$ TeV and 8 TeV datasets. The expected signal yields are shown for a SM Higgs boson with $m_H = 125$ GeV with the signal strength set to the measured value $\mu = 1.2$. The error on the total background yield includes all statistical and systematic uncertainties.

$\sqrt{s} = 7$ TeV, $\int$ Ldt $= 4.6\,\mathrm{fb}^{-1}$	VBF		Boosted	
	High-$p_{\mathrm{T}}^{\tau\tau}$	Low-$p_{\mathrm{T}}^{\tau\tau}$	High-$p_{\mathrm{T}}^{\tau\tau}$	Low-$p_{\mathrm{T}}^{\tau\tau}$
H($m_H = 125$ GeV)	0.8	0.8	1.9	2.9
$Z \to \tau\tau$	1.4	2.3	15	40
Multi-jet	0.5	2.9	1.6	46
$W, t, t\bar{t}, VV'$	0.2	0.3	1.9	3.6
Total Bkg.	2.1 ± 1.1	5.4 ± 1.4	18 ± 5	90 ± 11
Data	2	7	25	93

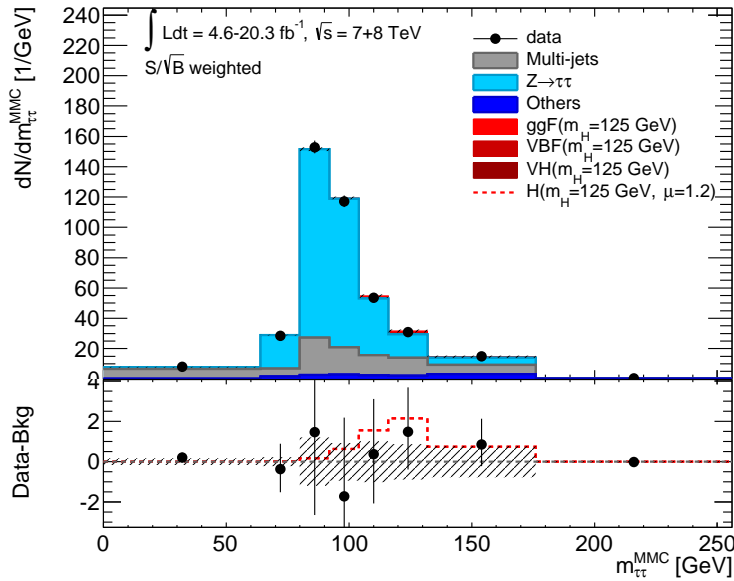| $\sqrt{s} = 8$ TeV, $\int$ Ldt $= 20.3\,\mathrm{fb}^{-1}$		Tight	Loose		
H($m_H = 125$ GeV)	4.2	3.7	2.5	12	19
$Z \to \tau\tau$	8.3	4.8	16	79	241
Multi-jet	0.4	13	19	8	201
$W, t, t\bar{t}, VV'$	0.5	0.5	2.2	2.9	47
Total Bkg.	9 ± 3	18 ± 4	37 ± 5	90 ± 14	490 ± 37
Data	13	19	38	113	477

Figure 4.21: Observed $s/\sqrt{b}$-weighted sum of $m_{\tau\tau}^{\mathrm{MMC}}$ distributions from all event categories. The lower panel compares the background-subtracted data with the expected signal distribution for the Higgs boson with $m_H = 125$ GeV and cross section 1.2 times higher than the SM prediction. The dashed area represents the expected total background uncertainty.

Moreover, the contribution of individual uncertainty components to the total uncertainty of the signal cross section is studied to probe for unexpectedly big uncertainty from less relevant uncertainties. In the third test, the fitted $m_{\tau\tau}^{\mathrm{MMC}}$ distributions are compared with observed data in the mass range outside the signal region. This test probes the convergence of the fit in the mass side-bands $0 < m_{\tau\tau}^{\mathrm{MMC}} < 110$ GeV and $m_{\tau\tau}^{\mathrm{MMC}} > 140$ GeV which are dominated by background contributions. All described tests do not single out any problem and confirm that the fit is stable and without unexpected features.

4.9 Results

The compatibility of the signal-plus-background and the background-only hypotheses with the observed data can be quantified by introducing a test statistics for the comparison of the two competing hypotheses. In general, the null hypothesis H_0 is assumed to be true in the observed data and the alternative hypothesis H_1 is tested with respect to H_0.

The test statistic recommended by the LHC Higgs Combination Group is the profile likelihood ratio [130], which provides almost as powerful hypothesis testing as the optimal test statistic with the likelihood ratio of the two hypotheses $T(n_{cb}) = \mathcal{P}(n_{cb}|H_1)/\mathcal{P}(n_{cb}|H_0)$. While $T(n_{cb})$ is parameter-independent, the profile likelihood ratio depends on the parameter of interest, whereas the nuisance parameters are absorbed (or "profiled") in the likelihood fit. In contrast to $T(n_{cb})$, the profile likelihood ratio allows for the analytic description of the test statistic distribution, which is crucial for the statistical interpretation

of large datasets.

The set of nuisance parameters $\boldsymbol{\theta} = \{\alpha_p, \phi_p, \gamma_p\}$ is profiled using two different likelihood estimators. The unconditional maximum likelihood estimator $L(\hat{\mu}, \hat{\theta})$ corresponds to the best fit result described in the previous section and the conditional maximum likelihood estimator $L(\mu, \hat{\hat{\theta}}(\mu))$ provides the maximum likelihood estimate of the nuisance parameters $\boldsymbol{\theta}$ for a given value of μ. The profile likelihood ratio is, thus, given as the ratio of the two estimators

$$\lambda(\mu) = \frac{L(\mu, \hat{\hat{\theta}}(\mu))}{L(\hat{\mu}, \hat{\theta})}, \tag{4.9}$$

independent of the nuisance parameters and depending only on the parameter of interest μ, the observed number of events n_{cb} and the auxiliary measurements. Even though the signal strength cannot assume negative values, one can allow for negative values of μ in the fit to account for downward background fluctuations. For that purpose, a modified profile likelihood ratio is used

$$\tilde{\lambda}(\mu) = \begin{cases} \frac{L(\mu, \hat{\hat{\theta}}(\mu))}{L(\hat{\mu}, \hat{\theta})} & \hat{\mu} \geq 0 \\[3mm] \frac{L(\mu, \hat{\hat{\theta}}(\mu))}{L(0, \hat{\hat{\theta}}(0))} & \hat{\mu} < 0 \end{cases}. \tag{4.10}$$

For the evaluation of the signal significance, the test statistic

$$\tilde{q}_0 = \begin{cases} -2 \ln \tilde{\lambda}(0) & \hat{\mu} > 0 \\[3mm] 0 & \hat{\mu} \leq 0 \end{cases} \tag{4.11}$$

probes the background-only null hypothesis ($\mu = 0$) with respect to the best fit (MLE) result. The compatibility of data with the background-only hypothesis is given by the probability

$$p_0 = \int_{\tilde{q}_{0,\text{obs}}}^{\infty} f(\tilde{q}_0 | 0, \hat{\hat{\theta}}(0)) d\tilde{q}_0, \tag{4.12}$$

where $\tilde{q}_{0,\text{obs}}$ is the observed value of the test statistic. This probability can be obtained from an ensemble of signal-plus-background pseudo-experiments with signal strength μ, although this can be computationally intensive. The advantage of the profile likelihood ratio compared to the direct likelihood ratio is that in the asymptotic limit of a large number of events, the distribution of the test statistic $f(\tilde{q}_\mu | \mu, \hat{\hat{\theta}}(\mu))$ can be described by an analytic function, independently of the nuisance parameters [130]. This analytic function depends only on the variance of $\hat{\mu}$ which is estimated from the so-called Asimov dataset, an artificial representative dataset defined by the previously described statistical model and assuming that the MLE defines the true parameter values in data. The p_0-value can be correlated to a signal significance expressed in terms of Gaussian standard deviations, defined as $Z = \Phi^{-1}(1-p_0)$, where Φ^{-1} is the inverse of the cumulative Gaussian distribution. The standard used in particle physics to claim an observation of a signal is $Z = 3\,\sigma$, corresponding to $p_0 = 1.3 \cdot 10^{-3}$. A discovery is claimed for signal significance of $5\,\sigma$, corresponding to $p_0 = 2.9 \cdot 10^{-7}$. The expected significance is computed assuming

that the background-plus-signal hypothesis is true. The corresponding expected p_0-value is obtained as the median of the $f(p_0|1, \hat{\hat{\theta}}(1, \text{obs}))$ distribution with the NPs profiled at $\mu = 1$.

For the evaluation of upper limits on the signal cross section, the signal-plus-background null hypothesis is tested assuming the presence of a signal with a given signal strength μ. The corresponding test statistic[2] is given by

$$
\tilde{q}_\mu = \begin{cases} -2\ln\tilde{\lambda}(\mu) & \hat{\mu} \leq \mu \\ 0 & \hat{\mu} > \mu \end{cases} = \begin{cases} -2\ln\frac{L(\mu,\hat{\hat{\theta}}(\mu))}{L(0,\hat{\hat{\theta}}(0))} & \hat{\mu} < 0 \\ -2\ln\frac{L(\mu,\hat{\hat{\theta}}(\mu))}{L(\hat{\mu},\hat{\theta})} & 0 \leq \hat{\mu} \leq \mu \\ 0 & \hat{\mu} > \mu \end{cases}. \tag{4.13}
$$

If no signal is present in data, upper limits can be set on its cross section by calculating the probability that the $\tilde{q}_\mu$ test statistic for a given signal strength μ assumes a value equal or higher than the observed value $\tilde{q}_{\mu,\text{obs}}$,

$$
p_\mu = \int_{\tilde{q}_{\mu,\text{obs}}}^{\infty} f(\tilde{q}_\mu|\mu, \hat{\hat{\theta}}(\mu))d\tilde{q}_\mu. \tag{4.14}
$$

The upper limit μ_{up} at 95% Confidence-Level (CL) is defined as the highest value of the signal strength μ for which the probability $p_{\mu,\text{up}}$ is still higher or equal to 5%.

In case of downward fluctuations of the observed background yield, this upper limit can be arbitrarily small leading to unphysical exclusion of rather small signal cross sections to which the search is not sensitive. This problem can be avoided by the use of the CL_S upper limits [131, 132] defined as the ratio of the p-values

$$
CL_S = \frac{p_\mu}{1 - p_b}, \tag{4.15}
$$

where p_b quantifies the agreement of data with the background-only hypothesis

$$
p_b = 1 - \int_{\tilde{q}_{\mu,\text{obs}}}^{\infty} f(\tilde{q}_\mu|0, \hat{\hat{\theta}}(0))d\tilde{q}_\mu. \tag{4.16}
$$

A downward background fluctuation in data will lead to small values of $1 - p_b$ increasing the CL_S upper limit and avoiding the exclusion of too small cross sections. The expected upper limits are obtained assuming the background-only hypothesis based on the median of the $f(\mu_{\text{up}}|0, \hat{\hat{\theta}}(0))$ distribution with the nuisance parameters profiled at $\mu = 0$.

Table 4.20 reports the expected CL_S upper limits at 95% confidence level on the signal strength μ for the hypothesised Higgs boson signal with a mass of $m_H = 125$ GeV separately for each event category. The VBF High-$p_{\text{T}}^{\tau\tau}$ is the most sensitive event category, closely followed by the Boosted High-$p_{\text{T}}^{\tau\tau}$ category. Table 4.21 compares the expected and observed 95% CL_S upper limits for the same assumed Higgs boson mass. Fig. 4.22 shows

---

[2] The test statistic $\tilde{q}_\mu$ is defined in order to derive one-sided confidence intervals.

Table 4.20: Expected 95% CL_S upper limits on the cross section times the branching ratio for a Higgs boson with $m_H = 125$ GeV, normalised to the SM prediction. Results are shown separately for each category.

Event Category	$\int \mathrm{Ldt} = 4.6\,\mathrm{fb}^{-1}$ $\sqrt{s} = 7$ TeV	$\int \mathrm{Ldt} = 20.3\,\mathrm{fb}^{-1}$ $\sqrt{s} = 8$ TeV	Combined 7 & 8 TeV
VBF High-$p_\mathrm{T}^{\tau\tau}$	6.2	2.2	
VBF Low-$p_\mathrm{T}^{\tau\tau}$ Tight	8.3	4.1	
VBF Low-$p_\mathrm{T}^{\tau\tau}$ Loose	-	8.0	
VBF All	4.4	1.7	1.5
Boosted High-$p_\mathrm{T}^{\tau\tau}$	6.1	2.7	
Boosted Low-$p_\mathrm{T}^{\tau\tau}$	10	4.8	
Boosted All	5.4	2.4	2.1
All Categories	3.2	1.3	1.2

Table 4.21: Observed (expected) 95% CL_S upper limits on the cross section times the branching ratio for a Higgs boson with $m_H = 125$ GeV, normalised to the SM prediction.

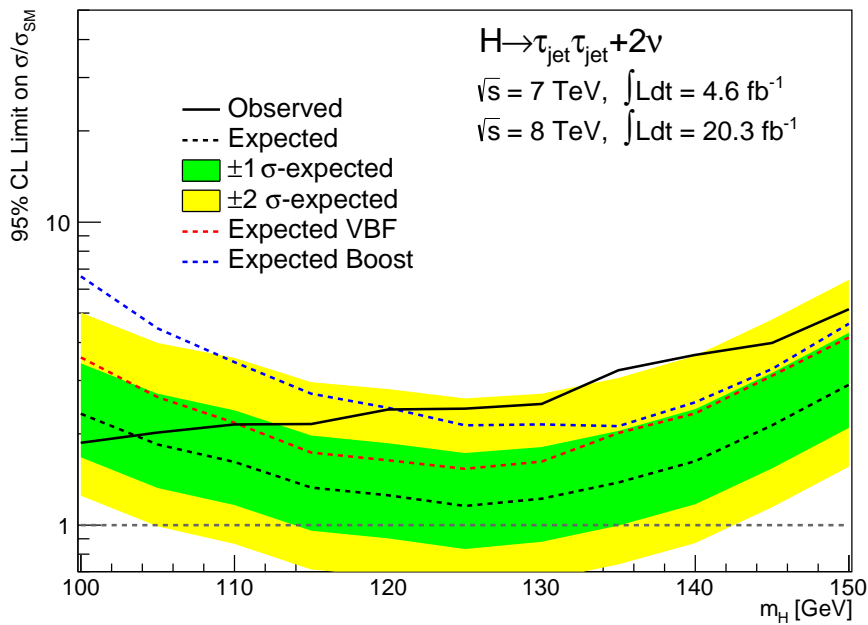| Event Category	$\int \mathrm{Ldt} = 4.6\mathrm{fb}^{-1}$ $\sqrt{s} = 7$ TeV	$\int \mathrm{Ldt} = 20.3\mathrm{fb}^{-1}$ $\sqrt{s} = 8$ TeV	Combined 7 & 8 TeV
VBF	5.4(4.4)	2.6(1.7)	2.4(1.5)
Boosted	9.0(5.4)	4.3(2.4)	4.3(2.1)
All Categories	4.7(3.2)	2.5(1.3)	2.4(1.2)

Figure 4.22: Observed and expected 95% CL_S upper limits on the Higgs boson cross section times the branching ratio normalised to the SM prediction as a function of the Higgs boson mass. The green (yellow) band indicates the $\pm 1(2)\sigma$ uncertainty on the expected upper limits. Expected upper exclusion limits obtained separately for the VBF and Boosted event categories are also shown for comparison.

the observed and expected 95% CL_S upper exclusion limits on the signal strength μ in dependence on the hypothesised Higgs boson mass. The limits are obtained by combining all event categories. The expected upper limits obtained separately for the VBF and Boosted categories are also shown for the comparison of the corresponding sensitivities. The green and yellow bands indicate the $\pm 1\sigma$ and $\pm 2\sigma$ uncertainty on the expected upper limits, respectively. The observed limits agree with the expected ones within approximately two standard deviations, although they are systematically higher than the expected ones for mass hypotheses above $m_H \gtrsim 115$ GeV.

The higher values of the observed upper limits compared to the expected can be related to an excess of the observed data above the expected background contribution (see Fig. 4.21). Assuming that the excess is due to the presence of the Higgs boson signal, the p_0 values and signal significances can be evaluated from Eq. 4.12. Table 4.22 reports the expected signal significance in each event category for $m_H = 125$ GeV. Fig. 4.23 shows the expected p_0-value and corresponding signal significance for different Higgs boson mass hypotheses after combining all event categories. The results obtained separately for the VBF and Boosted event categories are shown for comparison.   The highest analysis

Table 4.22: Expected signal significance for $m_H = 125$ GeV, shown separately for each event category.

Event Category	$\int \mathrm{Ldt} = 4.6\mathrm{fb}^{-1}$ $\sqrt{s} = 7$ TeV	$\int \mathrm{Ldt} = 20.3\mathrm{fb}^{-1}$ $\sqrt{s} = 8$ TeV	Combined 7 & 8 TeV
VBF High-$p_\mathrm{T}^{\tau\tau}$	0.6	1.3	
VBF Low-$p_\mathrm{T}^{\tau\tau}$ Tight	0.3	0.5	
VBF Low-$p_\mathrm{T}^{\tau\tau}$ Loose	-	0.3	
VBF All	0.7	1.5	1.6
Boosted High-$p_\mathrm{T}^{\tau\tau}$	0.5	0.9	
Boosted Low-$p_\mathrm{T}^{\tau\tau}$	0.2	0.5	
Boosted All	0.5	1.1	1.2
All Categories	0.8	1.9	2.1

sensitivity is expected for a hypothesised Higgs boson mass of about $m_H \simeq 125$ GeV. For smaller Higgs boson masses it is more difficult to discriminate the signal from the $Z \to \tau\tau$ background process, while for higher Higgs boson masses the expected signal yield is reduced by the decreasing $H \to \tau\tau$ branching ratio. The VBF High-$p_\mathrm{T}^{\tau\tau}$ event category provides the highest sensitivity, followed by the Boosted High-$p_\mathrm{T}^{\tau\tau}$ category.

Table 4.23 compares the expected and observed signal significances for $m_H = 125$ GeV. Fig. 4.24 shows the observed p_0-value and the corresponding signal significance as a function of the hypothesised Higgs boson mass. This is compared to the value assuming the presence of a SM Higgs boson with a mass of either 105, 125 or 145 GeV. The observed p_0-values agree best with the expected ones for $m_H = 125$ GeV, in agreement with the previous observation of a Higgs boson in the $\gamma\gamma$ and ZZ^* decays [71].
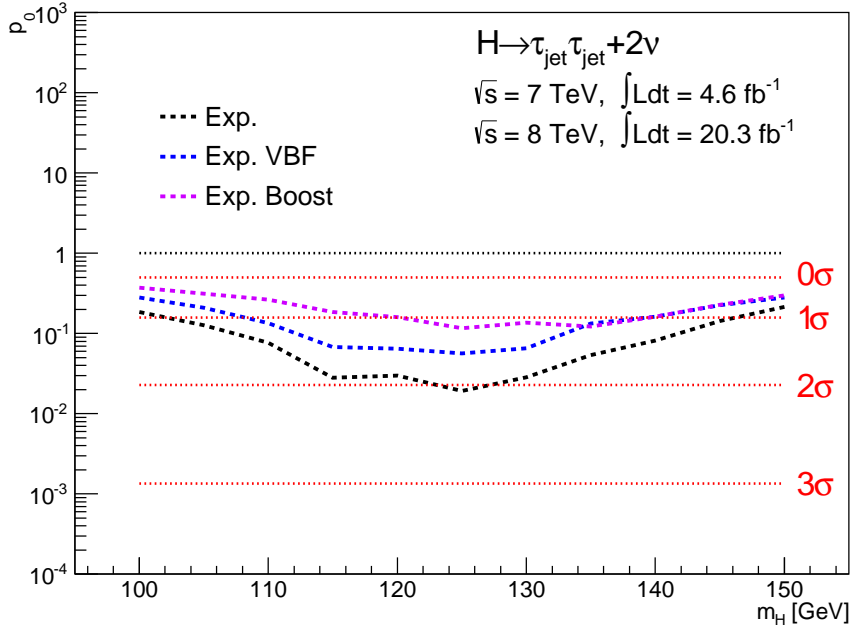
Figure 4.23: Expected p_0-value and the corresponding signal significance in dependence on the hypothesised value of m_H. Expected p_0-values obtained separately for the VBF and Boosted event categories are also shown for comparison.

Table 4.23: Observed (expected) signal significance for a hypothesised Higgs boson signal with a mass $m_H = 125$ GeV.

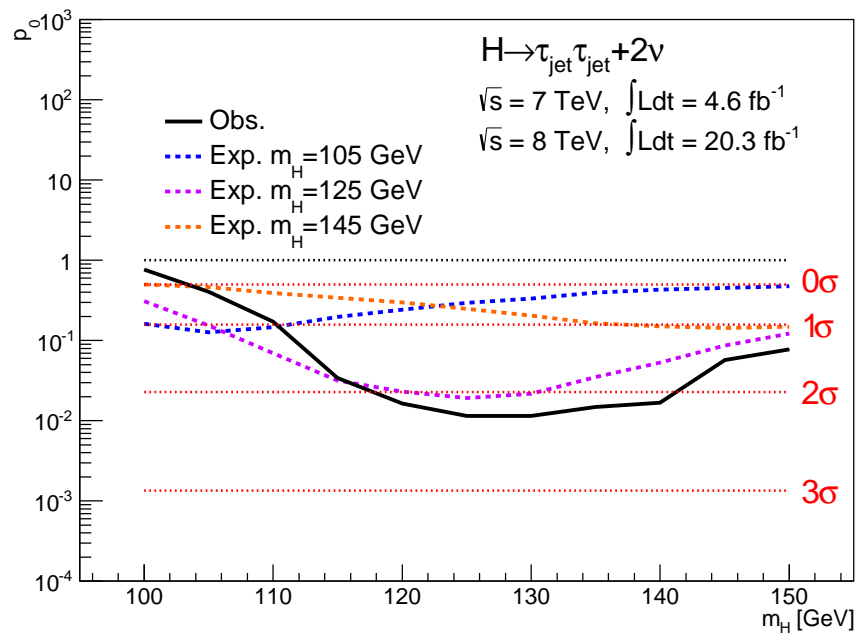| Event Category	$\int \text{L}dt = 4.6\text{fb}^{-1}$ $\sqrt{s} = 7$ TeV	$\int \text{L}dt = 20.3\text{fb}^{-1}$ $\sqrt{s} = 8$ TeV	Combined 7 & 8 TeV
VBF	-0.3(0.7)	1.6(1.5)	1.5(1.6)
Boosted	1.0(0.5)	1.6(1.1)	1.9(1.2)
All	0.3(0.8)	2.3(1.9)	2.3(2.1)

Figure 4.24: Observed p_0-value as a function of the hypothesised Higgs boson mass. Expected p_0-values obtained by assuming the existence of a SM Higgs boson with a given mass are also shown.

Although the observed signal significance of $2.3\,\sigma$ is not sufficient to claim evidence for the signal in the fully hadronic $H \to \tau^+\tau^-$ channel, the presence of the $H \to \tau^+\tau^-$ signal is a realistic interpretation strongly supported by the result of the combination of all three $H \to \tau^+\tau^-$ decay modes reported in Ref. [67] and outlined in the next section.

The size of the observed excess of events is measured in units of the signal strength μ relative to the predicted SM Higgs boson cross section times the branching ratio in dependence on the hypothesised mass m_H as shown in Fig. 4.25. The signal strength
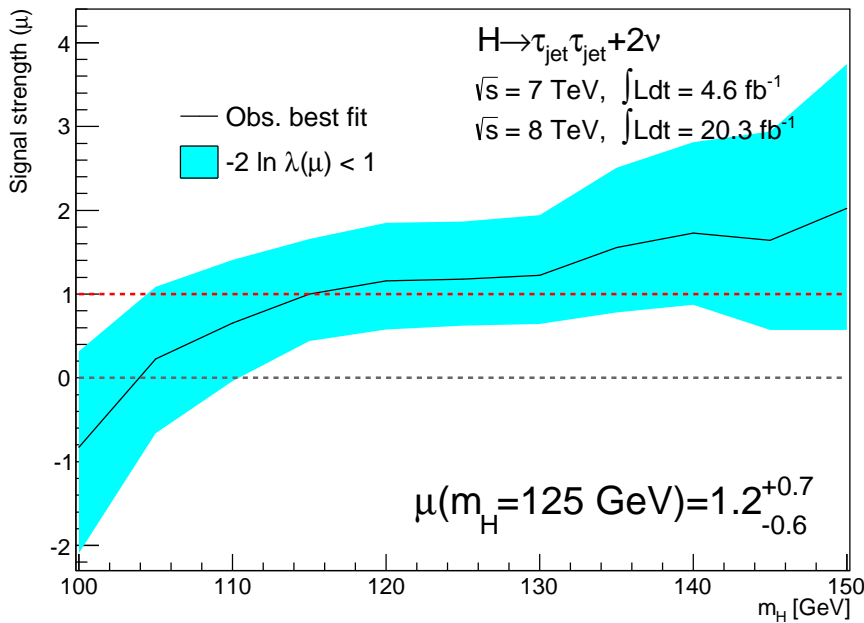

Figure 4.25: Measured signal strength μ in dependence on the Higgs boson mass m_H. Cyan band indicates the total uncertainty on the measured value of μ.

measured at $m_H = 125$ GeV is $\mu = 1.2 \pm 0.4(\text{stat})^{+0.5}_{-0.4}(\text{syst}) = 1.2^{+0.7}_{-0.6}$. The size of the uncertainty of μ is not directly related to the significance of the observed excess of events because the systematic uncertainties of the signal prediction are taken into account in the measurement of μ, but not in the significance of the excess. The total uncertainty of μ can be decomposed into several contributions as listed in Table 4.24. Given the low number of events in the most sensitive event categories, the uncertainty is dominated by the statistical uncertainty of the observed event yields. The statistical uncertainties of the background modelling procedure includes the statistical uncertainty of the expected total background contribution, while the uncertainty of the data-driven modelling includes the systematic uncertainties on the data-driven predictions based on the embedded $Z \to \tau\tau$ and the multi-jet control data samples. The signal strengths measured separately in the VBF and Boosted categories are $0.9^{+0.8}_{-0.6}$ and $1.9^{+1.3}_{-1.0}$, respectively, at $m_H = 125$ GeV. All measurements are in agreement with the SM expectation of unity.

