TECHNISCHE UNIVERSITÄT
MÜNCHEN

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

Study of the Higgs Boson Discovery Potential in the Process pp $\rightarrow$ Hqq, H $\rightarrow \tau^+\tau^-$ with the ATLAS Detector

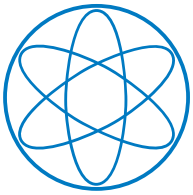
Manfred Groh

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 26.03.2009 bei der Technischen Universität München eingereicht und durch die Fakultät für Physik am 27.04.2009 angenommen.

PHYSIK-DEPARTMENT



Study of the Higgs Boson Discovery Potential in the Process $pp \to Hqq,\ H \to \tau^+\tau^-$ with the ATLAS Detector

Dissertation
von

Manfred Groh

7. Mai 2009

Technische Universität München

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

Abstract

The subject of this work is the evaluation of the discovery potential of the ATLAS detector at the Large Hadron Collider for the Standard Model Higgs boson in vector-boson fusion production and a subsequent decay into a τ-lepton pair. This is one of the most promising discovery channels of the Higgs boson in the low mass range, which is the mass range favored from precision measurements of the electroweak interaction. The decay modes where both τ leptons decay leptonically and where one τ lepton decays leptonically and the other one hadronically are studied in this thesis.
The main objective was to investigate possible improvements upon earlier cut-based analyses by using additional discriminating variables as well as by applying multivariate analysis methods which take into account correlations between the variables. The variables are carefully selected in order to avoid correlations with the reconstructed invariant $\tau\tau$ mass.
In an intermediate step, the sequential signal selection cuts have been optimized for maximum signal significance. With this strategy, one can expect to discover the Higgs boson with $\geq 5\,\sigma$ significance in the mass range $115\ \text{GeV} \leq m_H \leq 135\,\text{GeV}$ with an integrated luminosity of $30\,\text{fb}^{-1}$ corresponding to the first three years of ATLAS operation. The maximum signal significance of $5.9\,\sigma$ is obtained for a Higgs mass of 120 GeV.
Significant further improvement was found with multivariate selection methods. The best results are obtained with an Artificial Neural Network algorithm. The mass range for the $\geq 5\,\sigma$ Higgs discovery with $30\,\text{fb}^{-1}$ is extended to 110 GeV with a maximum signal significance of $6.5\,\sigma$ at $m_H = 125\,\text{GeV}$.
Systematic uncertainties are studied in detail for both methods and are included in the above predictions of the signal significance. The largest uncertainty is due to the jet energy scale. In the case of using only Monte Carlo simulations for estimating the background, the uncertainties on the detector performance lead to a big loss in discovery potential. It is demonstrated that a reliable method for background estimation from real data is essential. In this case, the systematic uncertainties on the expected signal significance are about 10 % for both analysis methods.

Acknowledgements

Let me devote this page to all the people who supported my work during the last years.
First, I want to express my gratitude to my supervisor Hubert Kroha for giving me the opportunity to do both my diploma and my PhD thesis in the MDT group at MPI. I thank him for his supervision of the thesis, for providing the means and giving me the possibility to become part of the team building up such a fascinating experiment as the ATLAS detector, for providing me the chance to participate in schools, workshops, seminars and many interesting conferences.
I thank Susanne Mohrdieck-Möck, Oliver Kortner, Jörg Dubbert and especially Sandra Horvat for introducing me to the field of experimental high energy physics and guiding me through the last years. All of them showed much patience with me and always took the time to discuss my thoughts, no matter how busy they were. Many thanks to Steffen Kaiser for his support especially during the last weeks.
Thank you to all the members of the MDT group for creating the stimulating and pleasant atmosphere, the coffee in the morning, the Cheeseburgers and Spezis, the beers, bars and Schnitzels, for the tabletennis matches and much more.

Bei all meinen Freunden bedanke ich mich für ihre Unterstützung, das Interesse an meiner Arbeit und deren Fortschritt sowie dafür, dass sie trotz meiner eingeschränkten Zeit uneingeschränkt hinter mir gestanden haben. Für die Hilfe beim Korrekturlesen und Ausdrucken möchte ich mich außerdem ganz herzlich bei Martin Mühlegger bedanken.
Meiner Familie, speziell meinen Eltern Magdalena und Bruno Groh, gebührt ganz besonderer Dank für den Rückhalt, den sie mir während der letzten nun schon fast 30 Jahre gegeben haben und dafür, dass sie mir mein Studium ermöglicht haben.

Unmöglich in Worte zu fassen will ich trotzdem meinen unendlichen Dank dem allerwichtigsten Menschen in meinem Leben aussprechen: meiner Irene. Trotz ihrer eigenen Arbeitsbelastung und ihres Studiums hat sie mich immer in allen alltäglichen und nicht alltäglichen Dingen unterstützt und mir die Kraft und die Geborgenheit gegeben, ohne die diese Arbeit nicht möglich gewesen wäre. Ganz besonders danke ich ihr für ihr Verständnis, wenn ich aufgrund meiner zahlreichen Aufenthalte am CERN oder bei Konferenzen, sowie vor allem in den letzten Monaten nur sehr wenig Zeit mit ihr verbringen konnte.

Contents

Chapter 1

Introduction

The Standard Model of particle physics describes the current understanding of fundamental particles and their interactions except gravity. It has been developed during the last century and has proven to be extremely successful. Nevertheless, the origin of particle masses remains an open question. The electroweak symmetry breaking mechanism as introduced by Peter Higgs [1] and others [2, 3] provides an elegant answer. However, it predicts also a yet undiscovered particle, the Higgs boson. The LEP[1] experiments excluded a Higgs boson with a mass lighter than 114.4 GeV [4], while the Tevatron experiments recently excluded the existence of a Higgs boson with a mass close to 170 GeV [5]. Precision tests of the electroweak interaction favor a light Higgs boson with a mass below 185 GeV [6]. Chapter 2 gives an introduction to the Higgs mechanism and the current status of experimental searches for the Higgs boson.

In autumn 2009, the Large Hadron Collider (LHC) at CERN[2] will start its operation. Proton-proton collisions with a center of mass energy of up to $\sqrt{s} = 14$ TeV and a luminosity of up to $10^{34}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$ will open up a new era in particle physics allowing for the discovery of the Higgs boson over its entire allowed mass range. Therefore, the LHC will provide an answer to the question whether the Higgs mechanism exists in nature or not. In particular, the general-purpose detector ATLAS[3] is designed to discover the Higgs boson and to study a wide range of other physics subjects. Chapter 3 summarizes the design and the properties of the LHC and the ATLAS detector.