In Fig. 4.26 the measured signal strength μ is compared with the values expected

Table 4.24: Contributions to the uncertainty on the measured signal strength μ.

Components	$\delta\mu$
Statistical	± 0.4
Systematic	$+0.5$ -0.4
Theoretical	$+0.3$ -0.1
Experimental	$+0.4$ -0.3
Background modelling, statistical	± 0.2
Data-driven modelling	± 0.2
τ-jet energy scale	± 0.2
τ-jet trigger and ID efficiency	± 0.1
Jet energy scale	± 0.1

in presence of a SM Higgs boson with mass $m_H = 105, 125$ or 145 GeV. The measured values are most compatible with the existence of a SM Higgs boson with $m_H = 125$ GeV.

Based on the linear correlation between the Higgs boson mass m_H and the corresponding average reconstructed mass $m_{\tau\tau}^{\text{MMC}}$ (Fig. 4.10) it is possible to measure the mass of the potential signal. For this purpose, the likelihood function (Eq. 4.8) is evaluated in dependence on m_H as shown in Fig. 4.27. The observed data is compatible with a Higgs boson with a mass $m_H = 125^{+16}_{-7}$ GeV. The uncertainties are estimated using a polynomial interpolation of the observed likelihood distribution between the scanned mass points. Even though the uncertainties are rather large, this result confirms the compatibility of the observed excess of events with the Higgs boson discovered in the di-boson final states.

4.9.1 Comparison with Results of Multivariate Analyses

In addition to the presented cut-based study, the same preselected events are analysed using a multivariate analysis (MVA) technique based on boosted decision trees (BDT) [67]. Rather than imposing cuts on discriminating variables, the MVA approach relies on several variables used as input to determine a probability for each event to be originating from a signal or background process (BDT score). Both cut-based and MVA analyses employ similar methods for the data-driven measurement of the background contributions and similar definitions of the VBF and Boosted event categories, but with a significantly different approaches for the statistical interpretation of the observed data. While the previously described results are based on the comparison of observed and expected $m_{\tau\tau}^{\text{MMC}}$ distributions in several event categories, the result of the MVA analysis is based on the distributions of the BDT score obtained separately for all VBF-like and Boosted-like events without any additional categorisation according to $p_T^{\tau\tau}$ or di-jet properties. Contrary to the presented analysis which is not optimised for any specific mass hypothesis within the probed mass range, the MVA-based study is optimised for the measurement of the signal cross section assuming a Higgs boson with mass $m_H = 125$ GeV.
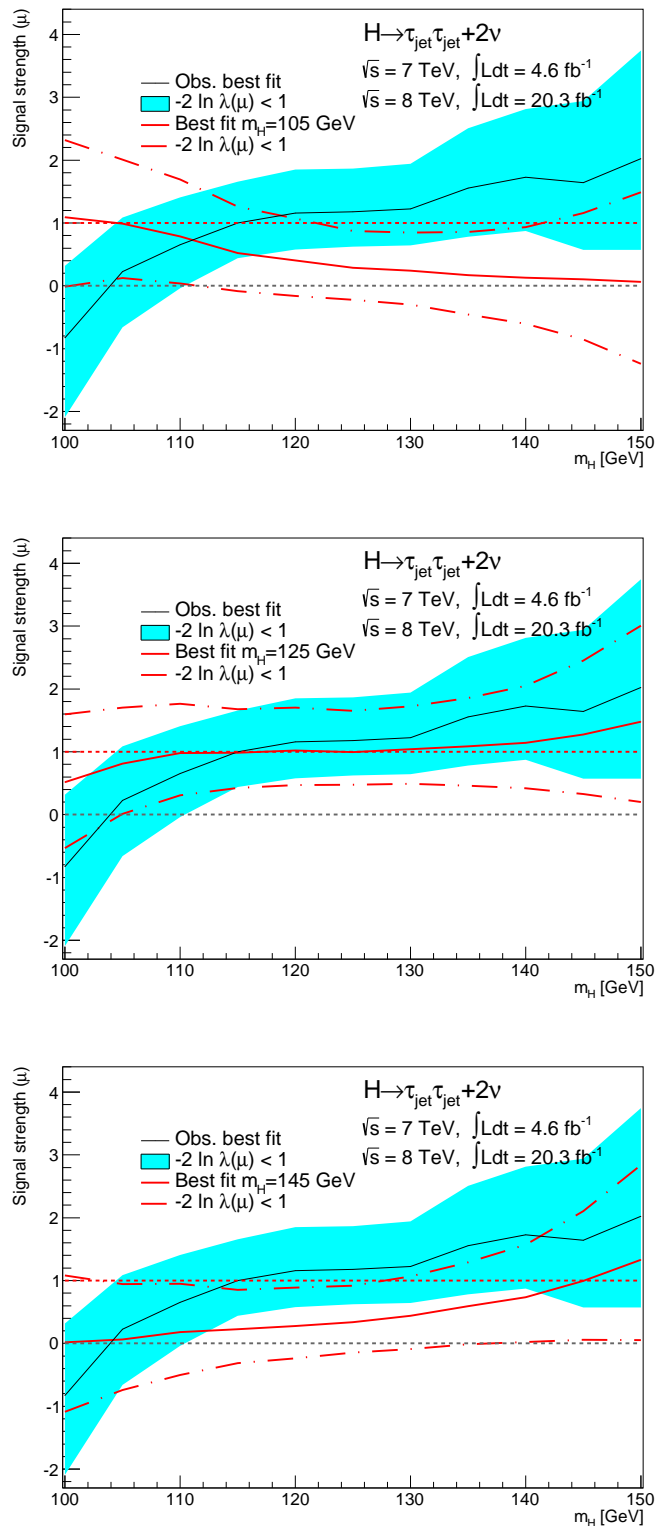
Figure 4.26: Measured signal strength μ in dependence on the Higgs boson mass, compared to the expected μ in presence of a SM Higgs boson with mass $m_H = 105$ (top), 125 (middle) or 145 GeV (bottom).
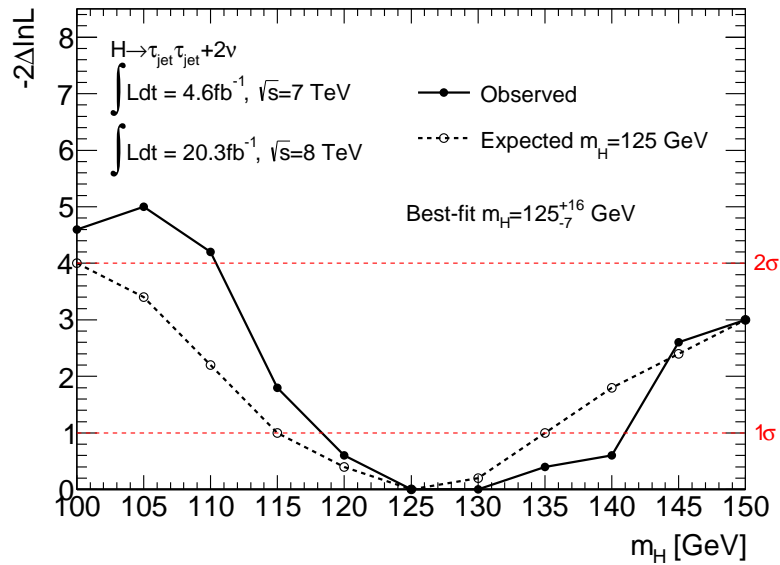
Figure 4.27: Observed negative log-likelihood values in dependence on the hypothesised mass m_H, compared with the expected values after assuming the existence of a Higgs boson with $m_H = 125$ GeV and a signal strength of $\mu = 1$.

Preliminary results obtained with the MVA analysis of $\sqrt{s} = 8$ TeV data are summarised in Ref. [67]. Fig. 4.28 shows the observed distributions of the BDT score in the VBF and Boosted event categories. An excess of events is observed in the high-score bins, where signal events are expected to accumulate. The corresponding observed signal strength of $\mu = 1.0^{+0.8}_{-0.6}$ is compatible with the result reported in this thesis. The comparable sizes of the uncertainties of μ obtained with the cut-based and MVA analysis indicate that the sensitivities of the two approaches are very similar.

The combination of this with the results of the MVA analyses performed for the fully leptonic and semi-leptonic $H \to \tau^+\tau^-$ final states leads to the observed (expected) significance of $4.1(3.2)\,\sigma$ at $m_H = 125$ GeV. This is the first evidence for the Higgs boson decay into fermions. Fig. 4.29 summarises the measured signal strengths separately for the three individual final states and for their statistical combination. Results are provided also separately for the VBF and Boosted event categories. The combined signal strength is $\mu = 1.4^{+0.5}_{-0.4}$ in agreement with the SM expectation. Signal strengths measured in the individual final states are compatible with each other as well as with the SM expectation. The uncertainties of the measured μ value obtained in each final state indicates the relative sensitivity of the searches performed in these different final states. The semi-leptonic channel is the most sensitive due to the highest $\tau\tau$ branching ratio, followed closely by the fully hadronic channel and then the fully leptonic channel.
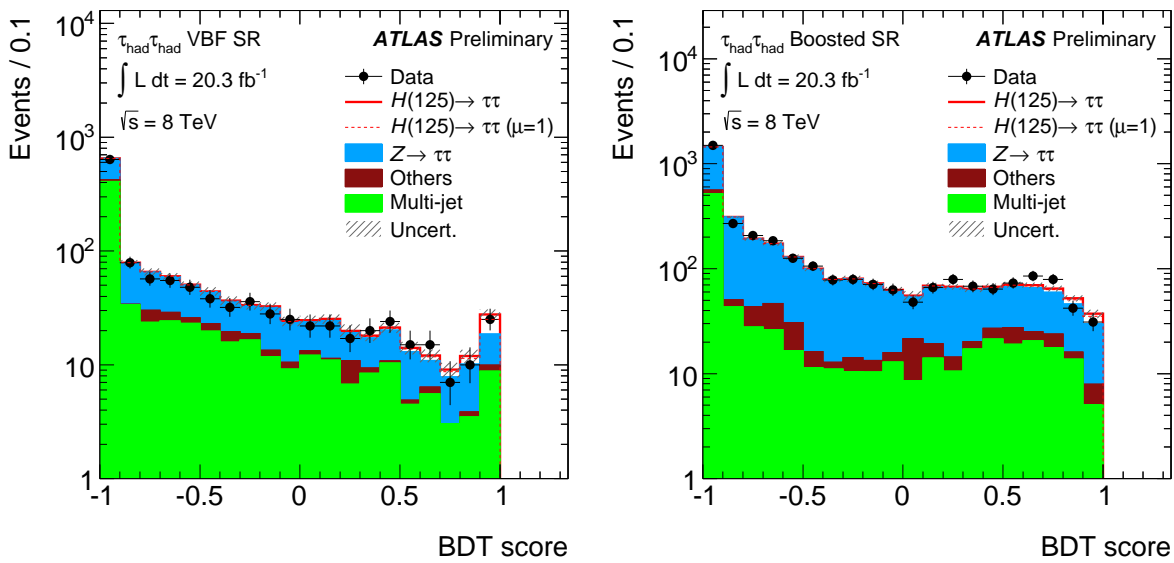
Figure 4.28: BDT score distributions in the VBF (left) and Boosted (right) event categories obtained by the MVA analysis of the fully hadronic final state at $\sqrt{s} = 8$ TeV. The expected Higgs boson signal ($m_H = 125$ GeV) is shown stacked with a signal strength of $\mu = 1$ (dashed line) and $\mu = 1.4$ (solid line). The normalisation of the background predictions is determined by the combined fit of all $H \rightarrow \tau^+\tau^-$ channels. The size of the statistical and normalisation systematic uncertainties is indicated by the hashed band [67].
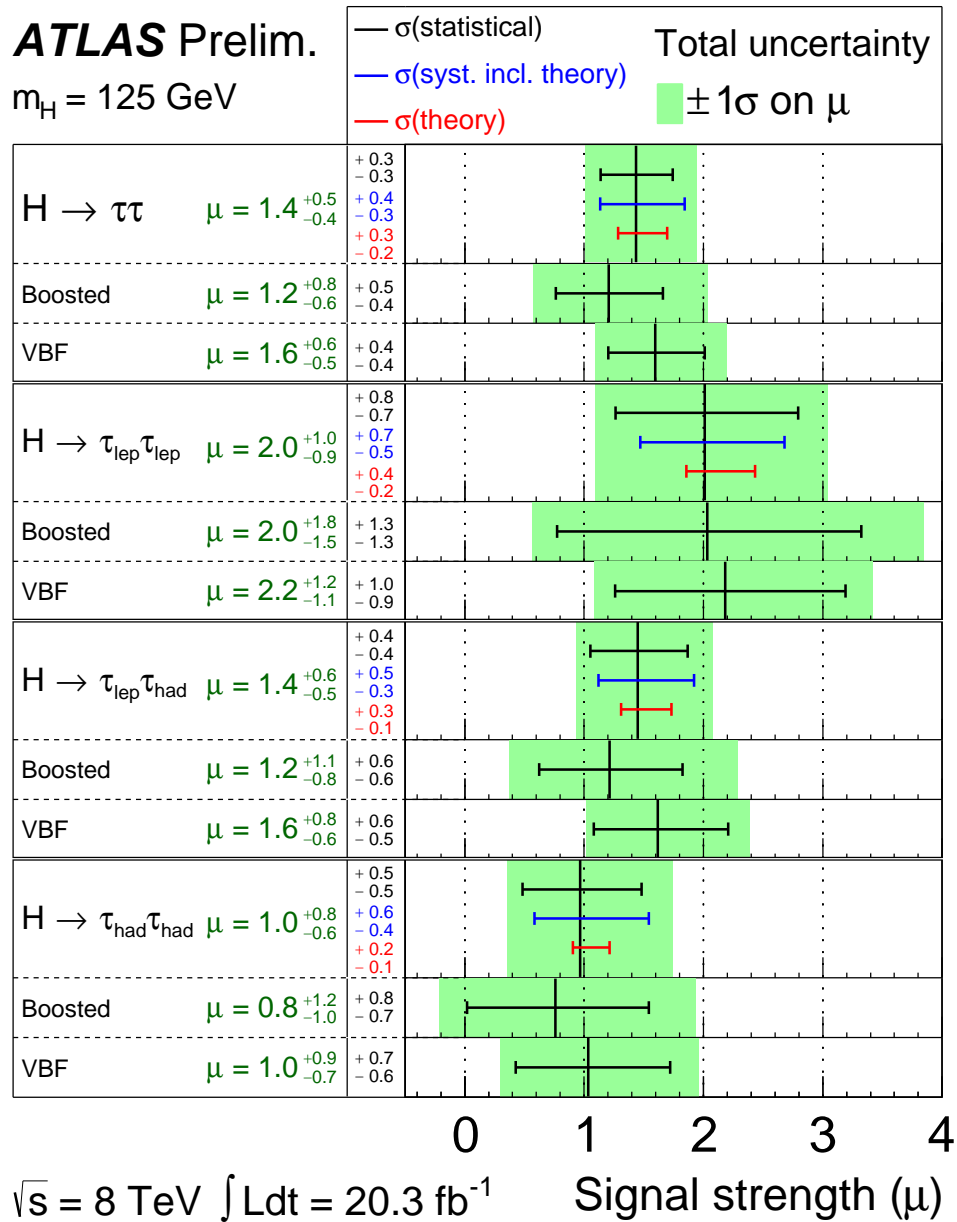
Figure 4.29: Signal strength μ in the three $H \to \tau^+\tau^-$ final states and their combination obtained with the MVA-based analyses. The corresponding total $\pm 1\sigma$ uncertainty is indicated by the shaded green band with the individual contributions from the statistical uncertainty (top, black), the total (experimental and theoretical) systematic uncertainty (middle, blue), and the theory uncertainty (bottom, red) alone indicated by the error bars and with the values listed in the central column [67].

Chapter 5

Conclusions

Electroweak symmetry breaking is an essential ingredient of the Standard Model of particle physics. It introduces masses to the weak gauge bosons and to the fermions without conflicting with the gauge symmetry of the electroweak interaction. The electroweak symmetry breaking manifests itself in the existence of an elementary scalar particle, the Higgs boson.

In 2012, the LHC experiments ATLAS and CMS discovered a Higgs boson of about $125\,\text{GeV}$ mass in $\gamma\gamma$, ZZ^* and WW^* decays with properties compatible with the Standard Model predictions. In 2013, also evidence for Higgs boson decays into τ lepton pairs was found.

The observation of Higgs boson decays into τ lepton pairs is important as it probes the coupling of the Higgs boson to fermions. The topic of this thesis is the search for $H \to \tau^+\tau^-$ production in the channel with subsequent hadronic decays of both τ leptons in proton-proton collisions at the LHC with the ATLAS detector. Challenging is the large multi-jet and $Z \to \tau\tau$ background in this channel. The large branching ratio of τ lepton decays into hadrons (τ-jets) plus τ neutrino and the better $\tau\tau$ invariant mass resolution are benefits compared to other τ lepton decay channels with leptons and, therefore, more neutrinos.

The sensitivity of this search strongly depends on the performance of the τ-jet trigger and reconstruction algorithms. The efficiency and discrimination power against quark and gluon jets of the trigger algorithm selecting events with at least two hadronically decaying τ leptons have been optimised. The τ-jet trigger and identification efficiencies have been measured using $W \to \tau\nu$ production data recorded at a center-of-mass energy of $\sqrt{s} = 7$ TeV. The measurement was used to correct the predictions of the Monte Carlo simulation of signal and background.

The search for the process $pp \to H + X \to \tau^+\tau^- + X$ in fully hadronic final states was performed with $4.6\,\text{fb}^{-1}$ and $20.3\,\text{fb}^{-1}$ of data collected by the ATLAS detector at center-of-mass energies of 7 and 8 TeV, respectively. The signal selection requirements are optimised for events with highly boosted Higgs bosons produced via gluon fusion with additional jet or via vector boson fusion. Several event categories with different background composition and signal-to-background ratio related to one of the two signal

production modes have been treated separately.

The two main background contributions from $Z \to \tau\tau$ and multi-jet production have been modelled using signal-free control data samples in order to reduce the systematic uncertainties from the Monte Carlo simulation of the background events. The remaining background contributions from W+jets, $t\bar{t}$ and di-boson production are estimated by simulation only.

The measured $\tau\tau$ invariant mass distributions in each event category are compared with the expected signal and background distributions. An excess of events above the expected background is observed in the signal region: the observed (expected) significance of a Higgs boson signal with a mass m_H of 125 GeV is $2.3\,\sigma\,(2.1\,\sigma)$. The measured signal strength relative to the Standard Model prediction is $\mu = 1.2 \pm 0.4(\text{stat})^{+0.5}_{-0.4}(\text{syst})$ in agreement with the Standard Model. The mass of the resonance is measured to be $m_H = 125^{+16}_{-7}$ GeV.

The results of this cut-based analysis are in agreement with those obtained with multivariate analysis techniques. The sensitivities are comparable after the optimisation of the event selection requirements in this analysis. The combination of searches for $H \to \tau^+\tau^-$ decays in fully leptonic and semi-leptonic and fully hadronic $\tau^+\tau^-$ final states in $\sqrt{s} = 8$ TeV data with multivariate analyses results in a $4.1\,\sigma$ evidence for the Higgs boson decays into τ leptons.

Bibliography

[1] S. L. Glashow, *Partial Symmetries of Weak Interactions*,
Nucl. Phys. **22** (1961) 579–588.

[2] S. Weinberg, *A Model of Leptons*, Phys. Rev. Lett. **19** (1967) 1264–1266.

[3] A. Salam, *in Elementary Particle Theory, p. 367*. Almqvist and Wiksell,
Stockholm, 1968.

[4] S. Glashow, J. Iliopoulos and L. Maiani, *Weak Interactions with Lepton-Hadron
Symmetry*, Phys.Rev. **D2** (1970) 1285–1292.

[5] ATLAS Collaboration, G. Aad et al., *Observation of a new particle in the search
for the Standard Model Higgs boson with the ATLAS detector at the LHC*,
Phys.Lett. **B716** (2012) 1–29, `arXiv:1207.7214 [hep-ex]`.

[6] CMS Collaboration, S. Chatrchyan et al., *Observation of a new boson at a mass of
125 GeV with the CMS experiment at the LHC*, Phys.Lett. **B716** (2012) 30–61,
`arXiv:1207.7235 [hep-ex]`.

[7] J. Goldstone, *Field theories with "superconductor" solutions*, Nuovo Cimento **19**
(1960) 154–164.

[8] P. W. Anderson, *Plasmons, Gauge Invariance, and Mass*,
Phys. Rev. **130** (1963) 439–442,
`http://link.aps.org/doi/10.1103/PhysRev.130.439`.

[9] F. Englert and R. Brout, *Broken Symmetry and the Mass of Gauge Vector
Mesons*, Phys.Rev.Lett. **13** (1964) 321–323.

[10] P. W. Higgs, *Broken symmetries, massless particles and gauge fields*,
Phys.Lett. **12** (1964) 132–133.

[11] P. W. Higgs, *Spontaneous Symmetry Breakdown without Massless Bosons*,
Phys.Rev. **145** (1966) 1156–1163.

[12] G. Guralnik, C. Hagen and T. Kibble, *Global Conservation Laws and Massless
Particles*, Phys.Rev.Lett. **13** (1964) 585–587.

[13] J. M. Cornwall, D. N. Levin and G. Tiktopoulos, *Uniqueness of Spontaneously
Broken Gauge Theories*, Phys. Rev. Lett. **30** (1973) 1268–1270,
`http://link.aps.org/doi/10.1103/PhysRevLett.30.1268`.

[14] J. M. Cornwall, D. N. Levin and G. Tiktopoulos, *Derivation of gauge invariance from high-energy unitarity bounds on the S matrix*,
Phys. Rev. D **10** (1974) 1145–1167,
http://link.aps.org/doi/10.1103/PhysRevD.10.1145.

[15] C. L. Smith, *High energy behaviour and gauge symmetry*,
Phys.Lett. B **46** (1973) no. 2, 233 – 236,
http://www.sciencedirect.com/science/article/pii/0370269373906928.

[16] B. W. Lee, C. Quigg and H. B. Thacker, *Weak interactions at very high energies: The role of the Higgs-boson mass*, Phys. Rev. D **16** (1977) 1519–1531,
http://link.aps.org/doi/10.1103/PhysRevD.16.1519.

[17] N. Cabibbo, L. Maiani, G. Parisi and R. Petronzio, *Bounds on the fermions and Higgs boson masses in grand unified theories*,
Nuclear Physics B **158** (1979) no. 2–3, 295 – 305,
http://www.sciencedirect.com/science/article/pii/0550321379901676.

[18] G. Altarelli and G. Isidori, *Lower limit on the Higgs mass in the standard model: An update*, Phys.Lett. B **337** (1994) no. 1–2, 141 – 144,
http://www.sciencedirect.com/science/article/pii/0370269394914583.

[19] J. Casas, J. Espinosa and M. Quirós, *Improved Higgs mass stability bound in the standard model and implications for supersymmetry*,
Phys.Lett. B **342** (1995) no. 1–4, 171 – 179,
http://www.sciencedirect.com/science/article/pii/037026939401404Z.

[20] J. Casas, J. Espinosa and M. Quirós, *Standard model stability bounds for new physics within {LHC} reach*, Phys.Lett. B **382** (1996) no. 4, 374 – 382,
http://www.sciencedirect.com/science/article/pii/037026939600682X.

[21] T. Hambye and K. Riesselmann, *Matching conditions and Higgs boson mass upper bounds reexamined*, Phys. Rev. D **55** (1997) 7255–7262,
http://link.aps.org/doi/10.1103/PhysRevD.55.7255.

[22] J. F. Gunion, H. E. Haber, G. L. Kane and S. Dawson, *The Higgs Hunter's Guide*,
Front.Phys. **80** (2000) 1–448.

[23] G. Degrassi, S. Di Vita, J. Elias-Miro, J. R. Espinosa, G. F. Giudice et al., *Higgs mass and vacuum stability in the Standard Model at NNLO*,
JHEP **1208** (2012) 098, arXiv:1205.6497 [hep-ph].

[24] M. Baak, M. Goebel, J. Haller, A. Hoecker, D. Kennedy et al., *The Electroweak Fit of the Standard Model after the Discovery of a New Boson at the LHC*,
Eur.Phys.J. **C72** (2012) 2205, arXiv:1209.2716 [hep-ph].

[25] Particle Data Group Collaboration, J. Beringer et al., *Review of Particle Physics*,
Phys. Rev. D **86** (2012) 010001,
http://link.aps.org/doi/10.1103/PhysRevD.86.010001.

[26] LHC Higgs Cross Section Working Group, S. Dittmaier, C. Mariotti, G. Passarino and R. Tanaka (Eds.), *Handbook of LHC Higgs Cross Sections: 1. Inclusive Observables*, CERN-2011-002, `arXiv:1101.0593 [hep-ph]`.

[27] LHC Higgs Cross Section Working Group, S. Dittmaier, C. Mariotti, G. Passarino and R. Tanaka (Eds.), *Handbook of LHC Higgs Cross Sections: 2. Differential Distributions*, CERN-2012-002, `arXiv:1201.3084 [hep-ph]`.

[28] D. Graudenz, M. Spira and P. Zerwas, *QCD corrections to Higgs boson production at proton proton colliders*, Phys.Rev.Lett. **70** (1993) 1372–1375.

[29] M. Spira, A. Djouadi, D. Graudenz and P. Zerwas, *Higgs boson production at the LHC*, Nucl.Phys. **B453** (1995) 17–82, `arXiv:hep-ph/9504378 [hep-ph]`.

[30] R. V. Harlander, *Virtual corrections to $gg \to H$ to two loops in the heavy top limit*, Phys.Lett. **B492** (2000) 74–80, `arXiv:hep-ph/0007289 [hep-ph]`.

[31] S. Catani, D. de Florian and M. Grazzini, *Higgs production in hadron collisions: Soft and virtual QCD corrections at NNLO*, JHEP **0105** (2001) 025, `arXiv:hep-ph/0102227 [hep-ph]`.

[32] R. V. Harlander and W. B. Kilgore, *Soft and virtual corrections to proton proton $\to H + x$ at NNLO*, Phys.Rev. **D64** (2001) 013015, `arXiv:hep-ph/0102241 [hep-ph]`.

[33] R. V. Harlander and W. B. Kilgore, *Next-to-next-to-leading order Higgs production at hadron colliders*, Phys.Rev.Lett. **88** (2002) 201801, `arXiv:hep-ph/0201206 [hep-ph]`.

[34] C. Anastasiou and K. Melnikov, *Higgs boson production at hadron colliders in NNLO QCD*, Nucl.Phys. **B646** (2002) 220–256, `arXiv:hep-ph/0207004 [hep-ph]`.

[35] V. Ravindran, J. Smith and W. L. van Neerven, *NNLO corrections to the total cross-section for Higgs boson production in hadron hadron collisions*, Nucl.Phys. **B665** (2003) 325–366, `arXiv:hep-ph/0302135 [hep-ph]`.

[36] J. Blumlein and V. Ravindran, *Mellin moments of the next-to-next-to leading order coefficient functions for the Drell-Yan process and hadronic Higgs-boson production*, Nucl.Phys. **B716** (2005) 128–172, `arXiv:hep-ph/0501178 [hep-ph]`.

[37] C. Anastasiou, R. Boughezal and F. Petriello, *Mixed QCD-electroweak corrections to Higgs boson production in gluon fusion*, JHEP **0904** (2009) 003, `arXiv:0811.3458 [hep-ph]`.

[38] S. Marzani, R. D. Ball, V. Del Duca, S. Forte and A. Vicini, *Higgs production via gluon-gluon fusion with finite top mass beyond next-to-leading order*, Nucl.Phys. **B800** (2008) 127–145, `arXiv:0801.2544 [hep-ph]`.

[39] R. V. Harlander and K. J. Ozeren, *Top mass effects in Higgs production at next-to-next-to-leading order QCD: Virtual corrections*, Phys.Lett. **B679** (2009) 467–472, `arXiv:0907.2997 [hep-ph]`.

[40] R. V. Harlander and K. J. Ozeren, *Finite top mass effects for hadronic Higgs production at next-to-next-to-leading order*, JHEP **0911** (2009) 088, `arXiv:0909.3420 [hep-ph]`.

[41] R. V. Harlander, H. Mantler, S. Marzani and K. J. Ozeren, *Higgs production in gluon fusion at next-to-next-to-leading order QCD for finite top mass*, Eur.Phys.J. **C66** (2010) 359–372, `arXiv:0912.2104 [hep-ph]`.

[42] A. Pak, M. Rogal and M. Steinhauser, *Finite top quark mass effects in NNLO Higgs boson production at LHC*, JHEP **1002** (2010) 025, `arXiv:0911.4662 [hep-ph]`.

[43] A. Pak, M. Rogal and M. Steinhauser, *Virtual three-loop corrections to Higgs boson production in gluon fusion for finite top quark mass*, Phys.Lett. **B679** (2009) 473–477, `arXiv:0907.2998 [hep-ph]`.

[44] R. D. Ball, M. Bonvini, S. Forte, S. Marzani and G. Ridolfi, *Higgs production in gluon fusion beyond NNLO*, Nucl.Phys. **B874** (2013) 746–772, `arXiv:1303.3590 [hep-ph]`.

[45] S. Buehler and A. Lazopoulos, *Scale dependence and collinear subtraction terms for Higgs production in gluon fusion at N3LO*, `arXiv:1306.2223 [hep-ph]`.

[46] S. Catani, D. de Florian, M. Grazzini and P. Nason, *Soft gluon resummation for Higgs boson production at hadron colliders*, JHEP **0307** (2003) 028, `arXiv:hep-ph/0306211 [hep-ph]`.

[47] A. Djouadi and P. Gambino, *Leading electroweak correction to Higgs boson production at proton colliders*, Phys.Rev.Lett. **73** (1994) 2528–2531, `arXiv:hep-ph/9406432 [hep-ph]`.

[48] U. Aglietti, R. Bonciani, G. Degrassi and A. Vicini, *Two loop light fermion contribution to Higgs production and decays*, Phys.Lett. **B595** (2004) 432–441, `arXiv:hep-ph/0404071 [hep-ph]`.

[49] G. Degrassi and F. Maltoni, *Two-loop electroweak corrections to Higgs production at hadron colliders*, Phys.Lett. **B600** (2004) 255–260, `arXiv:hep-ph/0407249 [hep-ph]`.

[50] S. Actis, G. Passarino, C. Sturm and S. Uccirati, *NNLO Computational Techniques: The Cases $H \to gamma\ gamma$ and $H \to g\ g$*, Nucl.Phys. **B811** (2009) 182–273, `arXiv:0809.3667 [hep-ph]`.

[51] S. Actis, G. Passarino, C. Sturm and S. Uccirati, *NLO Electroweak Corrections to Higgs Boson Production at Hadron Colliders*, Phys.Lett. **B670** (2008) 12–17, `arXiv:0809.1301 [hep-ph]`.

[52] M. Grazzini and H. Sargsyan, *Heavy-quark mass effects in Higgs boson production at the LHC*, JHEP **09** (2013) 129, `arXiv:1306.4581 [hep-ph]`.

[53] M. Ciccolini, A. Denner and S. Dittmaier, *Strong and electroweak corrections to the production of Higgs + 2jets via weak interactions at the LHC*, Phys.Rev.Lett. **99** (2007) 161803, `arXiv:0707.0381 [hep-ph]`.

[54] M. Ciccolini, A. Denner and S. Dittmaier, *Electroweak and QCD corrections to Higgs production via vector-boson fusion at the LHC*, Phys.Rev. **D77** (2008) 013002, `arXiv:0710.4749 [hep-ph]`.

[55] P. Bolzoni, F. Maltoni, S.-O. Moch and M. Zaro, *Higgs production via vector-boson fusion at NNLO in QCD*, Phys.Rev.Lett. **105** (2010) 011801, `arXiv:1003.4451 [hep-ph]`.

[56] G. Ferrera, M. Grazzini and F. Tramontano, *Associated WH production at hadron colliders: a fully exclusive QCD calculation at NNLO*, Phys.Rev.Lett. **107** (2011) 152003, `arXiv:1107.1164 [hep-ph]`.

[57] O. Brein, R. Harlander, M. Wiesemann and T. Zirke, *Top-Quark Mediated Effects in Hadronic Higgs-Strahlung*, Eur.Phys.J. **C72** (2012) 1868, `arXiv:1111.0761 [hep-ph]`.

[58] M. Ciccolini, S. Dittmaier and M. Kramer, *Electroweak radiative corrections to associated WH and ZH production at hadron colliders*, Phys.Rev. **D68** (2003) 073003, `arXiv:hep-ph/0306234 [hep-ph]`.

[59] W. Beenakker, S. Dittmaier, M. Kramer, B. Plumper, M. Spira et al., *Higgs radiation off top quarks at the Tevatron and the LHC*, Phys.Rev.Lett. **87** (2001) 201805, `arXiv:hep-ph/0107081 [hep-ph]`.

[60] W. Beenakker, S. Dittmaier, M. Kramer, B. Plumper, M. Spira et al., *NLO QCD corrections to t anti-t H production in hadron collisions*, Nucl.Phys. **B653** (2003) 151–203, `arXiv:hep-ph/0211352 [hep-ph]`.

[61] L. Reina and S. Dawson, *Next-to-leading order results for t anti-t h production at the Tevatron*, Phys.Rev.Lett. **87** (2001) 201804, `arXiv:hep-ph/0107101 [hep-ph]`.

[62] S. Dawson, L. Orr, L. Reina and D. Wackeroth, *Associated top quark Higgs boson production at the LHC*, Phys.Rev. **D67** (2003) 071503, `arXiv:hep-ph/0211438 [hep-ph]`.

[63] LEP Collaboration, R. Barate et al., *Search for the standard model Higgs boson at LEP*, Phys.Lett. **B565** (2003) 61–75, `arXiv:hep-ex/0306033`.

[64] CDF and D0 Collaborations, *Higgs Boson Studies at the Tevatron*, `arXiv:1303.6346 [hep-ex]`.

[65] CMS Collaboration, S. Chatrchyan et al., *Search for a standard-model-like Higgs boson with a mass in the range 145 to 1000 GeV at the LHC*, Eur.Phys.J. **C73** (2013) 2469, `arXiv:1304.0213 [hep-ex]`.

[66] ATLAS Collaboration, G. Aad et al., *Search for the bb decay of the Standard Model Higgs boson in associated W/ZH production with the ATLAS detector*, ATLAS-CONF-2013-079, CERN, Geneva, July 2013.

[67] ATLAS Collaboration, G. Aad et al., *Evidence for Higgs Boson Decays to the $\tau^+\tau^-$ Final State with the ATLAS Detector*, ATLAS-CONF-2013-108, CERN, Geneva, Nov. 2013.

[68] CMS Collaboration, S. Chatrchyan et al., *Evidence for the direct decay of the 125 GeV Higgs boson to fermions*, `arXiv:1401.6527 [hep-ex]`.

[69] CMS Collaboration, S. Chatrchyan et al., *Combination of standard model Higgs boson searches and measurements of the properties of the new boson with a mass near 125 GeV*, CMS-PAS-HIG-13-005, CERN, Geneva, 2013.

[70] CMS Collaboration, S. Chatrchyan et al., *Study of the Mass and Spin-Parity of the Higgs Boson Candidate Via Its Decays to Z Boson Pairs*, Phys.Rev.Lett. **110** (2013) 081803, `arXiv:1212.6639 [hep-ex]`.

[71] ATLAS Collaboration, G. Aad et al., *Measurements of Higgs boson production and couplings in diboson final states with the ATLAS detector at the LHC*, Phys.Lett. **B726** (2013) 88–119, `arXiv:1307.1427 [hep-ex]`.

[72] ATLAS Collaboration, G. Aad et al., *Evidence for the spin-0 nature of the Higgs boson using ATLAS data*, Phys.Lett. **B726** (2013) 120–144, `arXiv:1307.1432 [hep-ex]`.

[73] ATLAS Collaboration, G. Aad et al., *Measurements of the properties of the Higgs-like boson in the two photon decay channel with the ATLAS detector using 25 fb^{-1} of proton-proton collision data*, ATLAS-CONF-2013-012, CERN, Geneva, March 2013.

[74] ATLAS Collaboration, G. Aad et al., *Measurements of the properties of the Higgs-like boson in the four lepton decay channel with the ATLAS detector using 25 fb^{-1} of proton-proton collision data*, ATLAS-CONF-2013-013, CERN, Geneva, March 2013.

[75] ATLAS Collaboration, G. Aad et al., *Updated coupling measurements of the Higgs boson with the ATLAS detector using up to 25 fb^{-1} of proton-proton collision data*, ATLAS-CONF-2014-009, CERN, Geneva, March 2014.

[76] Y. Gao, A. V. Gritsan, Z. Guo, K. Melnikov, M. Schulze et al., *Spin determination of single-produced resonances at hadron colliders*, Phys.Rev. **D81** (2010) 075022, `arXiv:1001.3396 [hep-ph]`.

[77] L. Evans and P. Bryant, *LHC Machine*, JINST **3** (2008) S08001.

[78] ATLAS Collaboration, G. Aad et al., *Luminosity Determination in pp Collisions at $\sqrt{s} = 7$ TeV Using the ATLAS Detector at the LHC*, Eur.Phys.J. **C71** (2011) 1630, `arXiv:1101.2185 [hep-ex]`.

[79] ATLAS Collaboration, G. Aad et al., *The ATLAS Experiment at the CERN Large Hadron Collider*, JINST **3** (2008) S08003.

[80] CMS Collaboration, S. Chatrchyan et al., *The CMS experiment at the CERN LHC*, JINST **3** (2008) S08004.

[81] ALICE Collaboration, K. Aamodt et al., *The ALICE experiment at the CERN LHC*, JINST **3** (2008) S08002.

[82] LHCb Collaboration, J. Alves, A. Augusto et al., *The LHCb Detector at the LHC*, JINST **3** (2008) S08005.

[83] ATLAS Collaboration, G. Aad et al., *Improved luminosity determination in pp collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector at the LHC*, Eur.Phys.J. **C73** (2013) 2518, `arXiv:1302.4393 [hep-ex]`.

[84] ATLAS Collaboration, G. Aad et al., *Improved electron reconstruction in ATLAS using the Gaussian Sum Filter-based model for bremsstrahlung*, ATLAS-CONF-2012-047, CERN, Geneva, May 2012.

[85] ATLAS Collaboration, G. Aad et al., *Preliminary results on the muon reconstruction efficiency, momentum resolution, and momentum scale in ATLAS 2012 pp collision data*, ATLAS-CONF-2013-088, CERN, Geneva, Aug. 2013.

[86] M. Cacciari, G. P. Salam and G. Soyez, *The Anti-k(t) jet clustering algorithm*, JHEP **0804** (2008) 063, `arXiv:0802.1189 [hep-ph]`.

[87] ATLAS Collaboration, G. Aad et al., *Jet energy measurement with the ATLAS detector in proton-proton collisions at $\sqrt{s} = 7$ TeV*, Eur.Phys.J. **C73** (2013) 2304, `arXiv:1112.6426 [hep-ex]`.

[88] ATLAS Collaboration, G. Aad et al., *Performance of Missing Transverse Momentum Reconstruction in ATLAS studied in Proton-Proton Collisions recorded in 2012 at 8 TeV*, ATLAS-CONF-2013-082, CERN, Geneva, Aug. 2013.

[89] ATLAS Collaboration, G. Aad et al., *Determination of the tau energy scale and the associated systematic uncertainty in proton-proton collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector at the LHC in 2012*, ATLAS-CONF-2013-044, CERN, Geneva, April 2013.

[90] ATLAS Collaboration, G. Aad et al., *Performance of the Reconstruction and Identification of Hadronic Tau Decays with ATLAS*, ATLAS-CONF-2011-152, CERN, Geneva, Nov. 2011.

[91] ATLAS Collaboration, G. Aad et al., *Reconstruction, Energy Calibration, and Identification of Hadronically Decaying Tau Leptons*, ATLAS-CONF-2011-077, CERN, Geneva, May 2011.

[92] ATLAS Collaboration, G. Aad et al., *Identification of the Hadronic Decays of Tau Leptons in 2012 Data with the ATLAS Detector*, ATLAS-CONF-2013-064, CERN, Geneva, July 2013.

[93] ATLAS Collaboration, G. Aad et al., *Performance of the Reconstruction and Identification of Hadronic Tau Decays in ATLAS with 2011 Data*, ATLAS-CONF-2012-142, CERN, Geneva, Oct. 2012.

[94] ATLAS Collaboration, G. Aad et al., *Measurement of hadronic tau decay identification efficiency using $W \to \tau\nu$ events*, ATLAS-CONF-2011-093, CERN, Geneva, July 2011.

[95] ATLAS Collaboration, G. Aad et al., *Performance of the ATLAS tau trigger in 2011*, ATLAS-CONF-2013-006, CERN, Geneva, Jan. 2013.

[96] D. Zanzi and S. Xella, *Search for the Standard Model Higgs boson in the ATLAS Experiment*, Master Thesis of D. Zanzi, CERN-THESIS-2011-002, Copenhagen University, Copenhagen, 2011.

[97] ATLAS Collaboration, G. Aad et al., *Search for neutral MSSM Higgs bosons decaying to tau^+tau^- pairs in proton-proton collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector*, Phys.Lett. **B705** (2011) 174–192, `arXiv:1107.5003 [hep-ex]`.

[98] A. Elagin, P. Murat, A. Pranko and A. Safonov, *A New Mass Reconstruction Technique for Resonances Decaying to di-tau*, Nucl. Instr. and Meth. **A654** (2011) 481–489, `arXiv:1012.4686 [hep-ex]`.

[99] D. Zanzi, *Search for the SM Higgs Boson in the Fully Hadronic $\tau^+\tau^-$ Final State with the ATLAS Detector Using $25\,\mathrm{fb}^{-1}$ of pp Collision Data at the LHC*, ATLAS internal note ATL-COM-PHYS-2013-1558, CERN, Geneva, Nov. 2013.

[100] S. Alioli, P. Nason, C. Oleari and E. Re, *NLO Higgs boson production via gluon fusion matched with shower in POWHEG*, JHEP **0904** (2009) 002, `arXiv:0812.0578 [hep-ph]`.

[101] P. Nason and C. Oleari, *NLO Higgs boson production via vector-boson fusion matched with shower in POWHEG*, JHEP **1002** (2010) 037, `arXiv:0911.5299 [hep-ph]`.

[102] G. Bozzi, S. Catani, D. de Florian and M. Grazzini, *Transverse-momentum resummation and the spectrum of the Higgs boson at the LHC*, Nucl.Phys. **B737** (2006) 73–120, `arXiv:hep-ph/0508068 [hep-ph]`.

[103] D. de Florian, G. Ferrera, M. Grazzini and D. Tommasini, *Transverse-momentum resummation: Higgs boson production at the Tevatron and the LHC*, JHEP **1111** (2011) 064, `arXiv:1109.2109 [hep-ph]`.

[104] T. Sjostrand, S. Mrenna and P. Z. Skands, *A Brief Introduction to PYTHIA 8.1*, Comp.Phys.Comm. **178** (2008) 852–867, `arXiv:0710.3820 [hep-ph]`.

[105] LHC Higgs Cross Section Working Group Collaboration, S. Heinemeyer et al., *Handbook of LHC Higgs Cross Sections: 3. Higgs Properties*, `arXiv:1307.1347 [hep-ph]`.

[106] M. L. Mangano, M. Moretti, F. Piccinini, R. Pittau and A. D. Polosa, *ALPGEN, a generator for hard multiparton processes in hadronic collisions*, JHEP **0307** (2003) 001, `arXiv:hep-ph/0206293 [hep-ph]`.

[107] M. L. Mangano, M. Moretti and R. Pittau, *Multijet matrix elements and shower evolution in hadronic collisions: $Wb\bar{b} + n$ jets as a case study*, Nucl. Phys. **B632** (2002) 343–362, `arXiv:hep-ph/0108069`.

[108] S. Frixione and B. R. Webber, *Matching NLO QCD computations and parton shower simulations*, JHEP **0206** (2002) 029, `arXiv:hep-ph/0204244 [hep-ph]`.

[109] B. P. Kersevan and E. Richter-Was, *The Monte Carlo event generator AcerMC version 2.0 with interfaces to PYTHIA 6.2 and HERWIG 6.5*, `arXiv:hep-ph/0405247 [hep-ph]`.

[110] G. Corcella et al., *HERWIG 6.5 release note*, `arXiv:hep-ph/0210213`.

[111] J. Butterworth, J. R. Forshaw and M. Seymour, *Multiparton interactions in photoproduction at HERA*, Z.Phys. **C72** (1996) 637–646, `arXiv:hep-ph/9601371 [hep-ph]`.

[112] T. Binoth, M. Ciccolini, N. Kauer and M. Kramer, *Gluon-induced W-boson pair production at the LHC*, JHEP **0612** (2006) 046, `arXiv:hep-ph/0611170 [hep-ph]`.

[113] Z. Was, *TAUOLA the library for tau lepton decay, and KKMC / KORALB / KORALZ /... status report*, Nucl.Phys.Proc.Suppl. **98** (2001) 96–102, `arXiv:hep-ph/0011305 [hep-ph]`.

[114] N. Davidson, T. Przedzinski and Z. Was, *PHOTOS Interface in C++: Technical and Physics Documentation*, `arXiv:1011.0937 [hep-ph]`.

[115] H.-L. Lai, M. Guzzi, J. Huston, Z. Li, P. M. Nadolsky et al., *New parton distributions for collider physics*, Phys.Rev. **D82** (2010) 074024, `arXiv:1007.2241 [hep-ph]`.

[116] J. Pumplin, D. Stump, J. Huston, H. Lai, P. M. Nadolsky et al., *New generation of parton distributions with uncertainties from global QCD analysis*, JHEP **0207** (2002) 012, `arXiv:hep-ph/0201195 [hep-ph]`.

[117] GEANT4 Collaboration, S. Agostinelli et al., *GEANT4: A Simulation toolkit*, Nucl. Instr. and Meth. **A506** (2003) 250–303.

[118] R. K. Ellis, I. Hinchliffe, M. Soldate and J. van der Bij, *Higgs Decay to $\tau^+\tau^-$: A Possible Signature of Intermediate Mass Higgs Bosons at the SSC*, Nucl.Phys. **B297** (1988) 221.

[119] ATLAS Collaboration, G. Aad et al., *Performance of Missing Transverse Momentum Reconstruction in Proton-Proton Collisions at 7 TeV with ATLAS*, Eur.Phys.J. **C72** (2012) 1844, `arXiv:1108.5602 [hep-ex]`.

[120] ATLAS Collaboration, G. Aad et al., *Search for the Standard Model Higgs boson in the $H \to \tau^+\tau^-$ decay mode with $4.7 fb^-1$ of ATLAS data at 7 TeV*, ATLAS-CONF-2012-014, CERN, Geneva, March 2012.

[121] K. Hagiwara, T. Li, K. Mawatari and J. Nakamura, *TauDecay: a library to simulate polarized tau decays via FeynRules and MadGraph5*, Eur.Phys.J. **C73** (2013) 2489, `arXiv:1212.6247 [hep-ph]`.

[122] Z. Czyczula, T. Przedzinski and Z. Was, *TauSpinner Program for Studies on Spin Effect in tau Production at the LHC*, Eur.Phys.J. **C72** (2012) 1988, `arXiv:1201.0117 [hep-ph]`.

[123] ATLAS and CMS Collaborations, *Procedure for the LHC Higgs boson search combination in summer 2011*, ATL-PHYS-PUB-2011-011, CERN, Geneva, Aug. 2011.

[124] ATLAS Collaboration, G. Aad et al., *Measurements of the properties of the Higgs-like boson in the $WW^{(*)} \to \ell\nu\ell\nu$ decay channel with the ATLAS detector using $25 fb^{-1}$ of proton-proton collision data*, ATLAS-CONF-2013-030, CERN, Geneva, March 2013.

[125] E. L. Berger and J. M. Campbell, *Higgs boson production in weak boson fusion at next-to-leading order*, Phys.Rev. **D70** (2004) 073011, `arXiv:hep-ph/0403194 [hep-ph]`.