The high-radiation conditions at the LHC and the requirements for excellent solid angle coverage, efficiency and resolution for the reconstruction of particles emerging from the collisions demand high performance and reliability of all detector components. They had to pass stringent tests during production and after installation. A large fraction of the ATLAS muon detectors was built at the Max-Planck-Institut für Physik in Munich in collaboration with the Ludwig-Maximilians University. Chapter 4 gives an overview of the installation of the muon detectors and of the commissioning of the ATLAS muon spectrometer with cosmic ray muons.

The Higgs decay into a τ-lepton pair which further decays into leptons and/or hadrons is one of the most promising processes for the discovery of a light Higgs boson with a mass below ~150 GeV.

[1] LEP: Large Electron-Positron collider
[2] CERN: Conseil Européene pour la Recherche Nucléaire
[3] ATLAS: A Toroidal LHC ApparatuS

The exclusive Higgs production mechanism via vector-boson fusion is needed for the detection of this decay channel in order to sufficiently suppress the large background mainly caused by Z boson as well as top quark pair production. The vector-boson fusion process provides a characteristic signature with two highly energetic jets in the forward regions of the detector while the Higgs decay products are preferably emitted into the central detector region. Details of this Higgs decay channel and the most important background processes are discussed in Chapter 5 as well as the criteria which can be used to separate signal from background events.
Two approaches for the discrimination between signal and background are studied in the course of this work. In the first approach, a series of cuts is applied on discriminating variables. In Chapter 6 the cut optimization procedure for a maximum signal significance is presented. The second approach uses multivariate methods, which are able to take correlations into account. In this study, several different multivariate methods have been tested for their signal selection performance. In Chapter 7 the selection of discriminating variables and the results for the best performing multivariate analysis method are described.
Chapter 8 provides a comparison of the discovery potential of the two analysis methods for the Higgs mass range accessible with this channel. Finally, systematic uncertainties related to the imperfect knowledge of the ATLAS detector performance as well as their impact on the discovery potential are discussed.
The results presented in this work are preliminary and not yet officially approved by the ATLAS collaboration.

Chapter 2

Theoretical Background

The Standard Model is an extremely successful theory comprising our current understanding of fundamental particles and their interactions. It is based on a spontaneously broken local $SU(3) \times SU(2) \times U(1)$ gauge theory describing the strong, weak and electromagnetic interactions. A brief overview of the Standard Model is given in Section 2.1 including the Higgs mechanism which describes the origin of the masses of the fundamental particles. Further details can be found, for example, in [7]. Section 2.2 deals with the Higgs boson predicted by the theory. Limits on the Higgs boson mass from theory and experiments are discussed in Section 2.2.1. In Sections 2.2.2 and 2.2.3, the production mechanisms and decay channels of the Higgs boson in the Standard Model are described.

2.1 The Standard Model

In the Standard Model, three types of fundamental particles are distinguished according to their spin:

- Fermions with spin 1/2
  The Fermions f are the matter constituents and are grouped further into leptons l and quarks q. Both leptons and quarks exist in three generations with increasing mass. All known stable matter on earth is formed of fermions of the first generation. Table 2.1 gives an overview of the known fermions.

- Bosons with spin 1
  They are the vector bosons of the gauge fields mediating the three fundamental forces: the electromagnetic, the weak and the strong interaction.

- The Higgs boson with spin 0 is associated with the spontaneous breaking of the electroweak gauge symmetry.

All three forces are described by local gauge theories with the three simplest gauge symmetry groups $U(1)$, $SU(2)$ and $SU(3)$. The strong force between the quarks is described by Quantum Chromodynamics (QCD) [8, 9, 10, 11], a $SU(3)$ gauge theory. It contains eight gauge fields associated with eight massless vector bosons called gluons.

	Generation			Electric Charge
	1	2	3	[e]
Leptons	ν_e	ν_μ	ν_τ	0
	e	μ	τ	-1
Quarks	u	c	t	$+2/3$
	d	s	b	$-1/3$

Table 2.1: The three generations of fermions in the Standard Model.

Between 1960 and 1968, Glashow [12], Weinberg [13] and Salam [14] developed a $SU(2) \times U(1)$ gauge field theory which combines the electromagnetic and weak force and is thus called electroweak theory. The *Glashow-Salam-Weinberg Theory* includes the local $SU(2)$ gauge group of the weak isospin and the $U(1)$ gauge group of the weak hypercharge with two gauge couplings g and g' respectively. The $SU(2)$ gauge symmetry requires three gauge fields W_μ^1, W_μ^2 and W_μ^3 related to the three components $J_i (i = 1, 2, 3)$ of the weak isospin vector, while the $U(1)$ gauge theory contains only one gauge field B_μ related to the weak hypercharge. Linear combinations of these four gauge fields describe the observed particles mediating the weak and electromagnetic interactions, namely the charged weak gauge bosons W^+ and W^- described by the fields

$$W_\mu^\pm = \frac{1}{\sqrt{2}}(W_\mu^1 \pm i W_\mu^2) \tag{2.1}$$

and the neutral weak and electromagnetic gauge fields

$$\begin{pmatrix} A_\mu \\ Z_\mu \end{pmatrix} = \begin{pmatrix} \cos\theta_W & \sin\theta_W \\ \sin\theta_W & \cos\theta_W \end{pmatrix} \begin{pmatrix} B_\mu \\ W_\mu^3 \end{pmatrix} \tag{2.2}$$

which correspond to the photon and the Z^0 boson and are related to B_μ and W_μ^3 by a rotation with the weak mixing or Weinberg angle θ_W.
However, this theory has one severe problem: To conserve $SU(2)$ symmetry, the weak gauge bosons $W^\pm$ and Z^0 have to be massless which is in contradiction to observations.