[126] J. M. Campbell, R. K. Ellis and G. Zanderighi, *Next-to-Leading order Higgs + 2 jet production via gluon fusion*, JHEP **0610** (2006) 028, `arXiv:hep-ph/0608194 [hep-ph]`.

[127] ATLAS Collaboration, G. Aad et al., *Combined Standard Model Higgs boson searches with up to $2.3 fb^{-1}$ of pp collisions at $\sqrt{s} = 7$ TeV at the LHC*, ATLAS-CONF-2011-157, CERN, Geneva, Nov. 2011.

[128] ROOT Collaboration, K. Cranmer, G. Lewis, L. Moneta, A. Shibata, W. Verkerke et al., *HistFactory: A tool for creating statistical models for use with RooFit and RooStats*, CERN-OPEN-2012-016, New York University, New York, Jan. 2012.

[129] F. James and M. Roos, *Minuit: A System for Function Minimization and Analysis of the Parameter Errors and Correlations*, Comp.Phys.Comm. **10** (1975) 343–367.

[130] G. Cowan, K. Cranmer, E. Gross and O. Vitells, *Asymptotic formulae for likelihood-based tests of new physics*, Eur.Phys.J. **C71** (2011) 1554, `arXiv:1007.1727 [physics.data-an]`.

[131] A. L. Read, *Modified frequentist analysis of search results (The CL(s) method)*, CERN-OPEN-2000-205, CERN, Geneva, 2000.

[132] A. L. Read, *Presentation of search results: the CL s technique*, Journal of Physics G **28** (2002) no. 10, 2693, `http://stacks.iop.org/0954-3899/28/i=10/a=313`.

Appendices

Appendix A

Validation of the Background Modelling

This section shows distributions of discriminating variables observed in validation regions used to test the modelling of the background contributions. These regions are the Preselection and the VBF and Boosted validation regions defined in Section 4.6.3.

In all plots, backgrounds are estimated as described in Section 4.6. The $Z \to \tau\tau$ and multi-jet yields are determined fitting the $m_{\tau\tau}^{\mathrm{MMC}}$ distribution (Fig. 4.15a). The expected yield for a Higgs boson with $m_H = 125$ GeV and a cross section fifty or ten times larger than the one predicted by the SM is overlaid. The hashed area is the statistical uncertainty on the background model. The lower panel shows the ratio of the data to the total background. The yellow band includes the statistical uncertainty on the background model, the multi-jet shape uncertainty and the multi-jet and $Z \to \tau\tau$ yield uncertainties. The black error bars indicate the statistical uncertainty of data. In each histogram the first and the last bins contain under- and over-flows, respectively. All plots in this section show distributions in $\sqrt{s} = 8$ TeV data. Corresponding $\sqrt{s} = 7$ TeV plots are given in Ref. [99].

The first plots are related to the kinematic of each of the two selected τ-jets. Fig. A.1 shows the transverse momentum, Fig. A.2 the pseudo-rapidity and Fig. A.3 the azimuthal angle. The drop in acceptance around $|\eta| = 0$ is because that region is poorly instrumented. Fig. A.4 shows the $\Delta R_{\tau\tau}$ and $\Delta\eta_{\tau\tau}$ angles between the two τ-jets. Fig. A.5 shows the reconstructed $E_{\mathrm{T}}^{\mathrm{miss}}$ and its direction with respect to the two τ-jets as $\min\left\{\Delta\phi(E_{\mathrm{T}}^{\mathrm{miss}}, \tau_{\mathrm{jet},1}), \Delta\phi(E_{\mathrm{T}}^{\mathrm{miss}}, \tau_{\mathrm{jet},2})\right\}/\pi$, where the events with the $\mathbf{E}_{\mathrm{T}}^{\mathrm{miss}}$ is pointing in between the two τ-jets have value zero. Fig. A.6 shows the invariant mass of the two τ-jets $m_{\tau\tau}^{\mathrm{vis}}$ and the transverse momentum of the reconstructed Higgs boson candidate $p_{\mathrm{T}}^{\tau\tau}$. The following plots examine the modelling of jets reconstructed in the final state. Fig. A.7 shows the number of jets in preselected events. The depletion of events with no jets is due to the upper cuts on $\Delta R_{\tau\tau}$ at Preselection which rejects events with back-to-back τ-jets and enhances the fraction of events where the $\tau\tau$ pair is recoiling from some jet activity. Fig. A.8 shows the transverse momentum and the pseudo-rapidity of the jet with highest p_{T} in the VBF and Boosted VRs. Fig. A.9 displays the transverse momentum and the pseudo-rapidity of the jet with the second highest p_{T} in the VBF VR. Fig. A.10 shows the $|\Delta\eta_{\mathrm{jj}}|$ and the M_{jj} invariant mass of the two tagging jets in the VBF VR. Finally, Fig.

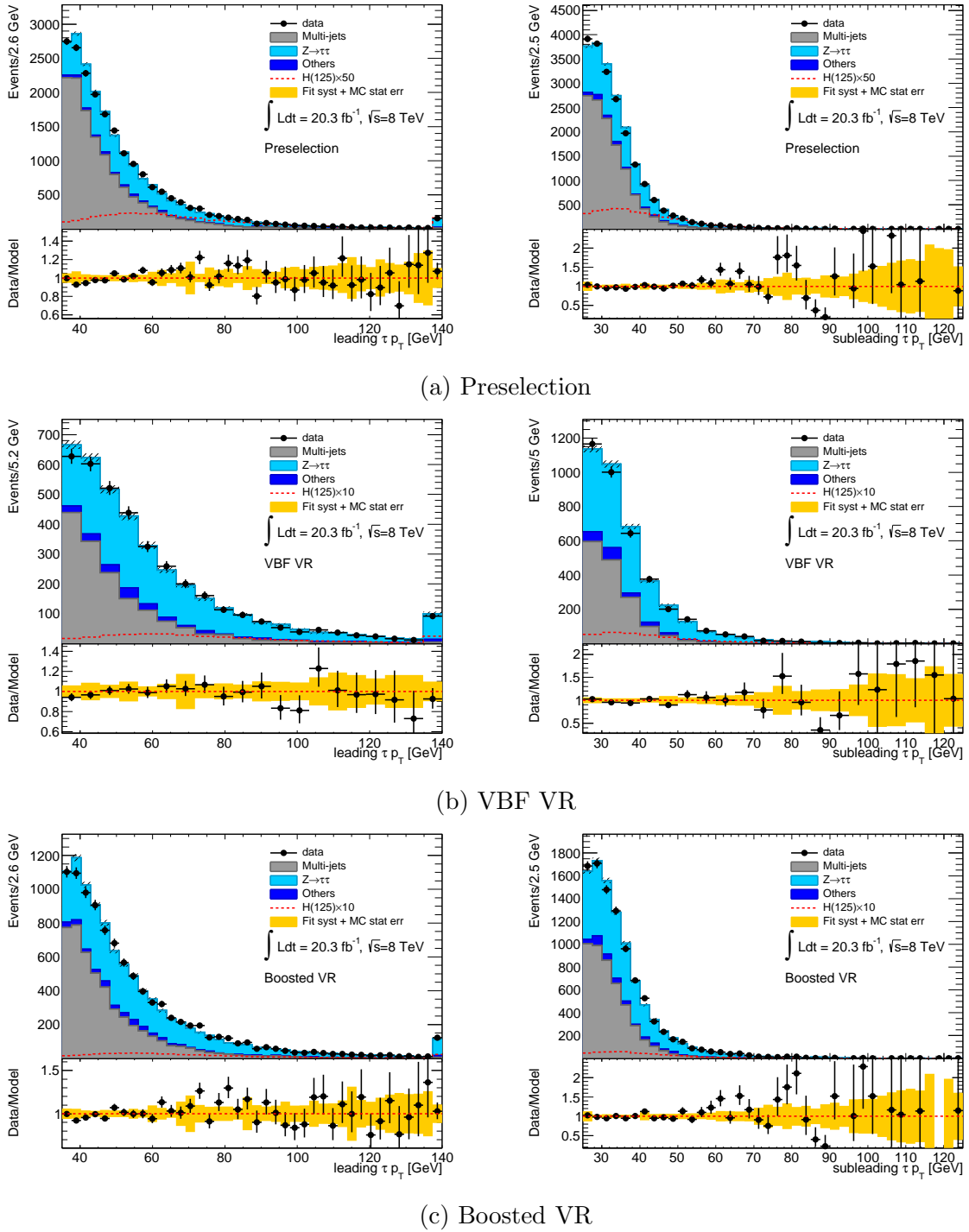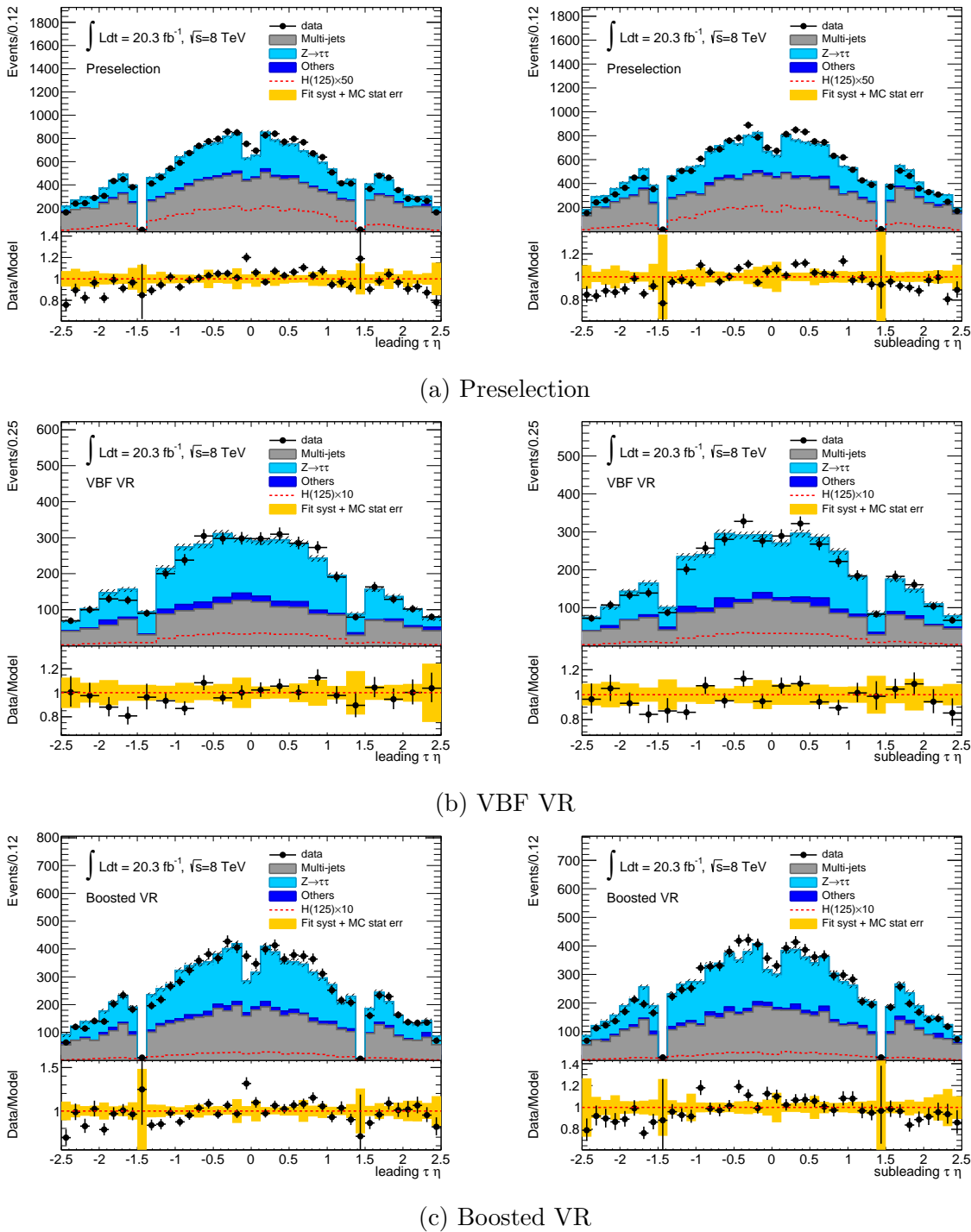(a) Preselection

(b) VBF VR

(c) Boosted VR

Figure A.1: Transverse momentum of the leading (left) and sub-leading (right) τ-jet in $\sqrt{s} = 8$ TeV data. The hashed area on each plot indicates the statistical uncertainty of the background model. The lower panel shows the ratio of the observed data to the total predicted background. The yellow band includes the statistical uncertainty of the background model, the multi-jet shape uncertainty and the multi-jet and $Z \rightarrow \tau\tau$ event yield uncertainties. The first and the last bins include the under- and over-flow events, respectively.
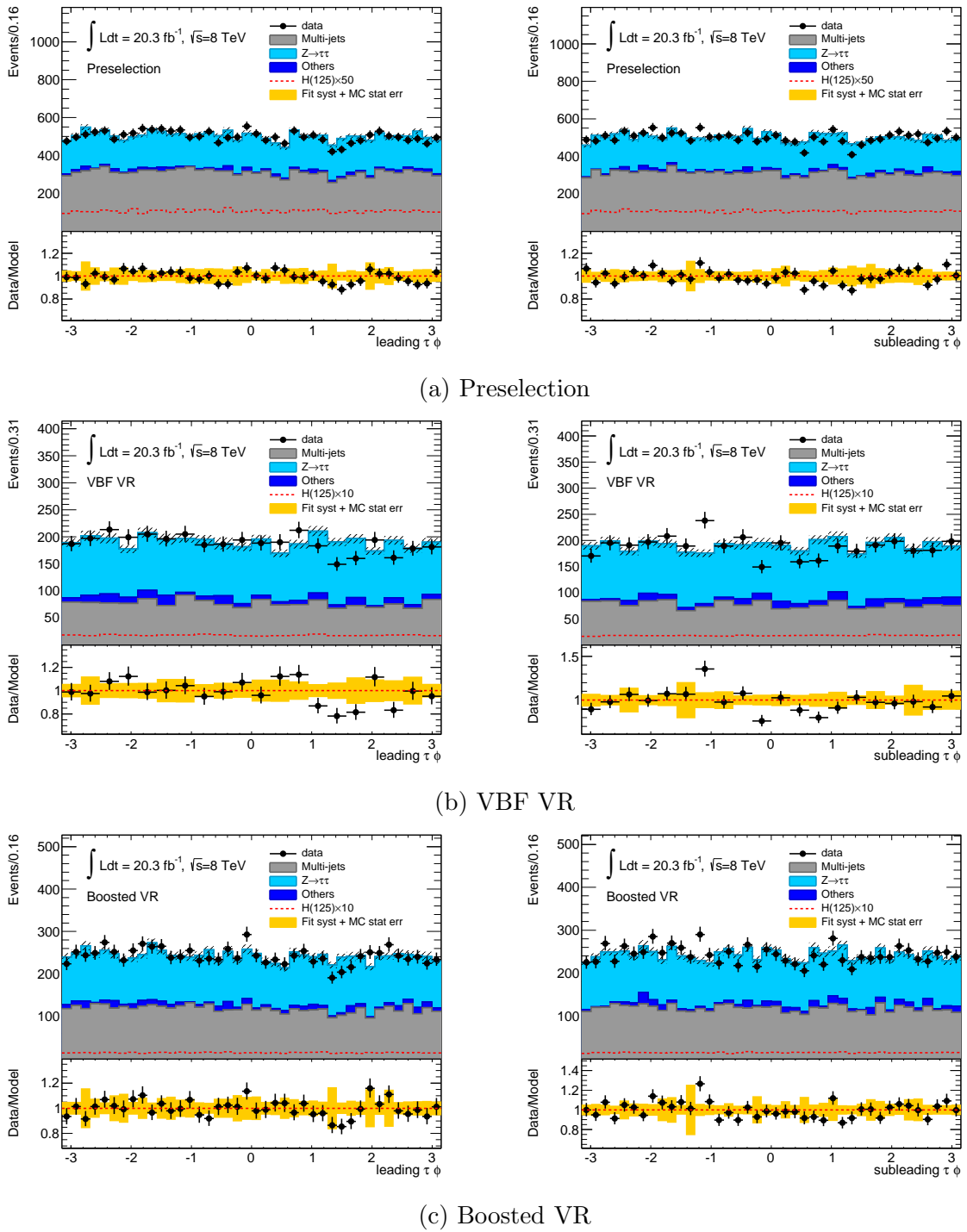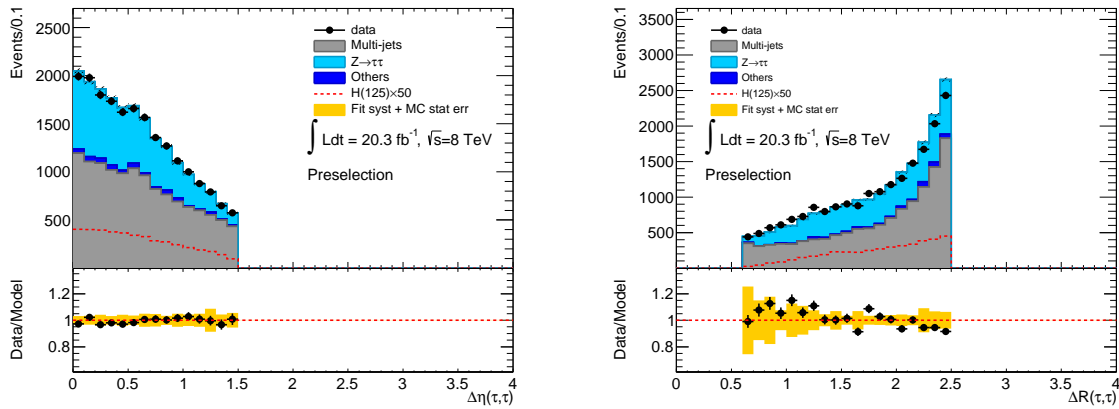
(a) Preselection

(b) VBF VR

(c) Boosted VR

Figure A.2: Pseudo-rapidity of the leading (left) and sub-leading (right) τ-jet in $\sqrt{s} = 8$ TeV data. The hashed area on each plot indicates the statistical uncertainty of the background model. The lower panel shows the ratio of the observed data to the total predicted background. The yellow band includes the statistical uncertainty of the background model, the multi-jet shape uncertainty and the multi-jet and $Z \to \tau\tau$ event yield uncertainties. The first and the last bins include the under- and over-flow events, respectively.

(a) Preselection

(b) VBF VR

(c) Boosted VR

Figure A.3: Azimuthal angle of the leading (left) and sub-leading (right) τ-jet in $\sqrt{s} = 8$ TeV data. The hashed area on each plot indicates the statistical uncertainty of the background model. The lower panel shows the ratio of the observed data to the total predicted background. The yellow band includes the statistical uncertainty of the background model, the multi-jet shape uncertainty and the multi-jet and $Z \to \tau\tau$ event yield uncertainties. The first and the last bins include the under- and over-flow events, respectively.

(a) Preselection
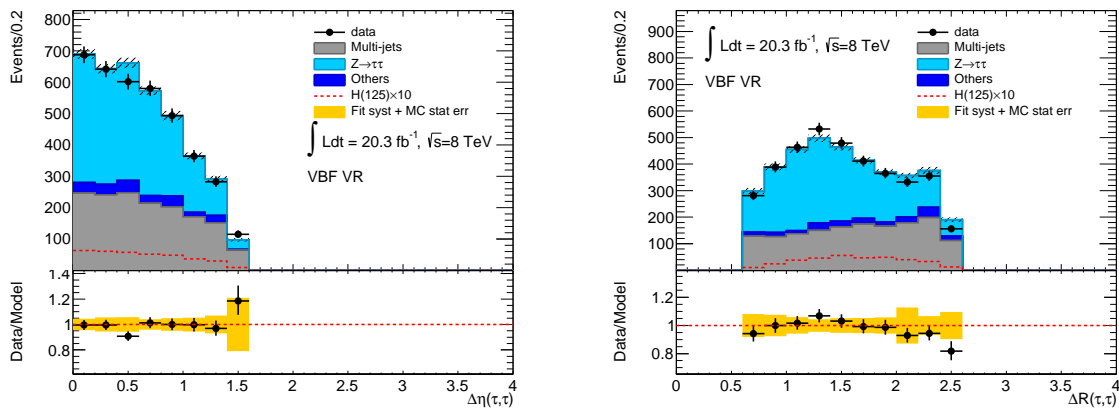
(b) VBF VR
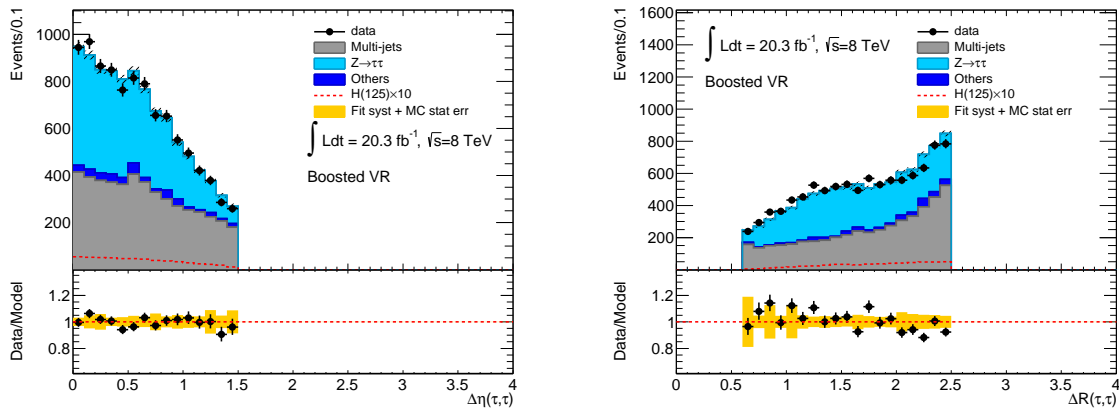
(c) Boosted VR

Figure A.4: $\Delta\eta_{\tau\tau}$ (left) and $\Delta R_{\tau\tau}$ (right) between the two τ-jets in $\sqrt{s} = 8$ TeV data. The hashed area on each plot indicates the statistical uncertainty of the background model. The lower panel shows the ratio of the observed data to the total predicted background. The yellow band includes the statistical uncertainty of the background model, the multi-jet shape uncertainty and the multi-jet and $Z \to \tau\tau$ event yield uncertainties. The first and the last bins include the under- and over-flow events, respectively.
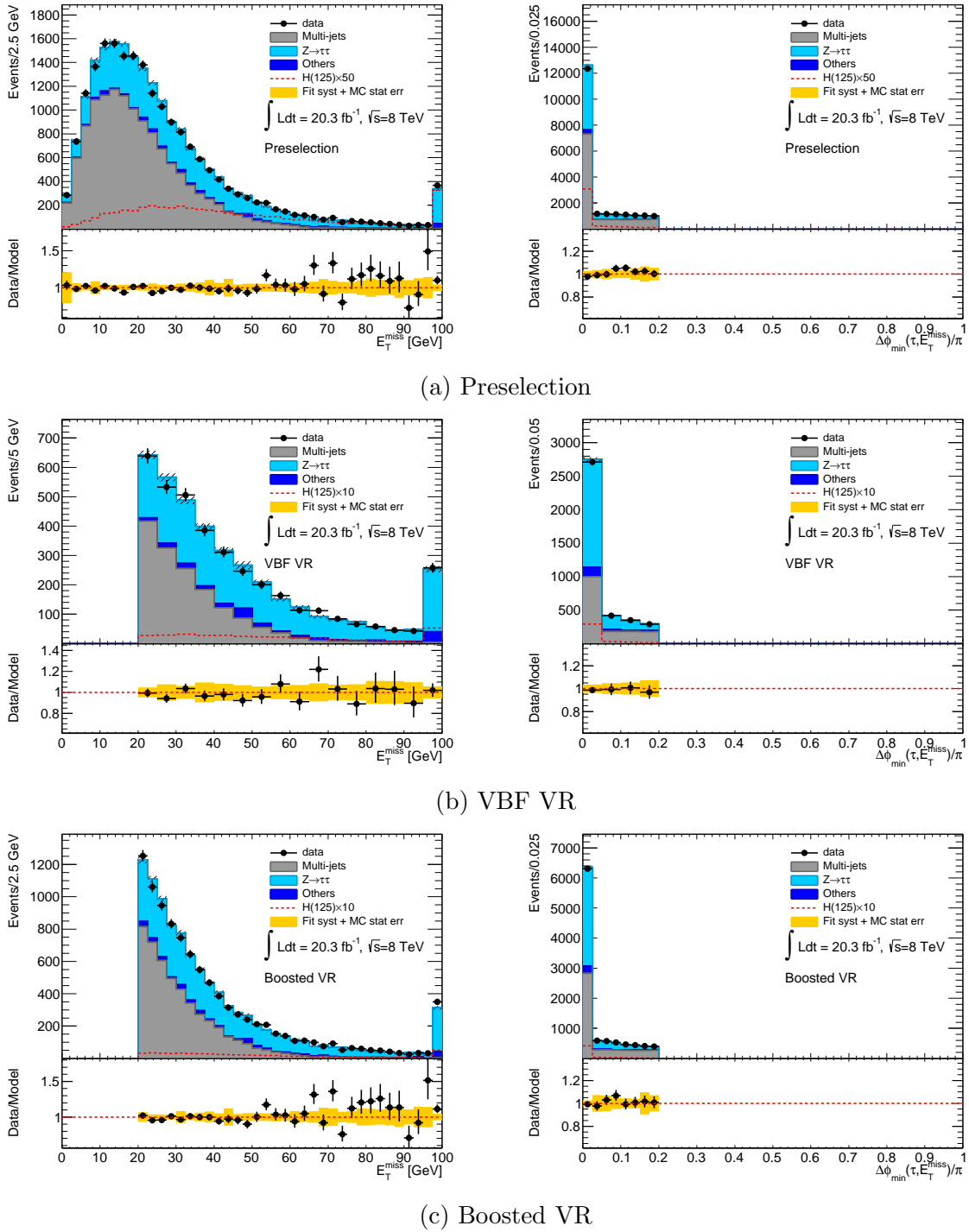
(a) Preselection

(b) VBF VR

(c) Boosted VR

Figure A.5: $E_{\mathrm{T}}^{\mathrm{miss}}$ (left) and $\min\left\{\Delta\phi(E_{\mathrm{T}}^{\mathrm{miss}}, \tau_{\mathrm{jet},1}), \Delta\phi(E_{\mathrm{T}}^{\mathrm{miss}}, \tau_{\mathrm{jet},2})\right\}/\pi$ (right) in $\sqrt{s} = $ 8 TeV data. The hashed area on each plot indicates the statistical uncertainty of the background model. The lower panel shows the ratio of the observed data to the total predicted background. The yellow band includes the statistical uncertainty of the background model, the multi-jet shape uncertainty and the multi-jet and $Z \to \tau\tau$ event yield uncertainties. The first and the last bins include the under- and over-flow events, respectively.

(a) Preselection

(b) VBF VR

(c) Boosted VR

Figure A.6: Invariant mass of the two τ-jets (left) and the transverse momentum of the Higgs boson candidate (right) in $\sqrt{s} = 8$ TeV data. The hashed area on each plot indicates the statistical uncertainty of the background model. The lower panel shows the ratio of the observed data to the total predicted background. The yellow band includes the statistical uncertainty of the background model, the multi-jet shape uncertainty and the multi-jet and $Z \rightarrow \tau\tau$ event yield uncertainties. The first and the last bins include the under- and over-flow events, respectively.
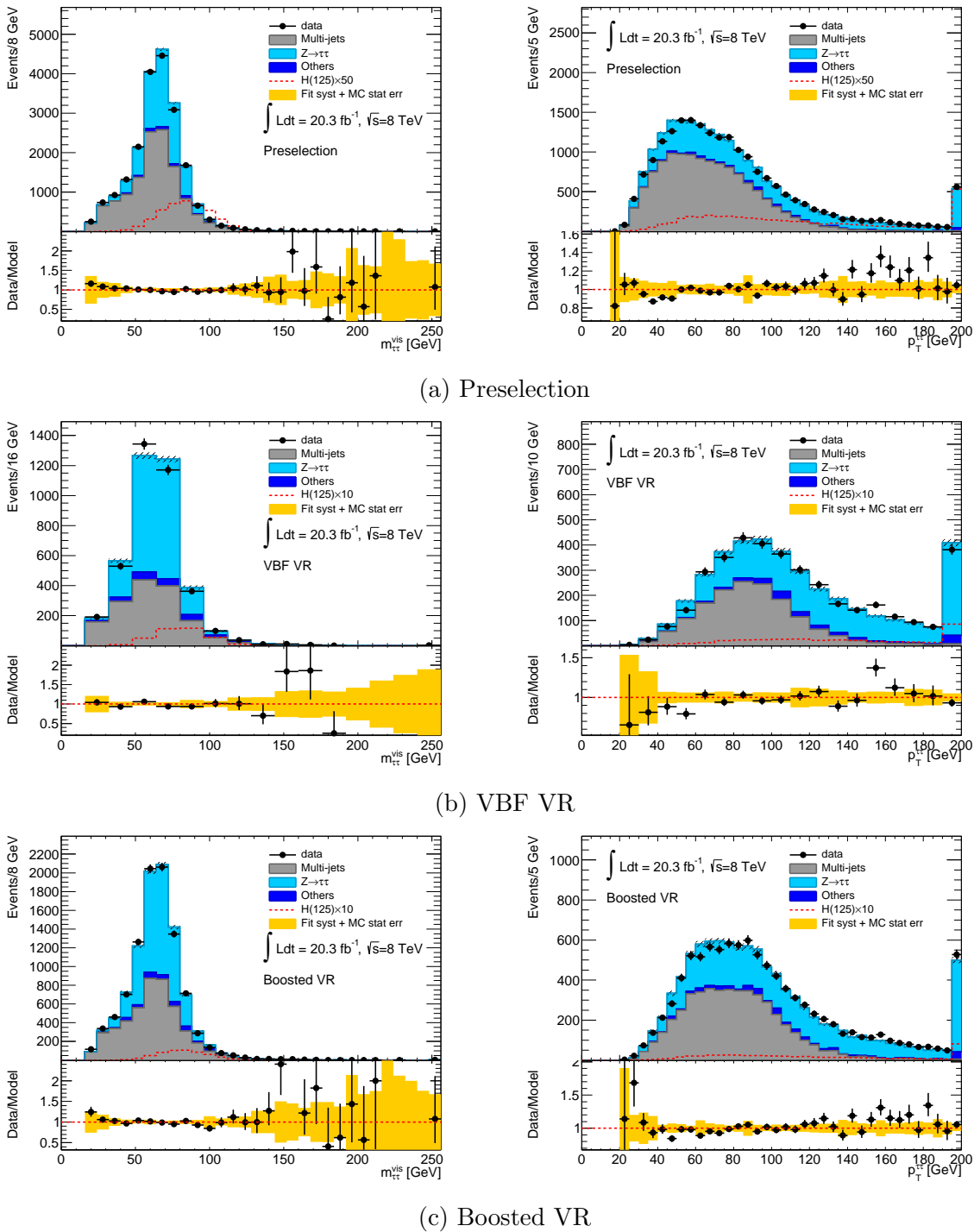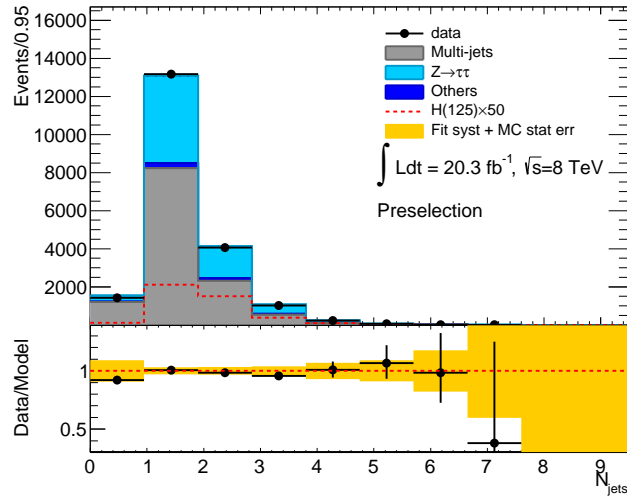
Figure A.7: Number of jets in preselected events in $\sqrt{s} = 8$ TeV data. The hashed area on each plot indicates the statistical uncertainty of the background model. The lower panel shows the ratio of the observed data to the total predicted background. The yellow band includes the statistical uncertainty of the background model, the multi-jet shape uncertainty and the multi-jet and $Z \to \tau\tau$ event yield uncertainties. The first and the last bins include the under- and over-flow events, respectively.

A.11 shows the $m_{\tau\tau}^{\mathrm{MMC}}$ and the $m_{\tau\tau}^{\mathrm{Coll}}$ invariant masses computed with the MMC algorithm and the collinear mass approximation, respectively.

The observed agreement in all plots indicates that the background estimation method is suitable to predict the background contributions in the event category used for the search.
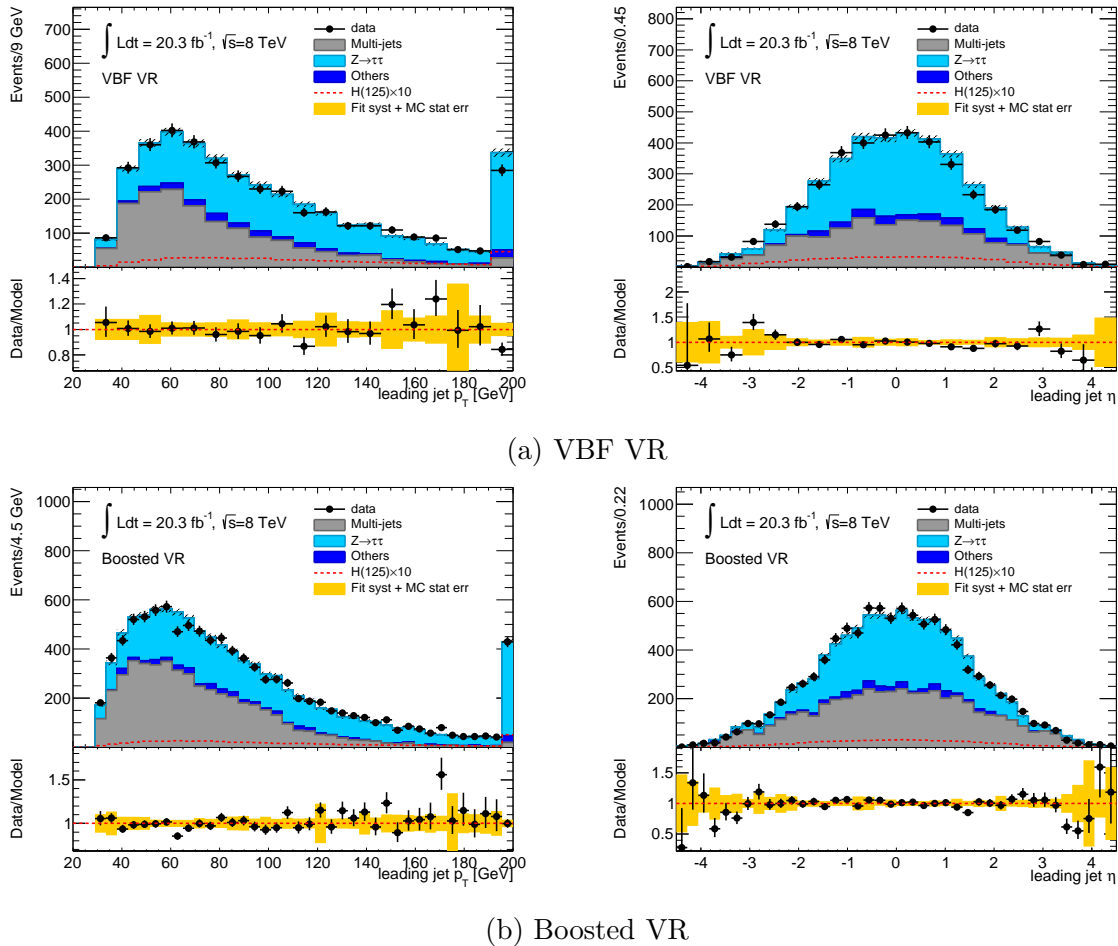
(a) VBF VR

(b) Boosted VR

Figure A.8: Transverse momentum (left) and pseudo-rapidity (right) of the jet with highest p_T in $\sqrt{s} = 8$ TeV data. The hashed area on each plot indicates the statistical uncertainty of the background model. The lower panel shows the ratio of the observed data to the total predicted background. The yellow band includes the statistical uncertainty of the background model, the multi-jet shape uncertainty and the multi-jet and $Z \to \tau\tau$ event yield uncertainties. The first and the last bins include the under- and over-flow events, respectively.
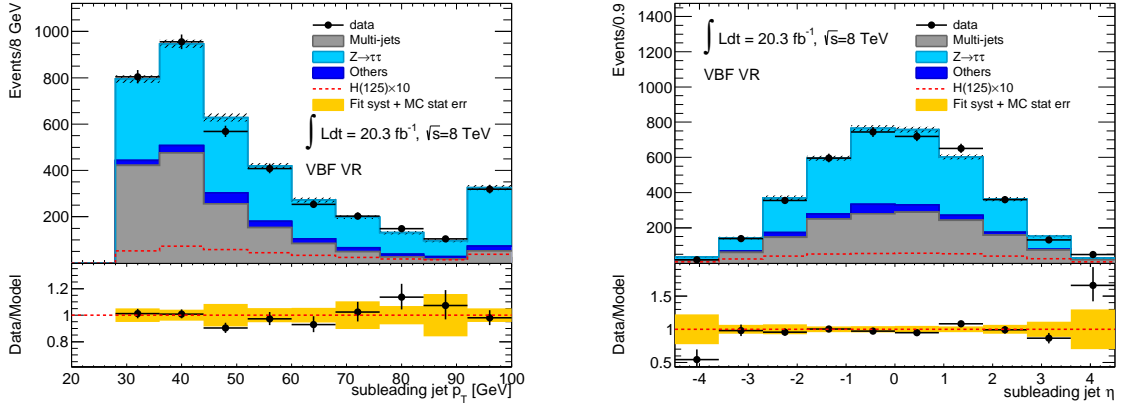
Figure A.9: Transverse momentum (left) and pseudo-rapidity (right) of the jet with the second highest p_T in the VBF VR in $\sqrt{s} = 8$ TeV data. The hashed area on each plot indicates the statistical uncertainty of the background model. The lower panel shows the ratio of the observed data to the total predicted background. The yellow band includes the statistical uncertainty of the background model, the multi-jet shape uncertainty and the multi-jet and $Z \to \tau\tau$ event yield uncertainties. The first and the last bins include the under- and over-flow events, respectively.
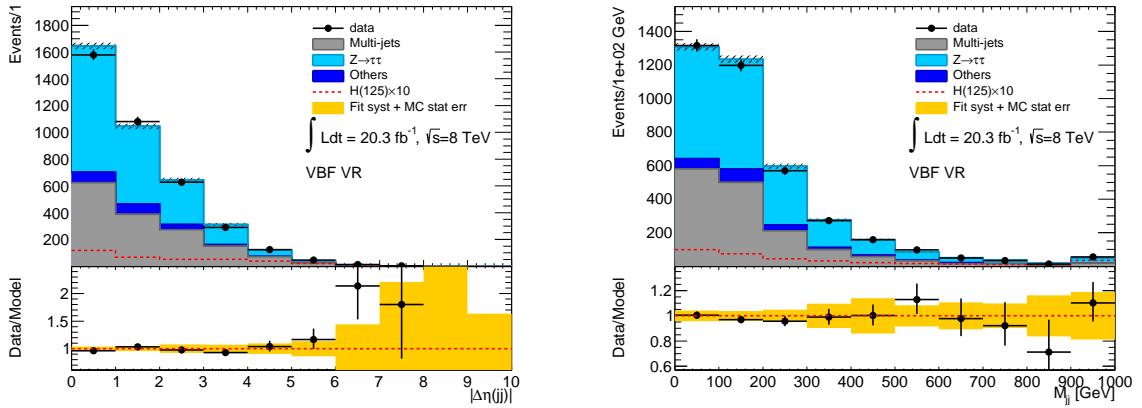
Figure A.10: $|\Delta\eta_{jj}|$ (left) and the M_{jj} invariant mass (right) of the two tagging jets in the VBF VR in $\sqrt{s} = 8$ TeV data. The hashed area on each plot indicates the statistical uncertainty of the background model. The lower panel shows the ratio of the observed data to the total predicted background. The yellow band includes the statistical uncertainty of the background model, the multi-jet shape uncertainty and the multi-jet and $Z \to \tau\tau$ event yield uncertainties. The first and the last bins include the under- and over-flow events, respectively.

(a) Preselection

(b) VBF VR

(c) Boosted VR
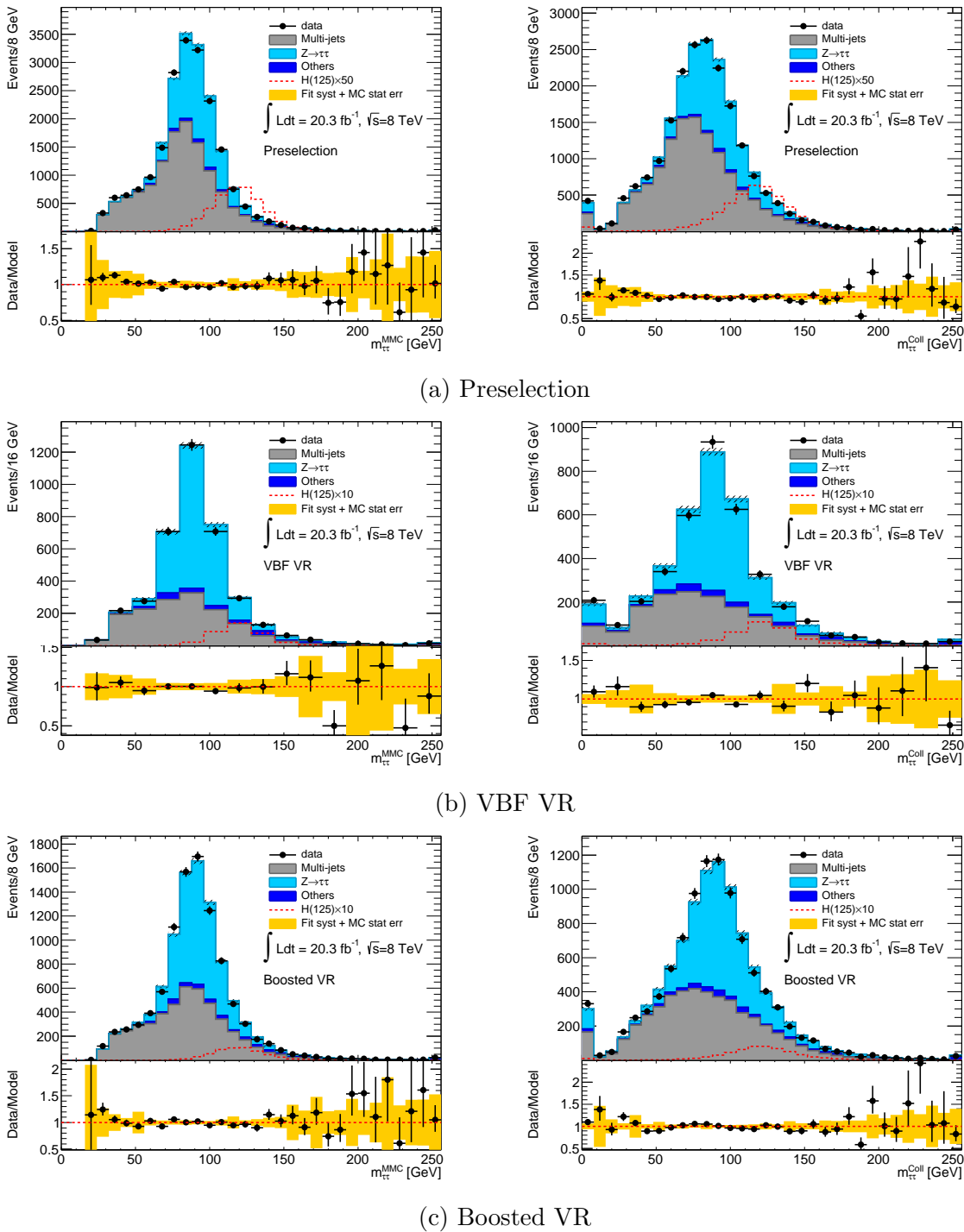
Figure A.11: $m_{\tau\tau}^{\text{MMC}}$ (left) and $m_{\tau\tau}^{\text{Coll}}$ (right) invariant masses of the two τ-jets in $\sqrt{s} = 8$ TeV data. The hashed area on each plot indicates the statistical uncertainty of the background model. The lower panel shows the ratio of the observed data to the total predicted background. The yellow band includes the statistical uncertainty of the background model, the multi-jet shape uncertainty and the multi-jet and $Z \rightarrow \tau\tau$ event yield uncertainties. The first and the last bins include the under- and over-flow events, respectively.

Appendix B

$m_{\tau\tau}^{\mathrm{MMC}}$ Fit

This section gives more details on the statistical model introduced in Section 4.8.

Table B.1 lists the nine orthogonal event categories with the associated samples. The labels "2011" and "2012" refer to the $\sqrt{s} = 7$ and 8 TeV data, respectively. The signal events produced via VH mode are not included in the VBF event categories because of their negligible contributions. In addition to the samples for the $Z \to \tau\tau$ and multi-jet backgrounds, the small contribution from the remaining backgrounds is treated as a single sample called "Others".

Table B.1: Event categories and samples implemented in the statistical model of the search.

Category	Samples						
VBF High-$p_{\mathrm{T}}^{\tau\tau}$ 2011	ggF	VBF			Z	Multi-jet	Others
VBF Low-$p_{\mathrm{T}}^{\tau\tau}$ 2011	ggF	VBF			Z	Multi-jet	Others
VBF High-$p_{\mathrm{T}}^{\tau\tau}$ 2012	ggF	VBF			Z	Multi-jet	Others
VBF Low-$p_{\mathrm{T}}^{\tau\tau}$ Tight 2012	ggF	VBF			Z	Multi-jet	Others
VBF Low-$p_{\mathrm{T}}^{\tau\tau}$ Loose 2012	ggF	VBF			Z	Multi-jet	Others
Boosted High-$p_{\mathrm{T}}^{\tau\tau}$ 2011	ggF	VBF	WH	ZH	Z	Multi-jet	Others
Boosted Low-$p_{\mathrm{T}}^{\tau\tau}$ 2011	ggF	VBF	WH	ZH	Z	Multi-jet	Others
Boosted High-$p_{\mathrm{T}}^{\tau\tau}$ 2012	ggF	VBF	WH	ZH	Z	Multi-jet	Others
Boosted Low-$p_{\mathrm{T}}^{\tau\tau}$ 2012	ggF	VBF	WH	ZH	Z	Multi-jet	Others

The observable used as discriminating variable in the statistical model is the binned $m_{\tau\tau}^{\mathrm{MMC}}$ distribution. The bin sizes used in each event category are listed in Table B.2. They are chosen so that there are no bins without any expected background events and that, whenever possible, there are predicted events from both $Z \to \tau\tau$ and multi-jet events.

The expected event yields from signal and Others background processes are determined by simulation and normalised to the measured integrated luminosity. The total yields of the $Z \to \tau\tau$ and multi-jet contributions are determined by the fit itself. This is done by using four unconstrained independent normalisation factors (NF), two for the $Z \to \tau\tau$

Table B.2: $m_{\tau\tau}^{\mathrm{MMC}}$ bin upper edges in each event category.

Category	$m_{\tau\tau}^{\mathrm{MMC}}$ Bin upper edges [GeV]
VBF High-$p_{\mathrm{T}}^{\tau\tau}$ 2011	[64,80,92,104,116,132,∞]
VBF Low-$p_{\mathrm{T}}^{\tau\tau}$ 2011	[64,80,92,104,116,132,152,∞]
VBF High-$p_{\mathrm{T}}^{\tau\tau}$ 2012	[64,80,92,104,116,132,176,∞]
VBF Low-$p_{\mathrm{T}}^{\tau\tau}$ Tight 2012	[64,80,92,104,116,132,152,176,∞]
VBF Low-$p_{\mathrm{T}}^{\tau\tau}$ Loose 2012	[64,80,92,104,116,132,152,176,∞]
Boosted High-$p_{\mathrm{T}}^{\tau\tau}$ 2011	[64,72,80,84,88,92,96,100,104,108,112,116,120,124,128,132,140,160,∞]
Boosted Low-$p_{\mathrm{T}}^{\tau\tau}$ 2011	[64,80,84,88,92,96,100,104,108,112,116,120,124,128,132,136,140,156,200,∞]
Boosted High-$p_{\mathrm{T}}^{\tau\tau}$ 2012	[64,72,80,84,88,92,96,100,104,108,112,116,120,124,128,132,140,156,176,∞]
Boosted Low-$p_{\mathrm{T}}^{\tau\tau}$ 2012	[64,80,84,88,92,96,100,104,108,112,116,120,124,128,132,136,140,148,156,176,∞]

contributions in $\sqrt{s} = 7$ and 8 TeV data and two for the multi-jet contributions in $\sqrt{s} = 7$ and 8 TeV data. These normalisation factors scale the yields of the corresponding background contributions in all event categories relative to the yields predicted in the preselected events by the fit to the observed $m_{\tau\tau}^{\mathrm{MMC}}$ distribution.