The Higgs Mechanism The solution to the problem of massive gauge bosons was provided by P. W. Higgs [1] and others [15, 16] in the year 1964. Based on the work of Nambu [2] and Goldstone [3], they developed a mechanism, the Higgs mechanism, in which massive gauge bosons can be accommodated by introducing a complex scalar field of the form

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} \tag{2.3}$$

described by the Lagrangian

$$\mathcal{L} = D_\mu \phi^\dagger D^\mu \phi - V \tag{2.4}$$

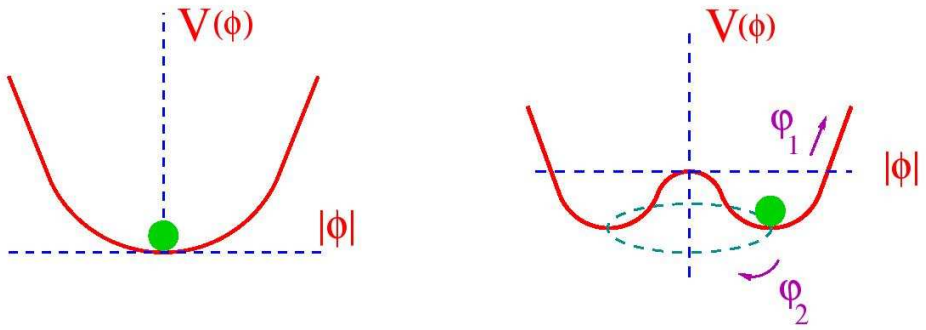


Figure 2.1: Illustration of the Higgs potential for $\mu^2 > 0$ (left) and $\mu^2 < 0$ (right). In the latter case, spontaneous symmetry breaking occurs when the system moves to one of the ground states [7].

with the covariant derivative of the $SU(2) \times U(1)$

$$D_\mu = \partial_\mu - igJ_iW_{\mu i} - ig'YB_\mu \tag{2.5}$$

and the potential energy function

$$V(\phi) = \mu^2\phi^\dagger\phi + \lambda(\phi^\dagger\phi)^2. \tag{2.6}$$

To ensure the existence of stable ground states, the potential is bound from below by requiring $\lambda > 0$.
Figure 2.1 illustrates the potential for $\mu^2 > 0$ and $\mu^2 < 0$. In the latter case, the minimum of the potential is not at $|\phi_0| = 0$ but at:

$$|\phi_0|^2 = \frac{-\mu^2}{2\lambda} \equiv \frac{v^2}{2} \tag{2.7}$$

with v called the vacuum expectation value of the scalar field. The $SU(2) \times U(1)$ symmetry is *spontaneously broken* when one of the ground states for $\mu^2 < 0$ is chosen. The Lagrangian remains invariant under *local* $SU(2) \times U(1)$ gauge transformations.
One can choose, for instance, the ground state

$$\phi_0 = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ v \end{pmatrix} \tag{2.8}$$

and parameterize excitations from this ground state by

$$\phi = \frac{1}{\sqrt{2}}e^{ij_i\zeta_i(x)}\begin{pmatrix} 0 \\ v + H(x) \end{pmatrix} \tag{2.9}$$

with a scalar field $H(x)$, the massive Higgs field which describes radial excitations from the ground state changing the potential energy, and massless scalar fields $\zeta_i(x)$, the Goldstone bosons, corresponding to angular excitations without potential energy change.

The latter can be eliminated by a local $SU(2)$ gauge transformation leading to the parameterization of the scalar field

$$\phi = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + H \end{pmatrix}. \tag{2.10}$$

Introducing the ansatz 2.10 into the Lagrangian 2.4 of the electroweak theory and using equations 2.1 and 2.2, one obtains the following expressions from the kinetic terms:

$$\frac{g^2 v^2}{8} W_\mu^+ W^{+\mu} + \frac{g^2 v^2}{8} W_\mu^- W^{-\mu} + \frac{g^2 v^2}{8 \cos^2 \theta_W} Z_\mu Z^\mu. \tag{2.11}$$

These terms can be identified as mass terms of the $W^\pm$ and Z^0 bosons which thus acquire masses

$$m_Z \cos \theta_W = m_W = \frac{vg}{2} \tag{2.12}$$

due to the coupling to the Higgs field.
The masses of the fermions are also generated by spontaneous $SU(2) \times U(1)$ breaking due to Yukawa couplings to the Higgs field. The coupling strength of a fermion to the Higgs field, which is not predicted by the theory, is proportional to the fermion mass.
The Glashow-Salam-Weinberg theory includes the Higgs mechanism. It predicts the W and Z boson and their properties, like mass and decay width. In 1983, the three weak gauge bosons have been discovered by the UA1 and UA2 experiments at CERN [17, 18, 19, 20] with the predicted properties.

2.2 The Higgs Boson

Although the Standard Model is widely accepted and very successful in describing the phenomena of particle physics, there is still one part missing: The experimental verification of the existence of the Higgs boson introduced in equation 2.8. The value of the mass of the Higgs boson

$$m_H = \sqrt{2\lambda}\, v \tag{2.13}$$

is not predicted by the Standard Model and has to be measured. However the possible mass range is restricted by theoretical and experimental constraints. These limits are briefly discussed in the following.

2.2.1 Limits on the Higgs Boson Mass

Theoretical Limits Several consistency requirements of the theory set upper and lower bounds on the Higgs mass in the Standard Model depending on the energy scale Λ up to which the Standard Model is valid and no new interactions or particles appear. Values of Λ up to the Planck Mass $M_{\text{Planck}} = 10^{19}$ GeV are considered above which gravitation becomes as strong as the other fundamental forces and the Standard Model must, at the latest, be extended by a quantum theory of gravitation. Here, only the results are summarized (further details can be found, for example, in [21] and references therein):

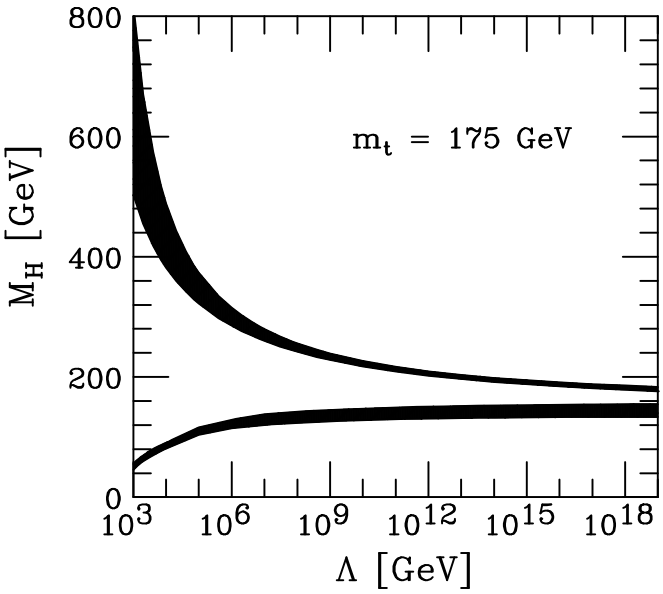


Figure 2.2: Upper and lower theoretical bounds on the Higgs boson mass as a function of the energy scale Λ up to which the Standard Model is valid. A top quark mass of $m_t = 175\,\text{GeV}$ is assumed. The bands indicate the theoretical uncertainties. [21]

- Unitarity of the electroweak interactions, in particular of the $W^+W^- \to W^+W^-$ scattering amplitude, limits the Higgs boson mass to $m_H \lesssim 1\,\text{TeV}$.
- The requirement of finite self-coupling of Higgs bosons, including Higgs and top quark loops, restricts the Higgs mass with an upper bound depending on Λ:
  $m_H \lesssim 600\,\text{GeV}$ for $\Lambda = 1\,\text{TeV}$ and $m_H \lesssim 180\,\text{GeV}$ for $\Lambda = M_{\text{Planck}}$.
- To ensure the stability of the Higgs ground state, the Higgs potential (see equation 2.6) has to have a lower bound ($\lambda(\Lambda) > 0$). This results in a lower limit on the Higgs boson mass of $m_H \gtrsim 55\,\text{GeV}$ for $\Lambda = 1\,\text{TeV}$ and $m_H \gtrsim 130\,\text{GeV}$ for $\Lambda = M_{\text{Planck}}$.

Figure 2.2 shows the theoretical upper and lower bounds for the Standard Model Higgs boson as a function of Λ. The measurement of the Higgs boson mass will constrain Λ. For example, a Higgs boson mass of $m_H = 500\,\text{GeV}$ implies that the Standard Model breaks down already at a much lower energy scale than M_{Planck}.

For $\Lambda = 1\,\text{TeV}$, theory predicts a Higgs boson in the wide mass range of

$$55\,\text{GeV} \lesssim m_H \lesssim 600\,\text{GeV}. \tag{2.14}$$

Limits from Experiments A lower bound of the Higgs boson mass comes from direct Higgs boson searches at the Large Electron Positron Collider LEP at CERN. By combining the data of

all four LEP experiments (ALEPH, DELPHI, L3 and OPAL) the existence of a Standard Model Higgs boson with a mass below $m_H < 114.4\,\text{GeV}$ is excluded at the 95 % confidence level [4].

The range of possible Higgs boson masses can be further constrained by combining the precision measurements of electroweak observables at LEP and at the Tevatron $p\bar{p}$ collider at Fermilab near Chicago with the experiments CDF and D0. Higher order corrections to the calculation of electroweak processes like $e^+e^- \to Z \to f\bar{f}$ include Higgs loops and thus depend on the Higgs mass. The measurements of LEP and Tevatron are combined in [6] with a precision sufficient to constrain the Higgs boson mass. The most recent result of this calculation is given in Figure 2.3 which shows $\Delta\chi^2 \equiv \chi^2 - \chi^2_{min}$ of the global least-squares fit of the Standard Model predictions to the electroweak data as a function of the Higgs boson mass. It can be seen that the fit to electroweak precision measurements favors a small Higgs boson mass. The most probable Higgs mass according to the fit is $m_H = 85^{+39}_{-28}\,\text{GeV}$, taking the lower Higgs mass limit from direct searches at LEP into account the 95 % confidence level upper limit being $m_H \leq 185\,\text{GeV}$ [6].

The latest results from direct Higgs boson searches at the Tevatron collider exclude the Standard Model Higgs boson with a mass around 170 GeV at a confidence level of 95 % [5].

2.2.2 Higgs Boson Production Mechanisms

At the LHC, the Higgs boson can be produced by several processes. The Feynman diagrams of the most important processes are shown in Figure 2.4.

The cross-sections of these processes depend on the unknown Higgs boson mass. Figure 2.5 shows the Standard Model Higgs boson production cross-sections in pp collisions at $\sqrt{s} = 14\,\text{TeV}$ in the Higgs mass range $100\,\text{GeV} < m_H < 500\,\text{GeV}$. The gluon fusion dominates over the whole mass range. The vector-boson fusion cross-section is roughly one order of magnitude smaller. However, the two outgoing quarks in this process form two characteristic jets in the very forward regions of the detector at high pseudorapidity values $|\eta|$ which provide a very good signature for background suppression. The other production processes have much lower cross-sections. However, they can still be exploited for specific searches. For example, requiring two additional b quarks from the t-quark decays of the associated Higgs boson production $gg, qq \to ttH$ is essential to suppress the background in searches for Higgs boson decays into a $b\bar{b}$ pair.

Considering that the maximum Higgs boson production cross-section for low Higgs masses is $\sigma_H \approx 50\,\text{pb}$ and the total pp cross-section at the LHC $\sigma_{total} \approx 125\,\text{mb}$ [24], a major challenge of the LHC experiments becomes clear: compared to other pp reactions, the Higgs boson signal is suppressed by ten orders of magnitude. Extreme care has to be taken to understand and reject the background processes.

2.2.3 Higgs Boson Decay Channels

Also the branching ratios of the Standard Model Higgs boson decays depend on m_H as unknown parameter. Since the Higgs boson couples to other particles proportional to their masses, it decays predominantly into the most massive particles accessible. Feynman diagrams of the Standard Model Higgs boson decays are shown in Figure 2.6: the tree-level decays into a fermion pair or a real or virtual weak gauge boson pair and, via loops, the decays into massless photons or gluons.