In the following, a detailed description of the nuisance parameters (NP) related to systematic uncertainties is given. Tables B.3 and B.4 show their application in each event category and in each sample. They can impact on the yield (Y) of a sample in a given category or on its $m_{\tau\tau}^{\mathrm{MMC}}$ shape (S) distribution or on both (YS). When either "2011" or "2012" is specified, the NP is applied only on the $\sqrt{s} = 7$ or 8 TeV event categories, respectively.

TAU ID 2011 yield NP for the uncertainty of the measurement of the τ-jet identification efficiency in 7 TeV data. It is not implemented in the Z sample because it affects only the total yield.

TAU IDsyst 2012 yield NP for the systematic uncertainty of the measurement of the τ-jet identification efficiency in 8 TeV data. The statistical uncertainty of the measurement is treated independently (see TAU SF 2012). This systematic uncertainty is not implemented in the Z sample because it affects only the total yield.

TAU TRIGGER 2011 yield NP for the uncertainty of the measurement of the τ-jet trigger efficiency in 7 TeV data. It is not implemented in the Z sample because it affects only the total yield.

TAU TRIGGERsyst 2012 yield NP for the systematic uncertainty of the measurement of the τ-jet trigger efficiency in 8 TeV data. The statistical uncertainty of the measurement is treated independently (see TAU SF 2012). This systematic uncertainty is not implemented in the Z sample because it affects only the total yield.

TAU FR 2011 yield NP for the uncertainty of the measurement of the τ-jet trigger and identification efficiency for QCD jets mis-identified as τ-jets in 7 TeV data. In 8

TeV data only the statistical component of this uncertainty is relevant and it is added to the statistical uncertainty of the $m_{\tau\tau}^{\mathrm{MMC}}$ distribution (see TAU SF 2012). This systematic uncertainty is implemented only in the Others sample.

TAU SF 2012 yield and shape NP for the statistical uncertainties of the measurements of the τ-jet trigger and identification efficiencies in 8 TeV data. It is applied only in the signal samples. In the background samples these statistical uncertainties are added to the statistical uncertainty of the $m_{\tau\tau}^{\mathrm{MMC}}$ distributions. The same is not done in the signal samples because the statistical uncertainty of their $m_{\tau\tau}^{\mathrm{MMC}}$ distributions is not taken into account in the fit. Shapes uncertainties are shown in Fig. B.1.

TAU TES TRUE INSITU/MODELING/SP 2012 yield and shape NPs for the uncertainties of the τ-jet energy calibration of true τ-jets in 8 TeV data and simulation. These three NPs correspond to the *in situ*, Modelling and Single Particle Response uncertainties described in Section 4.7.1. Uncertainties on the event yields are summarised in Table 4.12 and shapes uncertainties are shown in Figs. B.2 and following.

TAU TES TRUE 2011 yield and shape NP for the uncertainty of the τ-jet energy calibration of true τ-jets in 7 TeV data and simulation.

TAU TES FAKE 2011/2012 yield NP for the uncertainty of the τ-jet energy calibration of simulated QCD jets mis-identified as τ-jets. It is implemented in the Others sample only as overall systematic because the shape variations are not significant with respect to the large statistical uncertainty.

JES 2011/2012 Detector1/Modelling1 yield and shape NPs for two components of the *in situ* uncertainty of the jet energy calibration. They are treated independently in the 7 and 8 TeV data because they are based on different measurements. They are applied on signal, but not on the Others sample because on the latter they have statistically negligible impacts. Shapes uncertainties are shown in Fig. B.5.

JES Eta Modelling/StatMethod yield and shape NPs of the η inter-calibration uncertainties of the jet energy calibration. They are applied on signal, but not on the Others sample because on the latter they have statistically negligible impacts. Shapes uncertainties are shown in Fig. B.7.

JES PileRho tautauhh qq/gg yield NPs for the uncertainties of the correction for pileup effects on the jet energy calibration. It is treated independently for qq- (VBF and VH) and gg-initiated (ggF) processes. They are applied on signal, but not on the Others sample because on the latter they have statistically negligible impacts. The shape uncertainties are not significant compared to other JES shape uncertainties.

JES FlavComp tautauhh qq/gg yield NPs for the flavour composition uncertainty of the jet energy calibration. It is treated independently for qq- (VBF and VH) and gg-initiated (ggF) processes. They are applied on signal, but not on the Others sample because on the latter they have statistically negligible impacts. The shape uncertainties are not significant compared to other JES shape uncertainties.

JES FlavResp yield and shape NP for the uncertainties of the flavour response in the jet energy calibration. It is applied on signal, but not on the Others sample because on the latter it has a statistically negligible impact. Shapes uncertainties are shown in Fig. B.6.

JER yield NP for the uncertainty of the jet energy resolution. It is applied on signal, but not on the Others sample because on the latter it has a statistically negligible impact. The shape impact is not significant compared to other JES shape uncertainties.

QCDscale qqH yield NP for the uncertainty of the VBF cross section.

QCDscale ggH2in, QCDscale ggH1in yield NPs for the uncertainties of the ggF cross sections in different jet multiplicities (Table 4.16).

QCDscale VH yield NP for the uncertainty of the VH cross section.

QCDscale ttbar/V yield NPs for the uncertainties of the cross sections of the background processes included in the Others sample. The value of this systematic uncertainty in each category is assigned depending on whether the dominant contribution comes from $t\bar{t}$ or $W + j$ events.

pdf gg/qq yield NPs for the PDF uncertainties of the cross sections of the background processes included in the Others sample. The value of this systematic uncertainty in each category is assigned depending on whether the dominant contribution comes from $t\bar{t}$ or $W + j$ events.

pdf Higgs qq/gg yield NPs for the PDF uncertainty of the cross sections of signal processes. It is treated independently for qq- (VBF and VH) and gg-initiated (ggF) processes.

UE qq/gg yield NPs for the UE uncertainty of the cross sections of signal processes. It is treated independently for qq- (VBF and VH) and gg-initiated (ggF) processes.

BR tautau yield NP for the uncertainty of the $H \to \tau\tau$ decay branching ratio.

Gen Qmass ggH yield NP for the modelling of the $p_{\mathrm{T}}^{\mathrm{Higgs}}$ in the simulation of ggF events.

QCD BinMigration 2011/2012 yield NP for the uncertainty of the relative yields of the multi-jet contributions across event categories.

QCD mass 2011/2012 shape NP for the uncertainty of the modelling of the $m_{\tau\tau}^{\mathrm{MMC}}$ distribution of the multi-jet sample (Fig. 4.14).

EMB MFS 2011/2012 yield and shape NP for the uncertainty of the embedding procedure. Shapes uncertainties are shown in Fig. B.8.

EMB ISO 2011/2012 yield and shape NP for the uncertainty of the $Z \to \mu\mu$ event selection for the embedded sample. Shapes uncertainties are shown in Fig. B.8.

LUMI 2011/2012 yield NP for the uncertainty of the measurement of the integrated luminosity. It is applied on signal and Others samples.

Table B.3: Implementation of the nuisance parameters in the VBF categories. "Y" stands for uncertainties on the yield and "S" on the shape.

NP	VBF High-$p_T^{\tau\tau}$ 2011					VBF Low-$p_T^{\tau\tau}$ 2011					VBF High-$p_T^{\tau\tau}$ 2012					Low-$p_T^{\tau\tau}$ Tight 2012					VBF Low-$p_T^{\tau\tau}$ Loose 2012				
	Others	Multi-jet	VBFH	Z	ggFH	Others	Multi-jet	VBFH	Z	ggFH	Others	Multi-jet	VBFH	Z	ggFH	Others	Multi-jet	VBFH	Z	ggFH	Others	Multi-jet	VBFH	Z	ggFH
TAU ID 2011	Y		Y		Y	Y		Y		Y															
TAU TRIGGER 2011	Y		Y		Y	Y		Y		Y															
TAU FR 2011	Y					Y																			
TAU TES FAKE 2011	Y					Y																			
LUMI 2011	Y		Y		Y	Y		Y		Y															
QCDscale ttbar	Y					Y					Y					Y					Y				
pdf gg	Y					Y					Y					Y					Y				
QCD BinMigration 2011		Y					Y																		
JES 2011 Detector1			Y		Y			Y		Y															
JES 2011 Modelling1			Y		Y			Y		Y															
JES Eta Modelling			Y		Y			Y		Y			YS		Y			YS		Y			YS		YS
JES FlavComp tautauhh qq			Y					Y					Y					Y					Y		
JES FlavResp			Y		Y			Y		Y			YS		Y			YS		Y			YS		YS
JER			Y		Y			Y		Y			Y		Y			Y		Y			Y		Y
UE qq			Y					Y					Y					Y					Y		
BR tautau			Y		Y			Y		Y			Y		Y			Y		Y			Y		Y
pdf Higgs qq			Y					Y					Y					Y					Y		
QCDscale qqH			Y					Y					Y					Y					Y		
EMB ISO 2011		Y							Y																
JES FlavComp tautauhh gg					Y					Y					Y					Y					Y
TAU TES TRUE 2011			Y	Y	Y			YS	YS	Y															
UE gg					Y					Y					Y					Y					Y
pdf Higgs gg					Y					Y					Y					Y					Y
QCDscale ggH2in					Y					Y					Y					Y					Y
Gen Qmass ggH					Y					Y					Y					Y					Y
TAU TES TRUE INSITU 2012											Y		YS	Y	Y	Y		YS	YS	Y	Y		YS	YS	YS
TAU TES TRUE MODELING 2012											Y		YS	Y	Y	Y		YS	Y	Y	Y		YS	Y	YS
TAU TES TRUE SP 2012											Y		YS	Y	Y	Y		YS	Y	Y	Y		YS	Y	YS
TAU TES FAKE 2012											Y					Y					Y				
TAU IDsyst 2012											Y		Y		Y	Y		Y		Y	Y		Y		Y
LUMI 2012											Y		Y		Y	Y		Y		Y	Y		Y		Y
QCD BinMigration 2012												Y					Y					Y			
JES 2012 Eta StatMethod													Y		Y			Y		Y			Y		Y
JES 2012 PileRho tautauhh qq													Y					Y					Y		
TAU TRIGGERsyst 2012													Y		Y			Y		Y			Y		Y
JES 2012 PileRho tautauhh gg															Y					Y					Y
EMB MFS 2011		Y							YS																
EMB MFS 2012												Y					Y							YS	
EMB ISO 2012												Y					Y							YS	
JES 2012 Detector1													YS		Y			YS		Y			YS		YS
JES 2012 Modelling1													YS		Y			YS		Y			YS		YS
TAU SF 2012													YS		Y			YS		Y			YS		YS

Table B.4: Implementation of the nuisance parameters in the Boosted categories. "Y" stands for uncertainties on the yield and "S" on the shape.

NP	Boosted High-$p_T^{\tau\tau}$ 2011							Boosted Low-$p_T^{\tau\tau}$ 2011							Boosted High-$p_T^{\tau\tau}$ 2012							Boosted Low-$p_T^{\tau\tau}$ 2012						
	Others	Multi-jet	VBFH	WH	ZH	Z	ggFH	Others	Multi-jet	VBFH	WH	ZH	Z	ggFH	Others	Multi-jet	VBFH	WH	ZH	Z	ggFH	Others	Multi-jet	VBFH	WH	ZH	Z	ggFH
TAU TES FAKE 2011	Y							Y																				
TAU ID 2011	Y		Y	Y	Y		Y	Y		Y	Y	Y		Y														
TAU TRIGGER 2011	Y		Y	Y	Y		Y	Y		Y	Y	Y		Y														
TAU FR 2011	Y							Y																				
LUMI 2011	Y		Y	Y	Y		Y	Y		Y	Y	Y		Y														
QCDscale V	Y							Y							Y							Y						
pdf qq	Y							Y							Y							Y						
QCD BinMigration 2011		Y							Y																			
JES 2011 Detector1			Y	Y	Y		Y			Y	Y	Y		Y														
JES 2011 Modelling1			Y	Y	Y		Y			Y	Y	Y		Y														
JES Eta Modelling			Y	Y	Y		Y			Y	Y	Y		Y			YS	YS	YS		YS			YS	YS	YS		YS
JES FlavComp tautauhh qq			Y	Y	Y					Y	Y	Y					Y	Y	Y					Y	Y	Y		
JES FlavResp			Y	Y	Y		Y			Y	Y	Y		Y			YS	YS	YS		YS			YS	YS	YS		YS
JER			Y	Y	Y		Y			Y	Y	Y		Y			Y	Y	Y		Y			Y	Y	Y		Y
BR tautau			Y	Y	Y		Y			Y	Y	Y		Y			Y	Y	Y		Y			Y	Y	Y		Y
pdf Higgs qq			Y	Y	Y					Y	Y	Y					Y	Y	Y					Y	Y	Y		
QCDscale qqH			Y							Y							Y							Y				
TAU TES TRUE 2011			YS	Y	Y	YS	YS			YS	Y	Y	YS	YS														
QCDscale VH				Y	Y						Y	Y						Y	Y						Y	Y		
EMB MFS 2011						Y							Y															
EMB ISO 2011						Y							Y															
JES FlavComp tautauhh gg							Y							Y							Y							Y
pdf Higgs gg							Y							Y							Y							Y
Gen Qmass ggH							Y							Y							Y							Y
QCDscale ggH2in							Y							Y							Y							Y
QCDscale ggH1in							Y							Y							Y							Y
TAU TES TRUE INSITU 2012															Y		YS	YS	YS	YS	YS	Y		YS	YS	YS	YS	YS
TAU TES TRUE MODELING 2012															Y		YS	YS	YS	YS	YS	Y		YS	YS	YS	YS	YS
TAU TES TRUE SP 2012															Y		YS	YS	YS	YS	YS	Y		YS	YS	YS	YS	YS
TAU TES FAKE 2012															Y							Y						
TAU IDsyst 2012															Y		Y	Y	Y		Y	Y		Y	Y	Y		Y
LUMI 2012															Y		Y	Y	Y		Y	Y		Y	Y	Y		Y
QCD BinMigration 2012																Y							Y					
JES 2012 Eta StatMethod																	Y	Y	Y		Y			Y	Y	Y		Y
JES 2012 PileRho tautauhh qq																	Y	Y	Y					Y	Y	Y		
TAU TRIGGERsyst 2012																	Y	Y	Y		Y			Y	Y	Y		Y
JES 2012 PileRho tautauhh gg																					Y							Y
QCD mass 2011									S																			
QCD mass 2012																							S					
JES 2012 Detector1																	YS	YS	YS		YS			YS	YS	YS		YS
JES 2012 Modelling1																	YS	YS	YS		YS			YS	YS	YS		YS
TAU SF 2012																	YS	YS	YS		YS			YS	YS	YS		YS
EMB MFS 2012																				YS							YS	
EMB ISO 2012																				YS							YS	

(a) VBF sample in VBF High-$p_{\mathrm{T}}^{\tau\tau}$

(b) ggF sample in Boosted High-$p_{\mathrm{T}}^{\tau\tau}$
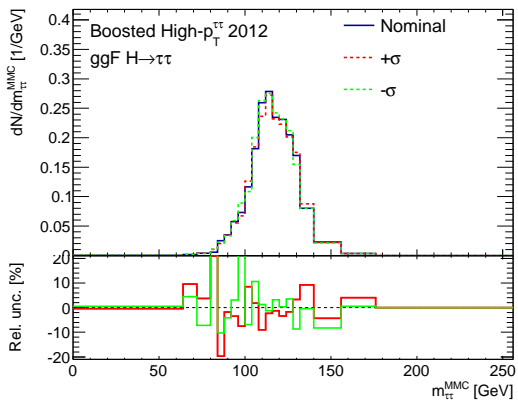
Figure B.1: $\pm 1\sigma$ variations of the $m_{\tau\tau}^{\mathrm{MMC}}$ distribution for the TAU SF NP in 8 TeV event categories for $m_H = 125$ GeV signal samples.

A series of tests are performed before measuring the observed signal strength in order to assess the quality of the fit.

The first test is about the stability of the fit. Scans of the negative log-likelihood (NLL) as a function of the most important nuisance parameters affecting the background prediction are shown in Figs. B.9, B.10, B.11 and B.12. On the x-axis is the variation of the NP in units of standard deviations of the pre-fit uncertainty. On the y-axis is the difference ΔNLL of the NLL at a given value of the NP with respect to the NLL global minimum, which corresponds to the best-fit value of the NP. The difference between the input (zero) and the best-fit value is the so-called "pull" and the range where ΔNLL is below unity is the 1σ uncertainty of the fitted NP in units of the input uncertainty. An error smaller than unity means that the NP is constrained by the fit to the observed data. In each plot, the NLL is shown for the fit of all event categories and also of the VBF and Boosted categories alone. These plots are aimed to detect unexpected double minima or non-parabolic behaviours of the NLL which might hint to instabilities of the fit. None of these are observed. Fig. B.12 shows the NLL scans for the normalisation factors of the Z and multi-jet backgrounds, whose pre-fit values is unity. No double minima or hints for instabilities are seen here either.

The second test is on the pulls of each nuisance parameter, i.e. the difference between the pre-fit and the post-fit values, the uncertainties on such pulls and the impact of them on the measurement of the signal strength, $\Delta\hat{\mu}/\Delta\hat{\mu}_{\mathrm{tot}}$. The pulls measured in the combined fit to the observed data are shown in Fig. B.13. The nuisance parameters are ranked by their impact on the total uncertainty on the signal strength and only the ones with $\Delta\hat{\mu}/\Delta\hat{\mu}_{\mathrm{tot}} > 0.04$ are shown. In contrast to the NPs, whose pre-fit value is zero, the normalisation factors are centered at one because their pre-fit values is unity. Negative pulls at about -1σ are observed in the EMB MFS, TES TRUE SP 2012 and TES TRUE MODELING 2012 nuisance parameters. Such uncertainties act on the $Z \to \tau\tau$ background in opposite directions, namely the pull in the EMB MFS gives a positive shift to higher masses, while the pulls in the TES NPs produce a negative shift. Given
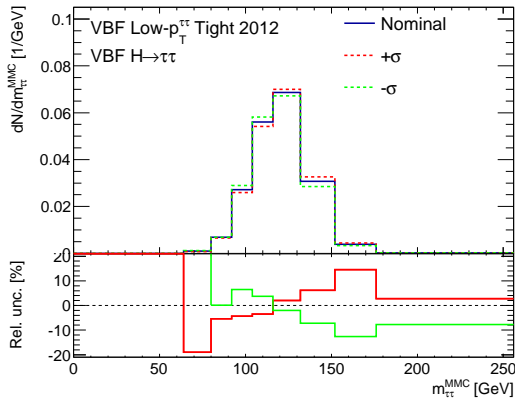
(a) VBF sample in VBF Low-$p_\mathrm{T}^{\tau\tau}$ Tight
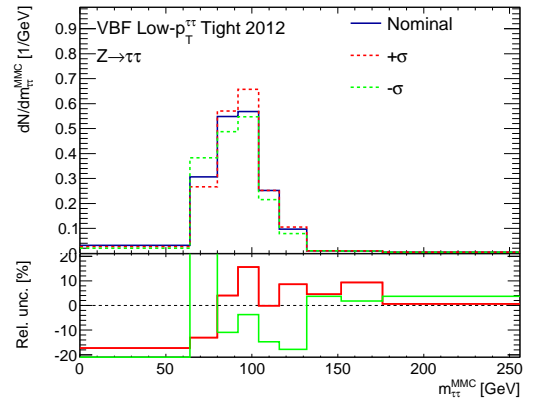
(b) Z sample in VBF Low-$p_\mathrm{T}^{\tau\tau}$ Tight
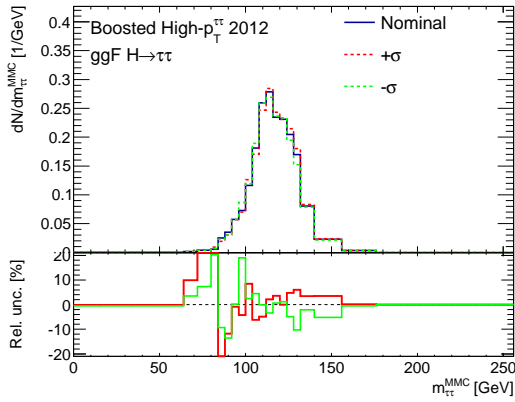
(c) ggF sample in Boosted High-$p_\mathrm{T}^{\tau\tau}$
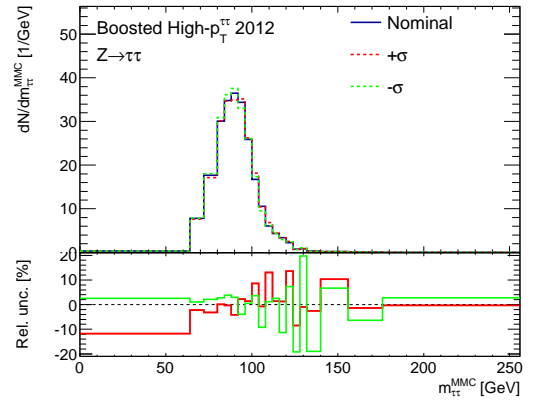
(d) Z sample in Boosted High-$p_\mathrm{T}^{\tau\tau}$

Figure B.2: $\pm 1\sigma$ variations of the $m_{\tau\tau}^{\mathrm{MMC}}$ distribution for the *in situ* TES NP in 8 TeV event categories for Z and signal samples.
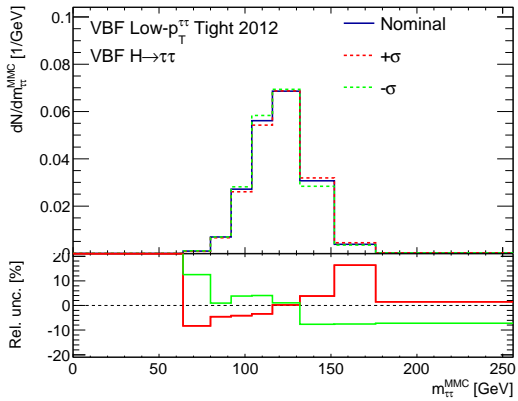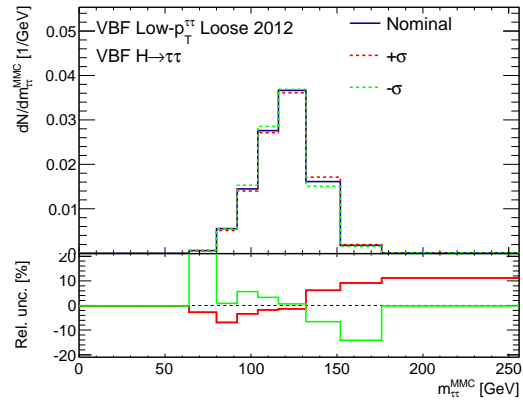
(a) VBF sample in VBF Low-$p_{\mathrm{T}}^{\tau\tau}$ Tight

(b) VBF sample in VBF Low-$p_{\mathrm{T}}^{\tau\tau}$ Loose

(c) ggF sample in Boosted High-$p_{\mathrm{T}}^{\tau\tau}$
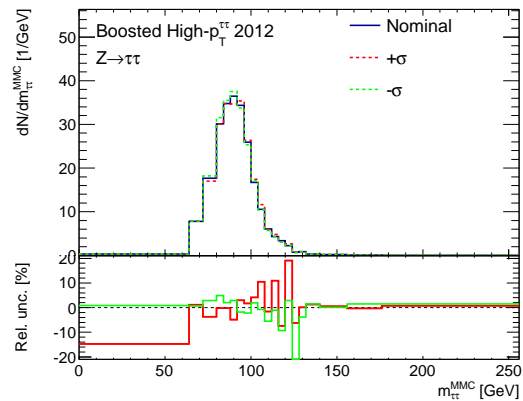
(d) Z sample in Boosted High-$p_{\mathrm{T}}^{\tau\tau}$

Figure B.3: $\pm 1\sigma$ variations of the $m_{\tau\tau}^{\mathrm{MMC}}$ distribution for the Modelling TES NP in 8 TeV event categories for Z and signal samples.
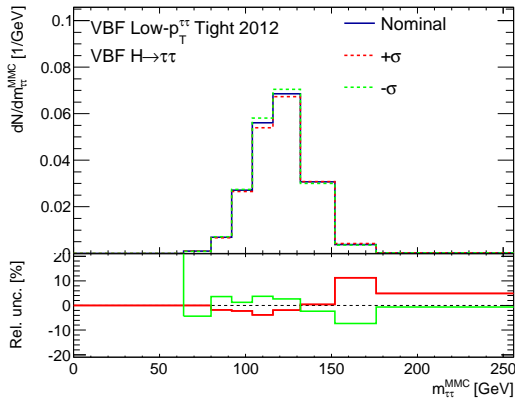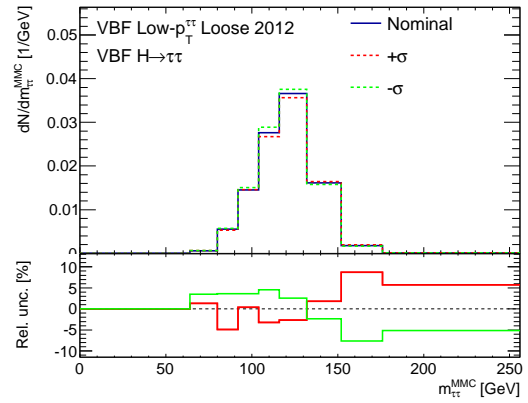
(a) VBF sample in VBF Low-$p_T^{\tau\tau}$ Tight