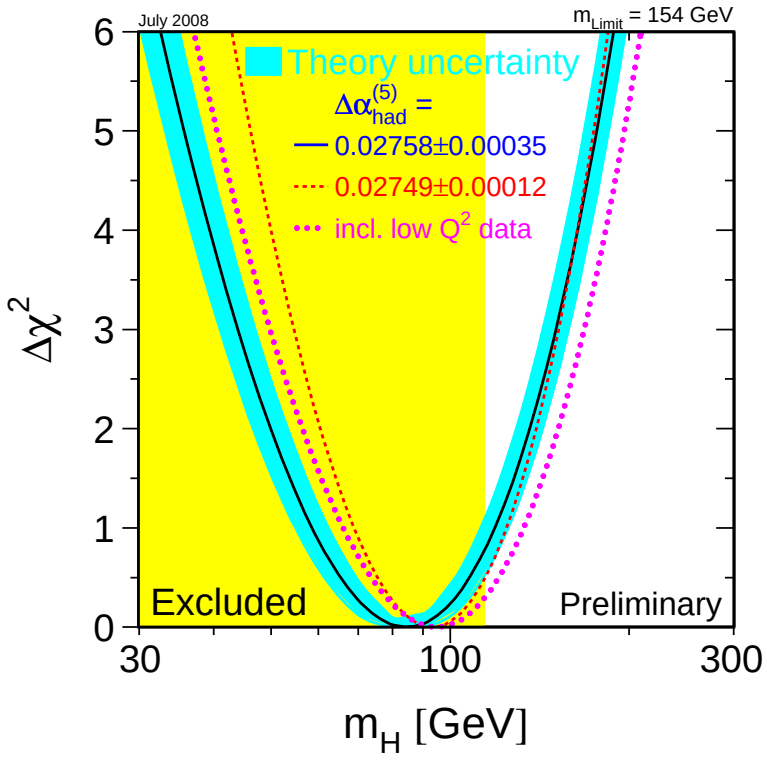
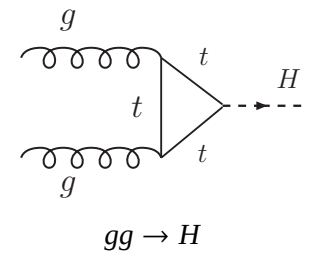
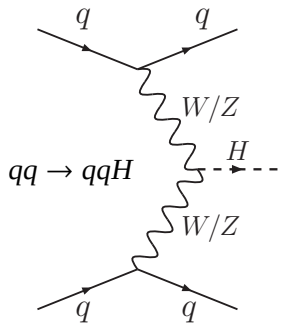


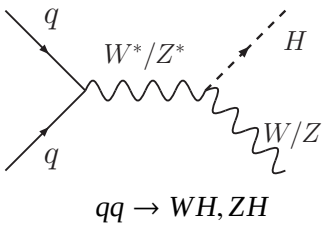
Figure 2.3: $\Delta\chi^2 = \chi^2 - \chi^2_{min}$ from the global least-squares fit to electroweak precision measurements as a function of the Higgs boson mass m_H [6]. The shaded band indicates the theoretical uncertainty including the error in the hadronic part $\Delta\alpha^5_{had}$ of radiative corrections to the fine structure constant α. The vertical band indicates the 95 % exclusion region from direct Higgs boson searches at LEP [4].



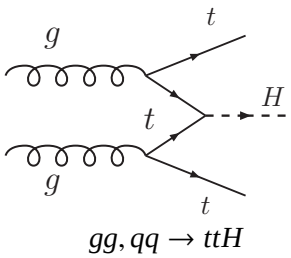
(a) Gluon fusion



(b) Vector-boson fusion



(c) Higgsstrahlung of W/Z



(d) Associated production with a top quark pair

Figure 2.4: Tree-level Feynman diagrams of the Higgs boson production processes in proton-proton collisions at the LHC.

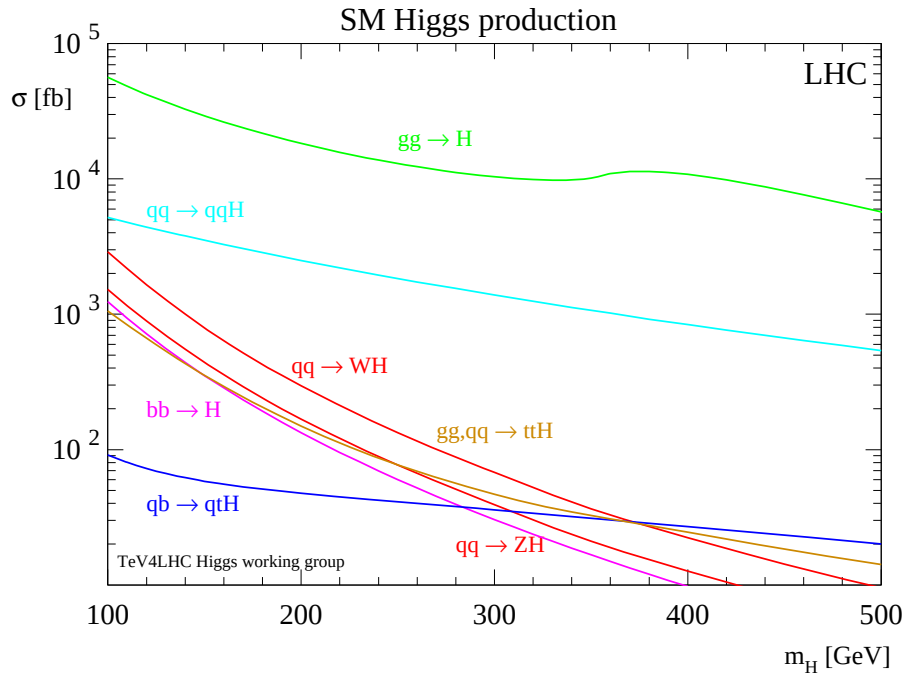


Figure 2.5: Standard Model Higgs boson production cross-sections for the processes in Figure 2.4 as a function of the Higgs boson mass m_H in pp-collisions at the nominal LHC center-of-mass energy of 14 TeV [22, 23].

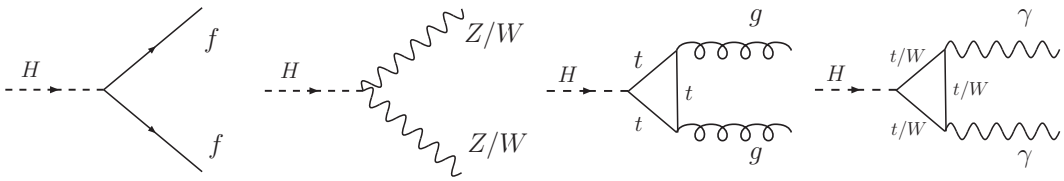


Figure 2.6: Feynman diagrams of the Standard Model Higgs boson decay channels.

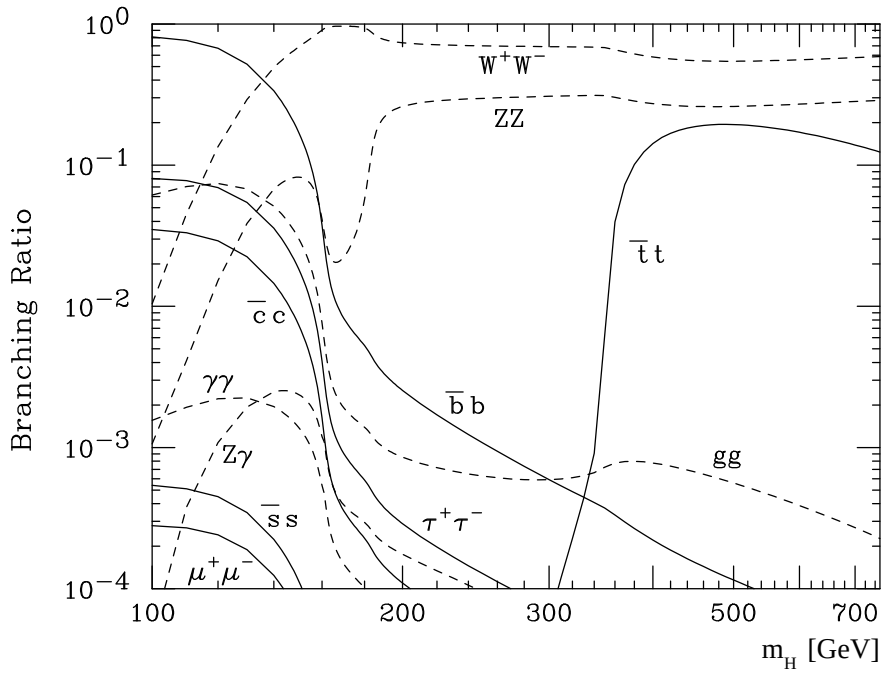


Figure 2.7: Branching ratios of the Standard Model Higgs boson decays as a function of the Higgs boson mass m_H [25].

Figure 2.7 shows the branching-ratio calculations of the most important decay channels. They are described in more detail in the following:

$H \to b\bar{b}$

Since the coupling of the Higgs boson to fermions is proportional to the fermion mass, the decay into a $b\bar{b}$ pair has the largest branching ratio for $m_H \lesssim 140\,\text{GeV}$. However, the discovery potential for this decay channel suffers from very large QCD background.

$H \to gg$

The problem of large QCD background is even more pronounced in the case of the decay into a gluon pair which has the second largest branching ratio in this mass region. The huge QCD jet production cross-sections make it practically impossible to identify $H \to gg$ decays at the LHC.

$H \to \tau^+\tau^-$

The Higgs boson decay into a τ-lepton pair also suffers from high background mainly from $Z \to \tau\tau$ decays. However, if one exploits the signatures of vector-boson fusion Higgs production, this decay channel is one of the most promising ones for $m_H \lesssim 140\,\text{GeV}$. The investigation of the discovery potential of the ATLAS experiment for this decay channel is the main subject of this thesis.

$H \to \gamma\gamma$

Another important decay channel for low Higgs masses is the decay $H \to \gamma\gamma$. Although it has only a very small branching ratio, its discovery potential is high due to the clean signature of two energetic photons and the high Higgs mass resolution in this channel.

$H \to W^+W^-$

The branching ratio of $H \to W^+W^-$ rises towards the threshold for real W-pair production. For $m_H \sim 160 - 180\,\text{GeV}$, the Higgs boson almost exclusively decays into W^+W^-. Unfortunately, the best identifiable leptonic decays of W bosons involve neutrinos, making it impossible to accurately reconstruct the Higgs boson mass.

$H \to ZZ$

Above $m_H \gtrsim 190\,\text{GeV}$, the decay $H \to ZZ$ is the most promising Higgs discovery channel at LHC. The further decay of the Z bosons into electron or muon pairs provides the cleanest signature and an excellent Higgs mass resolution. Therefore, the decay $H \to ZZ \to 4\ell$ is known as the gold-plated discovery channel for the Higgs boson.

$H \to t\bar{t}$

The $H \to t\bar{t}$ decay becomes kinematically possible above $m_H \gtrsim 350\,\text{GeV}$. However, due to the high background rate in this decay channel and the branching ratio being about ten times smaller than for $H \to W^+W^-$, the decay $H \to t\bar{t}$ is not a Higgs boson discovery channel at LHC.

Chapter 3

The LHC and ATLAS

Up to now, the existence of a Higgs boson has not been proven or excluded by any experiment. After more than 15 years of design and construction, a new proton-proton accelerator, the Large Hadron Collider (LHC) is put into operation at the European particle-physics laboratory CERN which will extend the accessible energy range up to $\sqrt{s} = 14\,\text{TeV}$ at high luminosity. With this accelerator it will be possible to answer the question whether the Standard Model Higgs boson exists and whether there are new phenomena beyond the Standard Model at the TeV scale. One of the general-purpose experiments at the LHC is the ATLAS detector. Simulations of the performance of the ATLAS detector are used in this work to study the discovery potential for the Higgs boson. In the following, the LHC and ATLAS are briefly introduced.

3.1 The Large Hadron Collider

The Large Hadron Collider (LHC) at CERN is a storage ring and accelerator which is equipped with superconducting dipole magnets and which will collide two proton beams with 7 TeV energy each [26]. It is installed in a tunnel of 26.7 km circumference which housed the Large Electron-Positron collider LEP until the year 2002. The tunnel is located 50 to 175 m underground, at the border between Switzerland and France near Geneva. The accelerator contains two vacuum beam pipes, one for each beam direction, which will guide $\sim$ 2800 bunches of up to 10^{11} protons each. With a diameter of 16.6 μm the beams collide at four interaction points at a rate of 40 MHz. The event rate of a pp interaction process is given by:

$$\frac{dN}{dt} = L \cdot \sigma(\sqrt{s})$$

where σ is the cross-section of the process depending on the proton-proton center-of-mass energy $\sqrt{s}$ and L is the instantaneous luminosity which depends only on the beam parameters. The design luminosity of the LHC is $10^{34}\,\text{cm}^{-2}\text{s}^{-1}$. The expected integrated luminosity after the first three years of operation at lower instantaneous luminosity of $10^{33}\,\text{cm}^{-2}\text{s}^{-1}$ is $30\,\text{fb}^{-1}$. This is the integrated luminosity for which the Higgs boson discovery potential is estimated in this work.
On average, 23 inelastic pp collisions will take place in every bunch collision. This means that each interesting event will be overlaid with products from about 22 additional pp interactions in