(b) VBF sample in VBF Low-$p_T^{\tau\tau}$ Loose
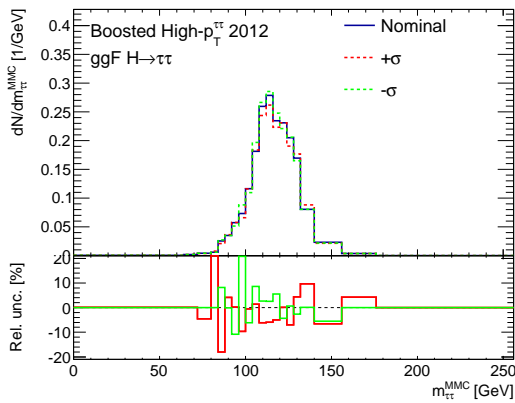
(c) ggF sample in Boosted High-$p_T^{\tau\tau}$
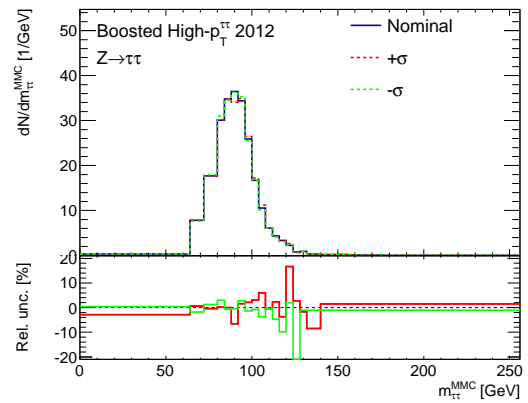
(d) Z sample in Boosted High-$p_T^{\tau\tau}$

Figure B.4: $\pm 1\sigma$ variations of the $m_{\tau\tau}^{\mathrm{MMC}}$ distribution for the Single Particle Response TES NP in 8 TeV event categories for Z and signal samples.
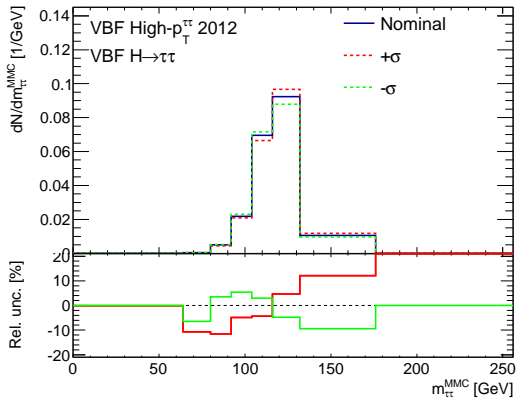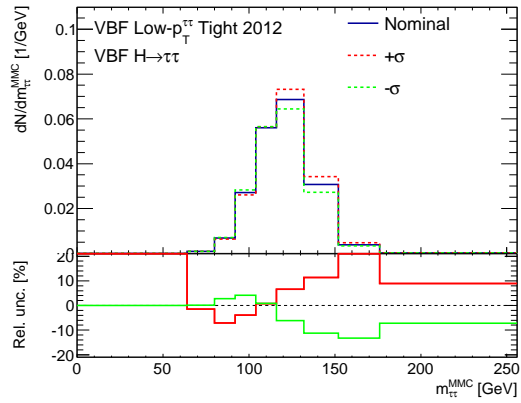
(a) VBF sample in VBF High-$p_{\mathrm{T}}^{\tau\tau}$

(b) VBF sample in VBF Low-$p_{\mathrm{T}}^{\tau\tau}$ Tight
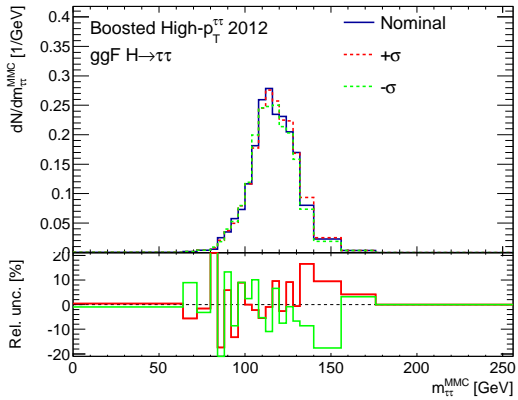
(c) ggF sample in Boosted High-$p_{\mathrm{T}}^{\tau\tau}$
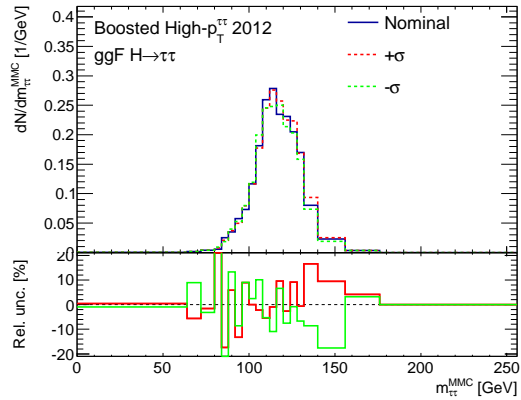
(d) ggF sample in Boosted High-$p_{\mathrm{T}}^{\tau\tau}$

Figure B.5: $\pm 1\sigma$ variations of the $m_{\tau\tau}^{\mathrm{MMC}}$ distribution for the JES 2012 Modelling1 NP in 8 TeV event categories for $m_H = 125$ GeV signal samples.
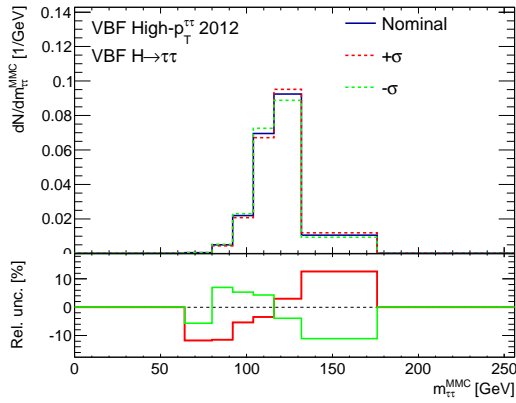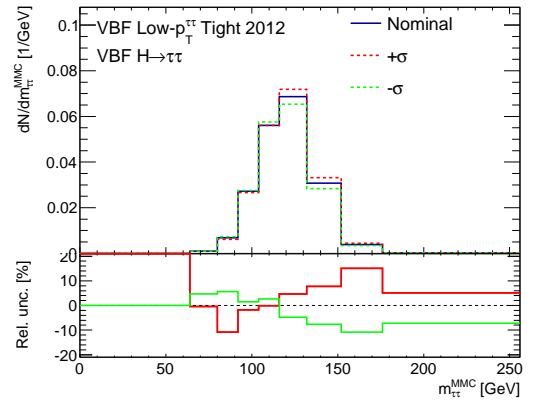
(a) VBF sample in VBF High-$p_{\mathrm{T}}^{\tau\tau}$

(b) VBF sample in VBF Low-$p_{\mathrm{T}}^{\tau\tau}$ Tight
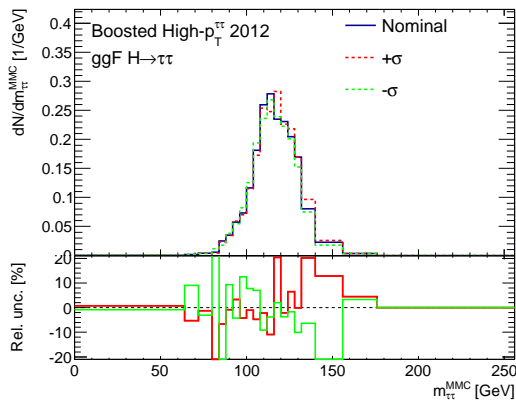
(c) ggF sample in Boosted High-$p_{\mathrm{T}}^{\tau\tau}$
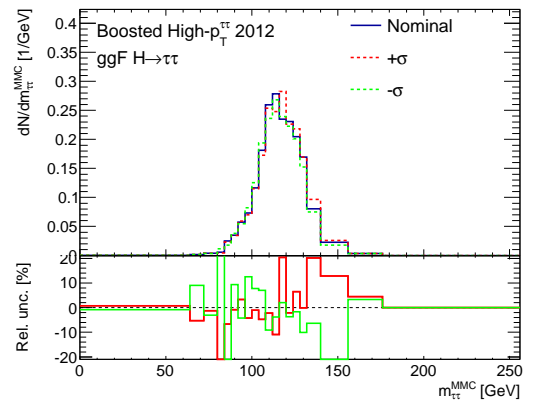
(d) ggF sample in Boosted High-$p_{\mathrm{T}}^{\tau\tau}$

Figure B.6: $\pm 1\sigma$ variations of the $m_{\tau\tau}^{\mathrm{MMC}}$ distribution for the JES FlavResp NP in 8 TeV event categories for $m_H = 125$ GeV signal samples.
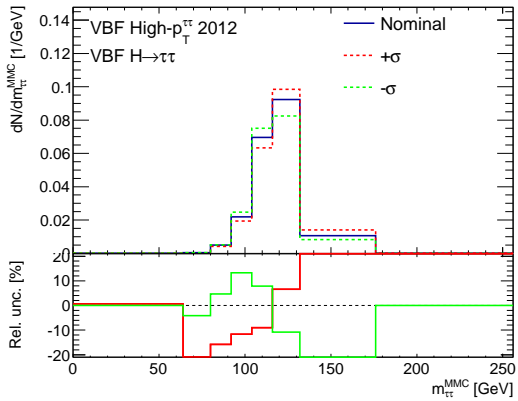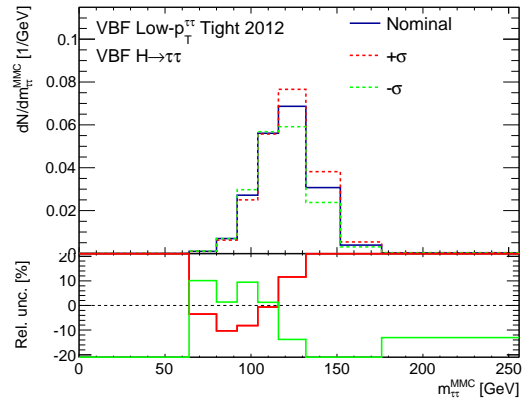
(a) VBF sample in VBF High-$p_{\mathrm{T}}^{\tau\tau}$

(b) VBF sample in VBF Low-$p_{\mathrm{T}}^{\tau\tau}$ Tight
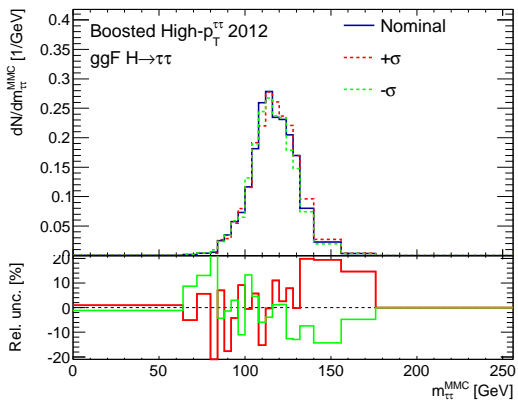
(c) ggF sample in Boosted High-$p_{\mathrm{T}}^{\tau\tau}$
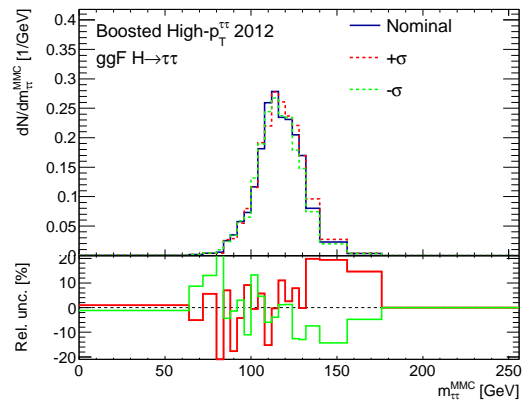
(d) ggF sample in Boosted High-$p_{\mathrm{T}}^{\tau\tau}$

Figure B.7: $\pm 1\sigma$ variations of the $m_{\tau\tau}^{\mathrm{MMC}}$ distribution for the JES Eta Modelling NP in 8 TeV event categories for $m_H = 125$ GeV signal samples.
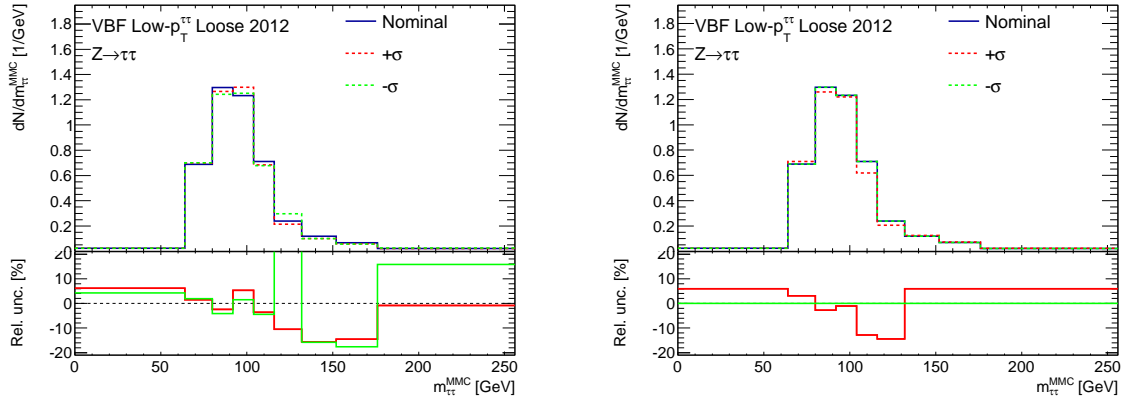
(a) VBF Low-$p_T^{\tau\tau}$ Loose
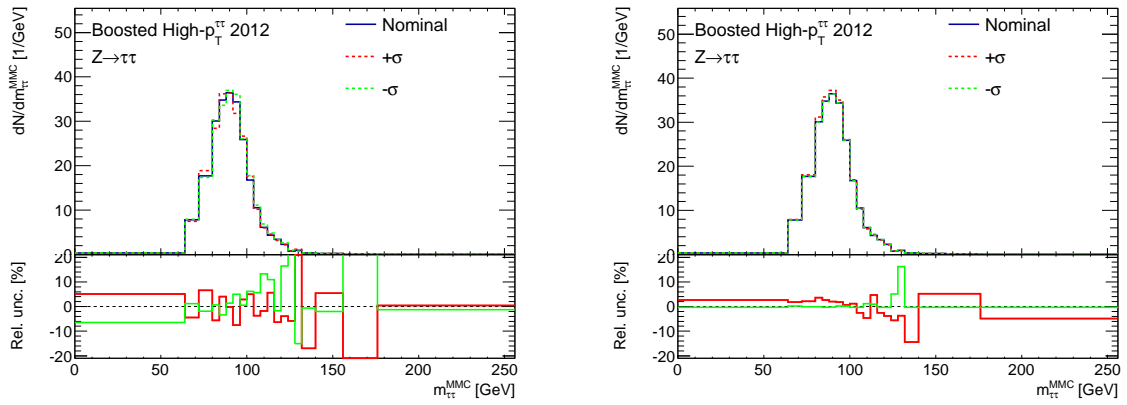
(b) Boosted High-$p_T^{\tau\tau}$
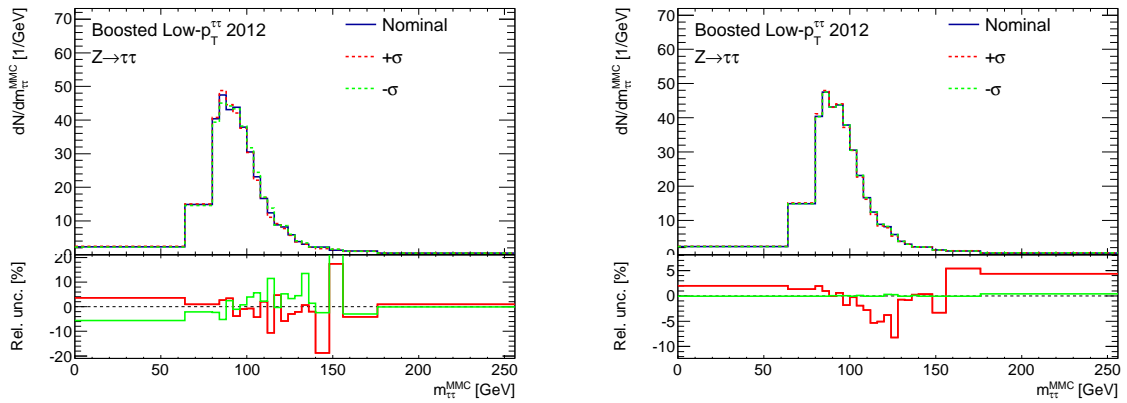
(c) Boosted Low-$p_T^{\tau\tau}$

Figure B.8: $\pm 1\sigma$ variations of the $m_{\tau\tau}^{\mathrm{MMC}}$ distribution for the EMB MFS 2012 (left) and EMB ISO 2012 (right) NPs on the Z sample in the 8 TeV event categories.
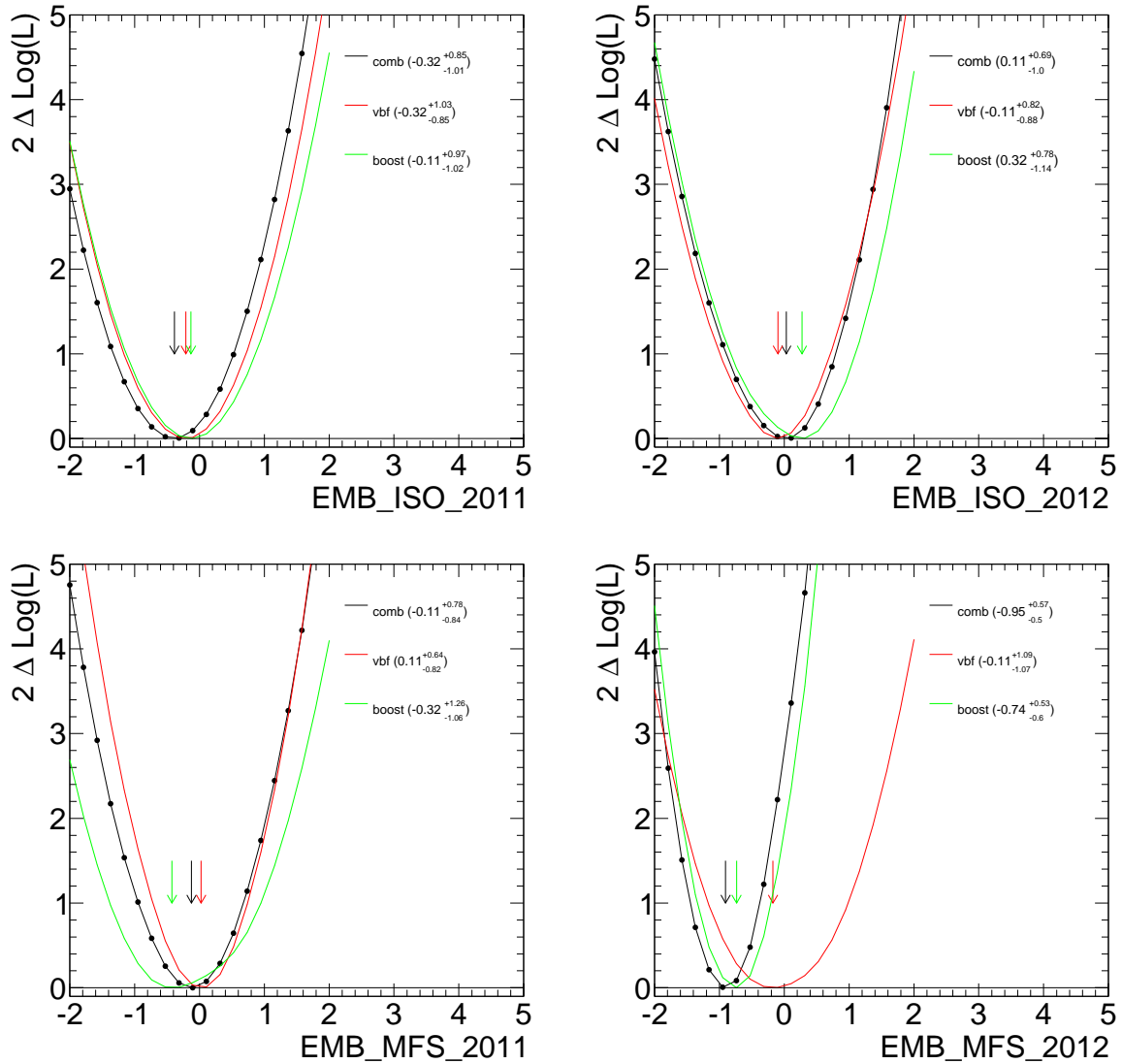
Figure B.9: NLL scans for the EMB ISO (top) and EMB MFS (bottom) NPs impacting on the Z background in 2011 (left) and 2012 (right). Scans are done in the fit of all event categories (black). Scans performed in the VBF and Boosted categories alone are also shown. Arrows indicate the minima of the NLL distributions.
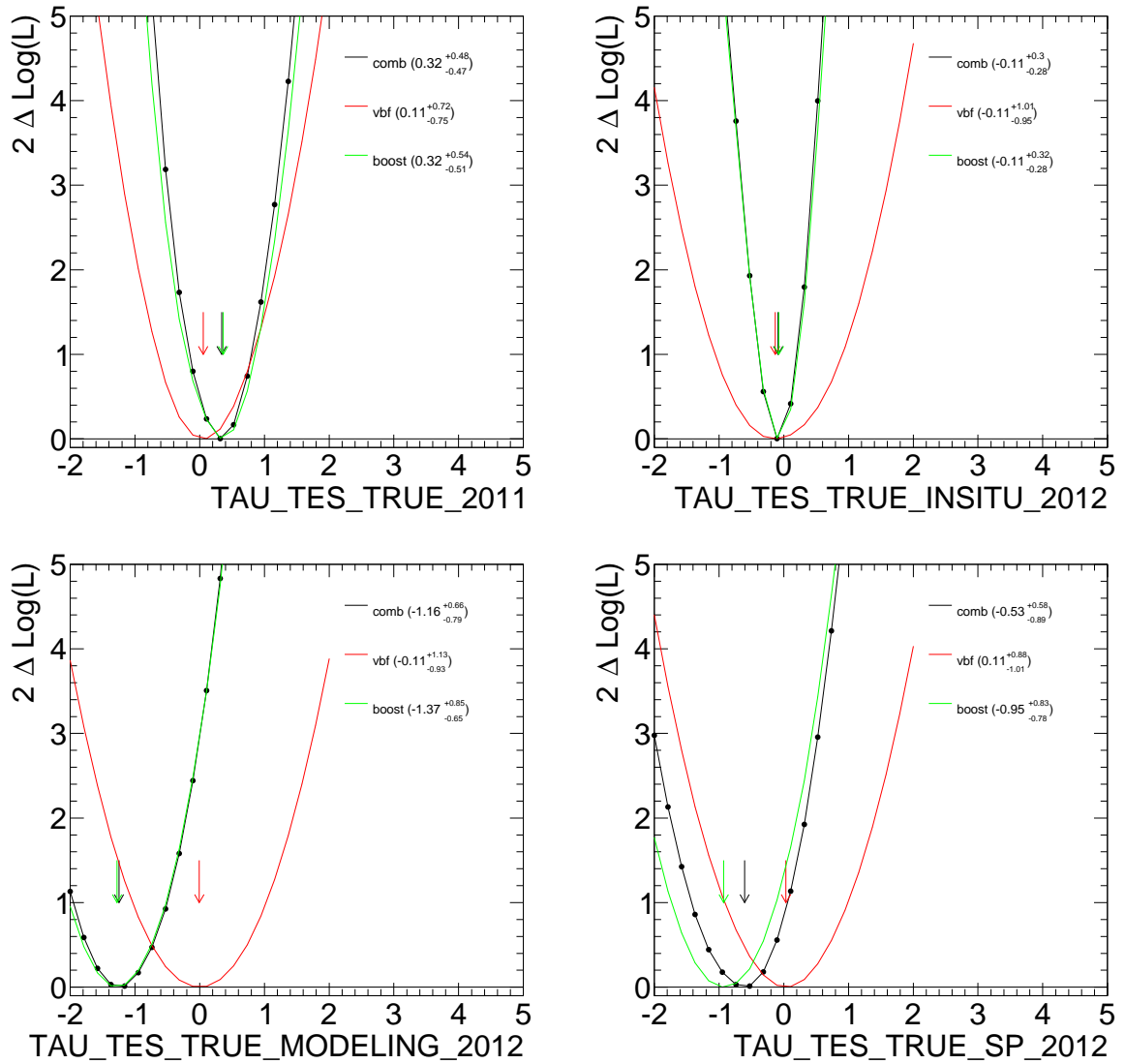
Figure B.10: NLL scans for the TES NPs. Scans are done in the fit of all event categories (black). Scans performed in the VBF and Boosted categories alone are also shown. Arrows indicate the minima of the NLL distributions.
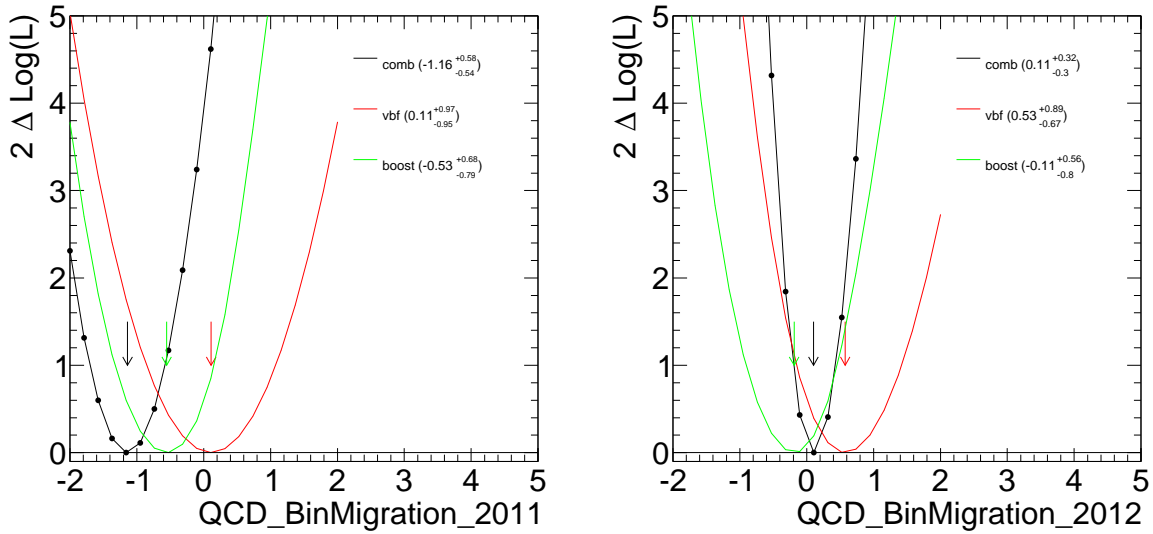
Figure B.11: NLL scans for the NPs impacting on the relative yields of the multi-jet background in 2011 (left) and 2012 (right). Scans are done in the fit of all event categories (black). Scans performed in the VBF and Boosted categories alone are also shown. Arrows indicate the minima of the NLL distributions.