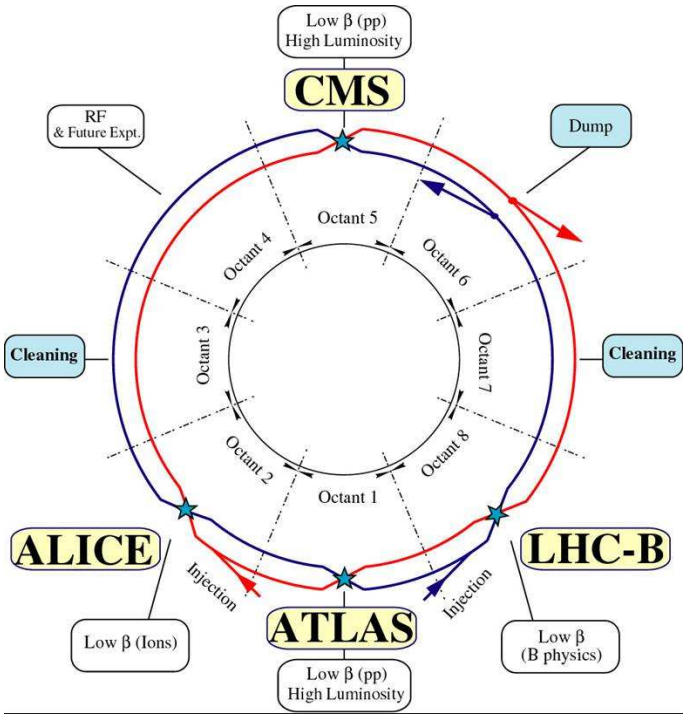


Figure 3.1: Schematic picture of the LHC at CERN with the experiments at four interaction points [26].

the detector, which is called pile-up of interactions. These conditions are extremely challenging for the detectors at the LHC.

Instead of protons, the LHC can also collide lead nuclei with an energy of 2.76 TeV per nucleon and a luminosity of 10^{27} cm^{-2}s^{-1} to probe the formation of a quark-gluon plasma [27].

The following experiments are housed at four interaction points: The general-purpose experiments ATLAS (A Toroidal LHC ApparatuS) [28] and CMS (Compact Muon Solenoid) [29], the LHCb experiment dedicated to B-meson physics [30] and the ALICE experiment [27], specialized to probe heavy ion collisions. Figure 3.1 shows a schematic drawing of the LHC ring, with the positions of the four experiments indicated.

During the LHC commissioning phase in September 2008, the first proton beams circulated in the beam pipe in stable condition in both directions. On September 19th, a fault of a superconducting connection between two dipole magnets occurred which resulted in the damage of several dozens of magnets and the beam pipe. The repair work is currently going on and the restart of the accelerator is scheduled for fall 2009.

3.2 The ATLAS Experiment

ATLAS is an acronym for A Toroidal LHC ApparatuS which refers to the configuration of the magnetic field of the outermost detector part, the muon spectrometer. Figure 3.2 shows a sketch of the ATLAS detector.

3.2.1 Physics Goals and Detector Requirements

The ATLAS experiment aims to study a broad spectrum of physics topics:

- Top quark physics
  Since the LHC will produce dozens of top quarks per second, precision measurements of their production cross-section, mass, coupling and spin can be performed.

- Higgs boson physics
  Searches for Higgs bosons associated with electroweak symmetry breaking in the Standard Model will be performed over the whole allowed mass range up to 1 TeV. Depending on the production and decay mode, this requires efficient identification and precise momentum measurement of electrons and muons, a hermetic detector for missing energy measurement, identification of b and τ jets and measurement of jets in the very forward region. The Higgs boson searches are the benchmark for the detector design and performance.

- Supersymmetric particles
  Many supersymmetric extensions of the Standard Model predict a lightest stable supersymmetric particle that interacts only weakly and therefore escapes the detector leading to a substantial amount of missing energy which has to be reliably reconstructed.

- New physics searches
  The LHC opens up a completely new energy regime. Searches for any kind of new particles or physics processes will be performed with the ATLAS experiment, including searches for new heavy gauge bosons W' and Z' with masses up to ~ 6 TeV, production of mini black-holes and rare decays of heavy quarks and leptons.

The studies put stringent requirements on the detector performance:

- Fast and radiation-hard detectors and readout electronics which can cope with the high radiation level at the LHC and are able to distinguish the decay products of 10^9 proton-proton interactions per second at design luminosity.

- Hermetic detector coverage of the solid angle around the interaction region up to the very forward regions in order to measure the decay products and the energy released in the collisions as completely as possible.

- High momentum resolution and reconstruction efficiency of charged particles, in particular electrons and muons.

- Precise tracking of charged particles to reconstruct the decay vertices of unstable particles.

- An electromagnetic calorimeter with a very good energy and spatial resolution to efficiently identify electrons and photons and to accurately measure their energy.
- A hermetic hadron calorimeter to reliably measure jet energies and the missing transverse energy.
- A highly efficient trigger system which allows for the detection of processes even with very small cross-sections and which provides strong background rejection at the high event rate of the LHC.

3.2.2 The ATLAS Coordinate System

The global right-handed coordinate system of ATLAS is defined as follows:

- The origin is the nominal interaction point.
- The positive x-direction points towards the center of the LHC ring.
- The positive y-direction points upwards.
- The z-direction points along the beam line.

The azimuthal and polar angles with respect to the beam axis are denoted by ϕ and θ. A commonly used quantity in collider experiments is the pseudorapidity η which is defined by

$$\eta = -\ln\tan(\theta/2).$$

Distances in the η-ϕ-space are usually given by

$$\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}.$$

Important observables are defined in the transverse x-y plane: the transverse momentum p_{T}, the transverse energy E_{T} and the missing transverse energy $E_{\mathrm{T}}^{\mathrm{miss}}$. The advantage of these quantities is that they are invariant under boosts along the beam line which usually are present in hadron-hadron collisions.

3.2.3 The ATLAS Detector

Figure 3.2 shows a cut-away view of the ATLAS detector. Like all typical collider experiments, it consists of three concentric layers of subdetectors enclosing the interaction point. The innermost part is the tracking detector, followed by the calorimeters which finally are surrounded by the muon spectrometer. The backbone of the detector is a huge superconducting magnet system. The various detector elements are described in the following.