that no instabilities are seen in the NLL scans and that such pulls are not statistically significant, these pulls are considered not a source of concern. The nuisance parameters with the biggest impacts on $\Delta\hat{\mu}_{\mathrm{tot}}$ are the expected ones. In addition to the ones previously mentioned, "stat hh vbf tight lowdr 2012 bin 5" and "stat hh boost tight 2012 bin 14" are the nuisance parameters for the bin statistical uncertainties with highest rankings because they are related to the most sensitive bins, namely $116 < m_{\tau\tau}^{\mathrm{MMC}} < 132$ GeV in the VBF High-$p_{\mathrm{T}}^{\tau\tau}$ 2012 category and $128 < m_{\tau\tau}^{\mathrm{MMC}} < 132$ GeV in the Boosted High-$p_{\mathrm{T}}^{\tau\tau}$ 2012 category, respectively. Moreover, the normalisation factors on the $Z \to \tau\tau$ and multi-jet backgrounds are also expected to have a significant impact on $\Delta\hat{\mu}_{\mathrm{tot}}$. As a further check, Fig. B.14 shows the NP pulls for the fits in the VBF and Boosted categories separately. The observed pulls reflect the ones seen in the combined fit and there are no significant deviations from the input values.

The test is the fit to the observed data in the mass sidebands $0 < m_{\tau\tau}^{\mathrm{MMC}} < 110$ GeV and $m_{\tau\tau}^{\mathrm{MMC}} > 140$ GeV, where negligible signal contribution is expected. Figs. B.15 and B.16 show the observed and predicted distributions in the VBF and Boosted categories, respectively. The measured signal strength is zero as expected. The lower panels show the data with the subtracted predicted total background. The observed agreement confirms the good modelling of the background contribution and the proper convergence of the fit in the mass sidebands.

All these findings do not single out any issue and build confidence about the observed results.
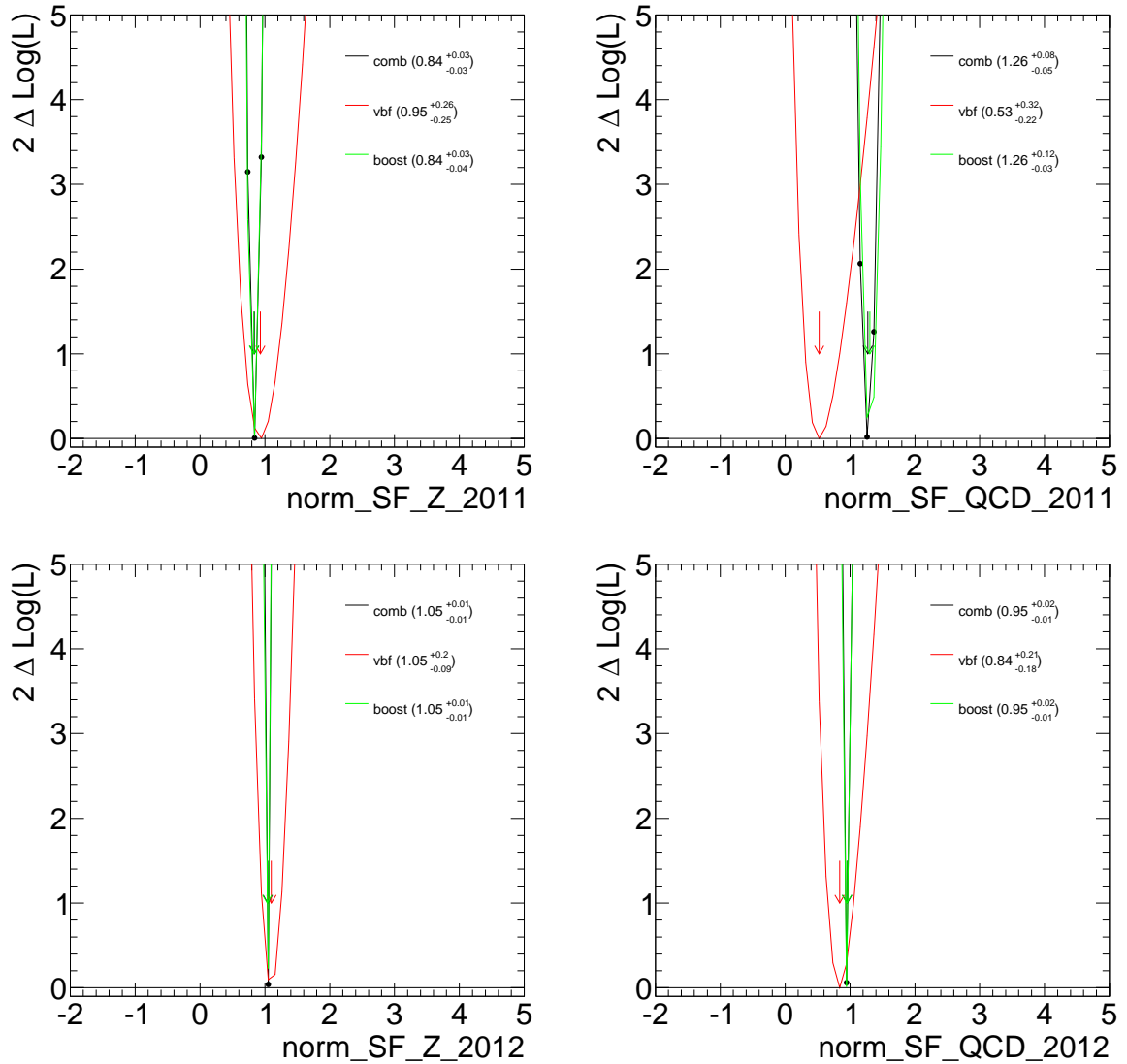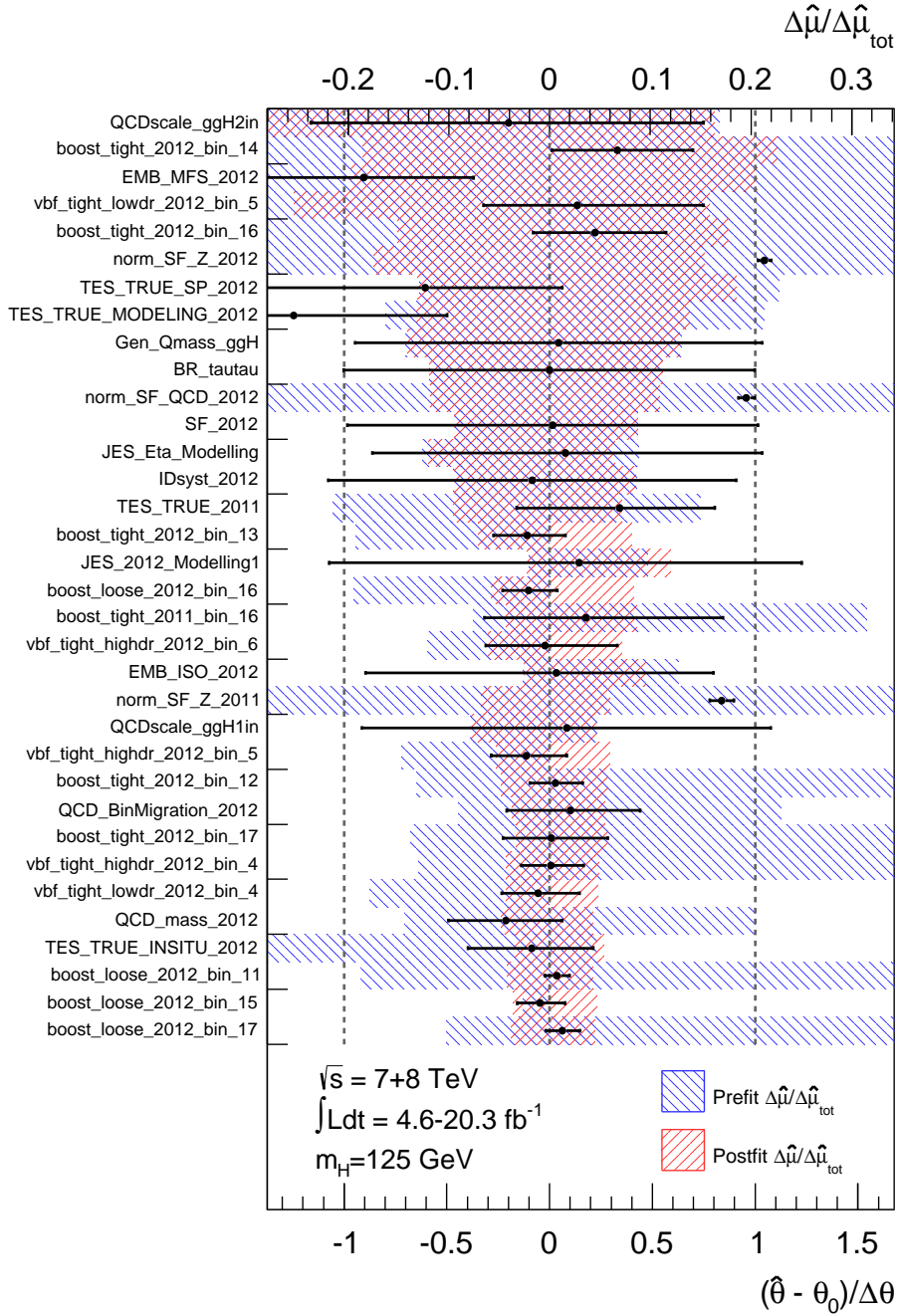
Figure B.12: NLL scans for the normalisation factors on the Z (left) and on the multi-jet (right) backgrounds in 2011 (top) and in 2012 (bottom). Scans are done in the fit of all event categories (black). Scans performed in the VBF and Boosted categories alone are also shown. Arrows indicate the minima of the NLL distributions.

Figure B.13: NP pulls in the combined fit to the observed data. NPs are ranked by the size of their impact on the $\Delta\hat{\mu}_{\mathrm{tot}}$ and only the ones with $\Delta\hat{\mu}/\Delta\hat{\mu}_{\mathrm{tot}} > 0.04$ are shown. The red and blue band indicate the pre- and post-fit $\Delta\hat{\mu}/\Delta\hat{\mu}_{\mathrm{tot}}$. The initial values of NP is zero and of NF is unity. The "*_bin_*" are NPs relative to the statistical uncertainties on the background prediction in a given $m_{\tau\tau}^{\mathrm{MMC}}$ bin.
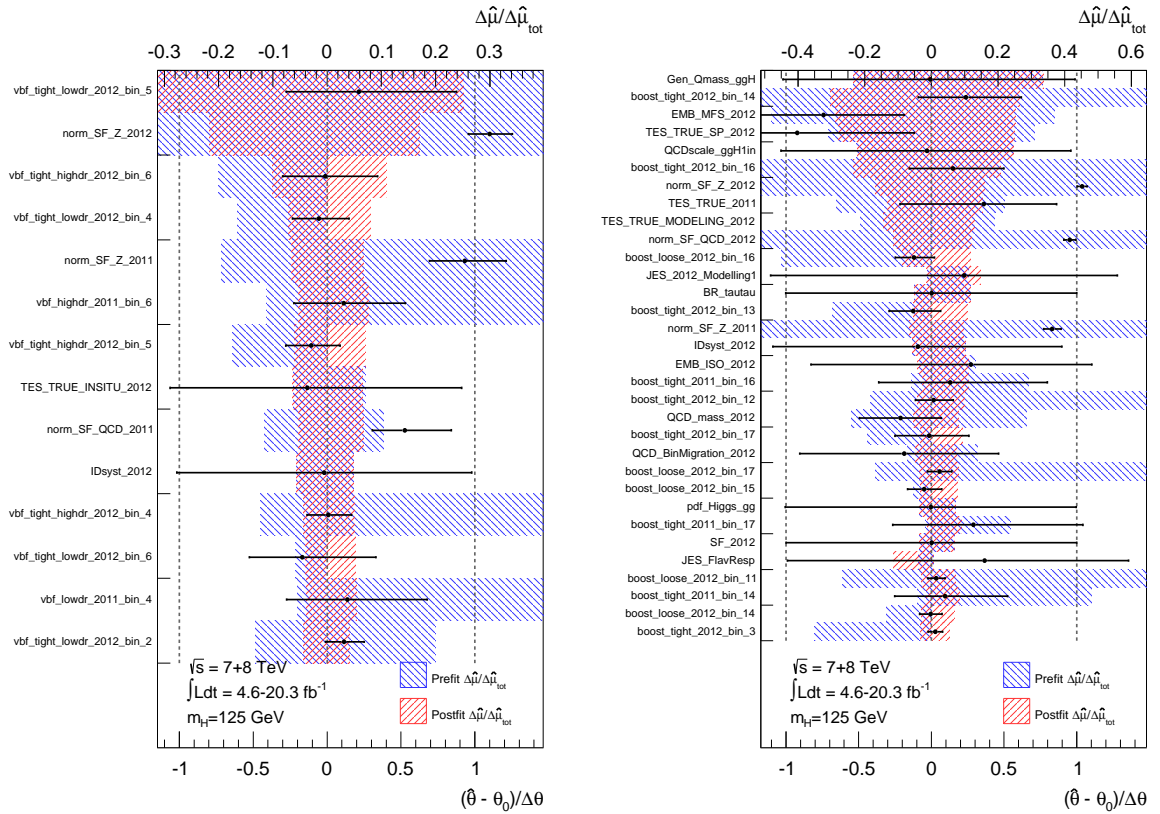
Figure B.14: Observed NP pulls in the fits of the VBF (top left) and Boosted (top right) categories separately. NPs are ranked by the size of their impact on the $\Delta\hat{\mu}_{\text{tot}}$ and only the ones with $\Delta\hat{\mu}/\Delta\hat{\mu}_{\text{tot}} > 0.04$ are shown. The red and blue band indicates the pre- and post-fit $\Delta\hat{\mu}/\Delta\hat{\mu}_{\text{tot}}$. The initial values for NP is zero and for NF is unity. The "*_bin_*" NP are relative to the statistical uncertainties on the background prediction in a given $m_{\tau\tau}^{\text{MMC}}$ bin.
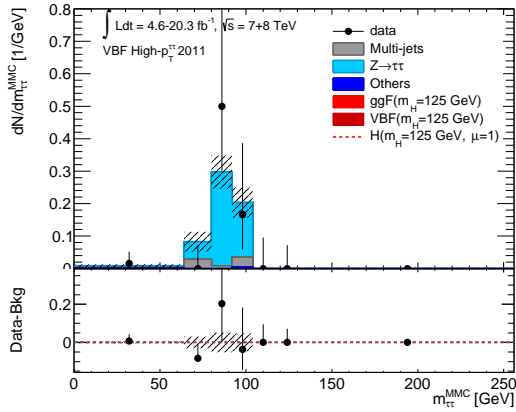
(a) VBF High-$p_{\mathrm{T}}^{\tau\tau}$ at $\sqrt{s} = 7$ TeV
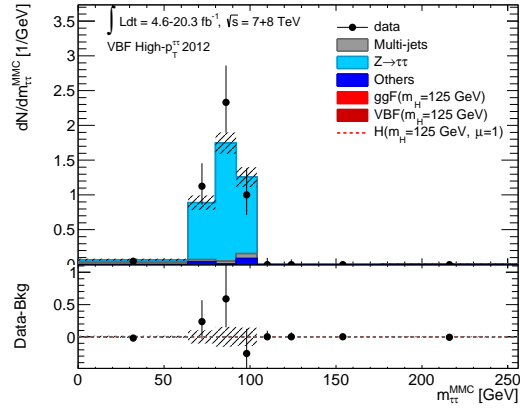
(b) VBF High-$p_{\mathrm{T}}^{\tau\tau}$ at $\sqrt{s} = 8$ TeV
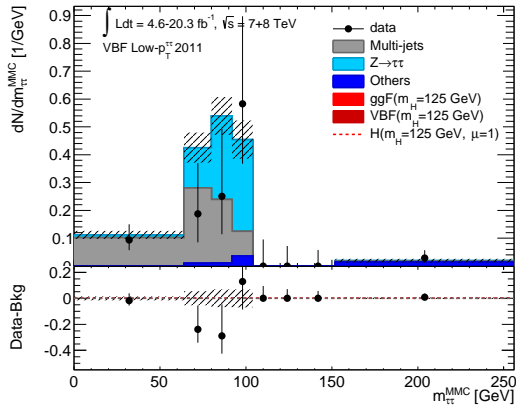
(c) VBF Low-$p_{\mathrm{T}}^{\tau\tau}$ at $\sqrt{s} = 7$ TeV
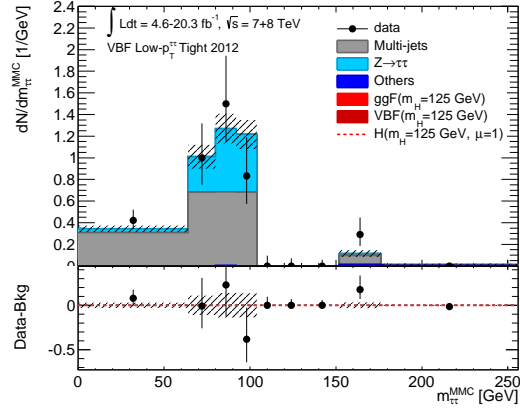
(d) VBF Low-$p_{\mathrm{T}}^{\tau\tau}$ Tight at $\sqrt{s} = 8$ TeV
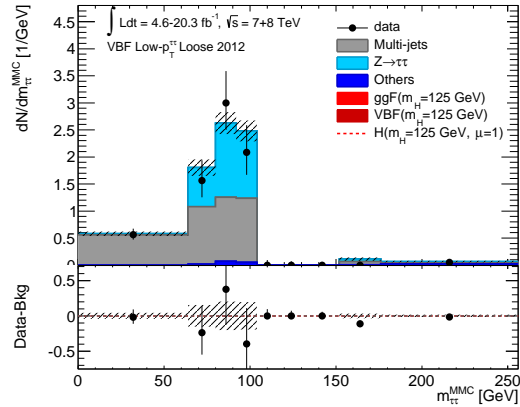
(e) VBF Low-$p_{\mathrm{T}}^{\tau\tau}$ Loose at $\sqrt{s} = 8$ TeV

Figure B.15: Observed $m_{\tau\tau}^{\mathrm{MMC}}$ distributions in the mass sidebands $0 < m_{\tau\tau}^{\mathrm{MMC}} < 110$ GeV and $m_{\tau\tau}^{\mathrm{MMC}} > 140$ GeV in the VBF categories. The lower panel shows the data with the predicted total background subtracted. The dashed area represents the total background uncertainty.
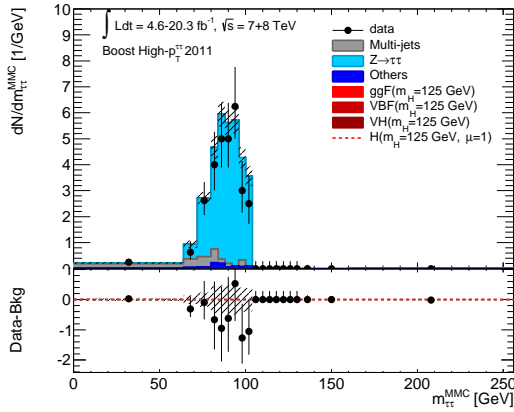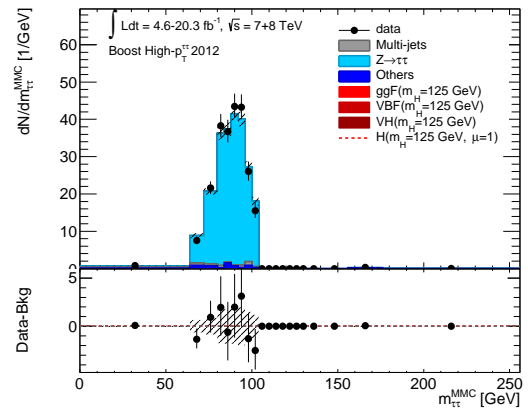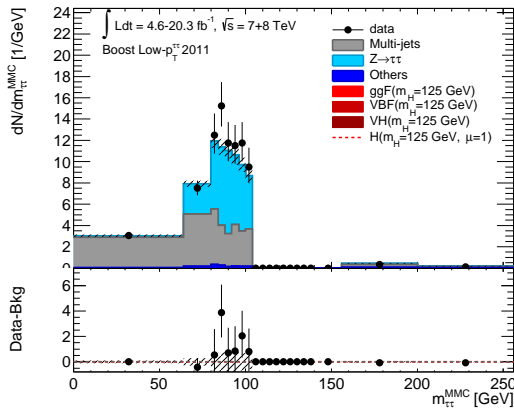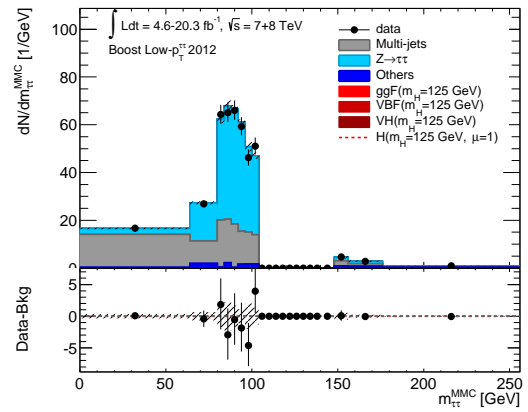
(a) Boosted High-$p_{\mathrm{T}}^{\tau\tau}$ at $\sqrt{s} = 7$ TeV

(b) Boosted High-$p_{\mathrm{T}}^{\tau\tau}$ at $\sqrt{s} = 8$ TeV

(c) Boosted Low-$p_{\mathrm{T}}^{\tau\tau}$ at $\sqrt{s} = 7$ TeV

(d) Boosted Low-$p_{\mathrm{T}}^{\tau\tau}$ at $\sqrt{s} = 8$ TeV

Figure B.16: Observed $m_{\tau\tau}^{\mathrm{MMC}}$ distributions in the mass sidebands $0 < m_{\tau\tau}^{\mathrm{MMC}} < 110$ GeV and $m_{\tau\tau}^{\mathrm{MMC}} > 140$ GeV in the Boosted categories. The lower panel shows the data with the predicted total background subtracted. The dashed area represents the total background uncertainty.