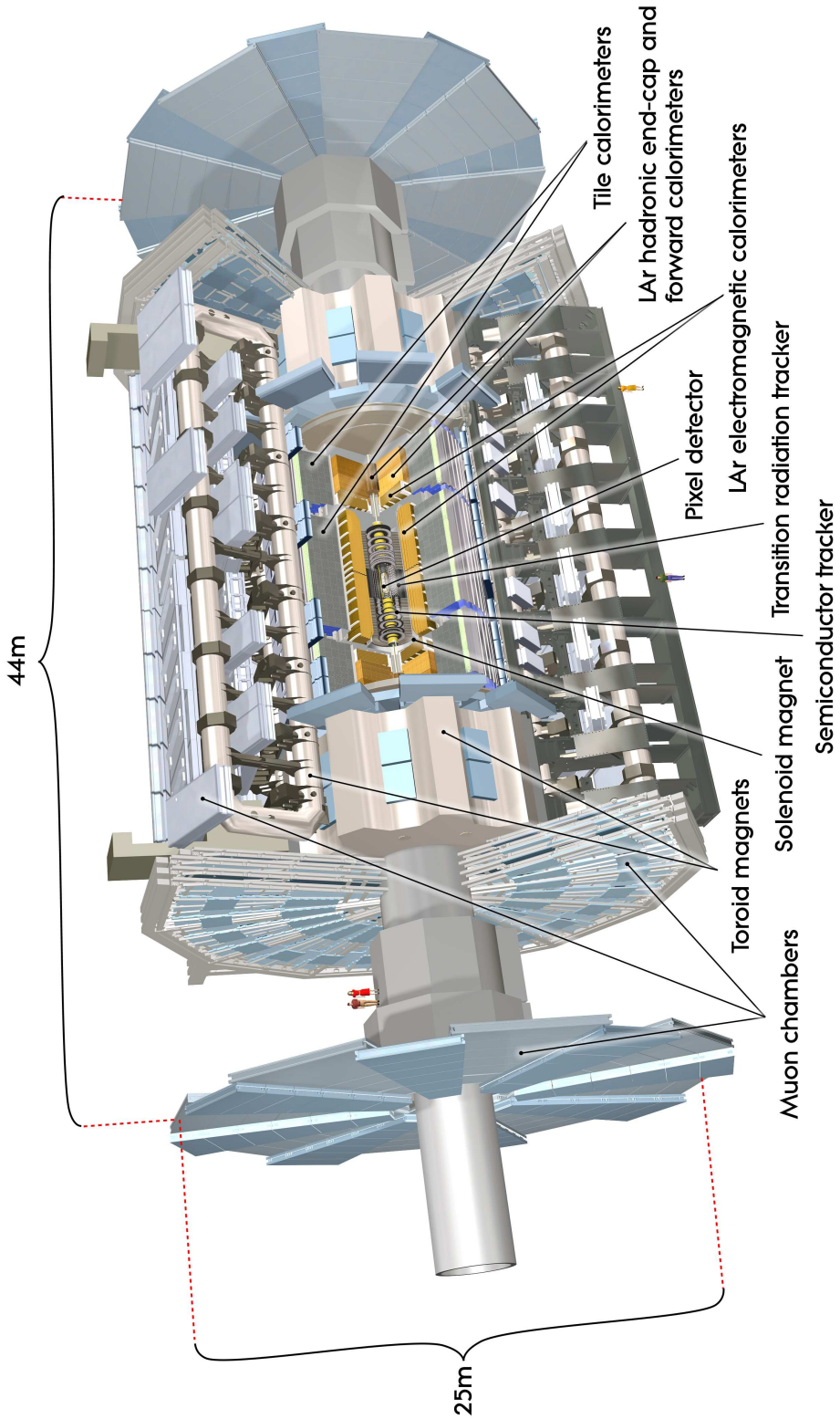


Figure 3.2: Cut-away view of the ATLAS detector. It is 44 m long, 25 m high and weighs ~ 7 000 tons [28].

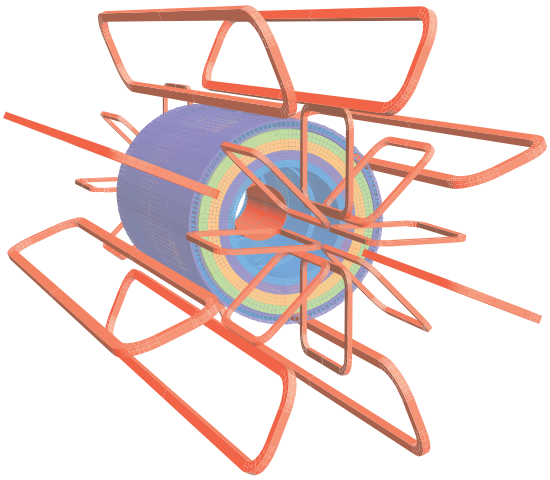


Figure 3.3: Sketch of the ATLAS magnet system with the central solenoid coil (cylinder) and the three toroid magnets around it [28].

The Magnet System

In order to measure the momentum of charged particles, two superconducting magnet systems provide a solenoidal and a toroidal magnetic field in the ATLAS detector. A sketch of the ATLAS magnet system is shown in Figure 3.3.
The central solenoid coil encloses the inner detector and provides a homogeneous magnetic field of 2 T pointing parallel to the beam line. The coil has a diameter of 2.6 m and a length of 5.8 m. In order to minimize the amount of material in front of the calorimeter, the central solenoid shares one vacuum vessel with the central electromagnetic liquid-argon calorimeter. The iron absorber of the electromagnetic calorimeter serves as return yoke.
For the muon spectrometer, a toroidal field configuration has been chosen which is divided into three toroid magnets consisting of eight coils each: The central (barrel) toroid and two end-cap toroids covering the forward regions of the detector. With a length of 25 m and an outer diameter of 20 m, the barrel toroid is the largest component of the ATLAS detector. While each coil of the barrel toroid is housed in its own vacuum vessel, the eight coils of the end-cap toroids are contained in a common cryostat. The magnetic field provided by the toroid magnets is not uniform. The field strength varies between 0.2 T and 2.5 T in the barrel region and from 0.2 T to 3.5 T in the end-caps, depending on the radial distance to the beam line and the azimuthal angle.

The Inner Detector

The inner detector is designed to accurately reconstruct the trajectories of charged particles. It also measures their point of origin (production vertex) to distinguish between particles from the primary hard interaction, from secondary decays, or from additional pile-up interactions. Since the inner detector is immersed in the magnetic field of the central solenoid, the tracks of charged particles are bent in the transverse plane allowing for measuring particle momenta from the track curvature. A radiation hard, fast and highly granular detector is needed to cope with the high track

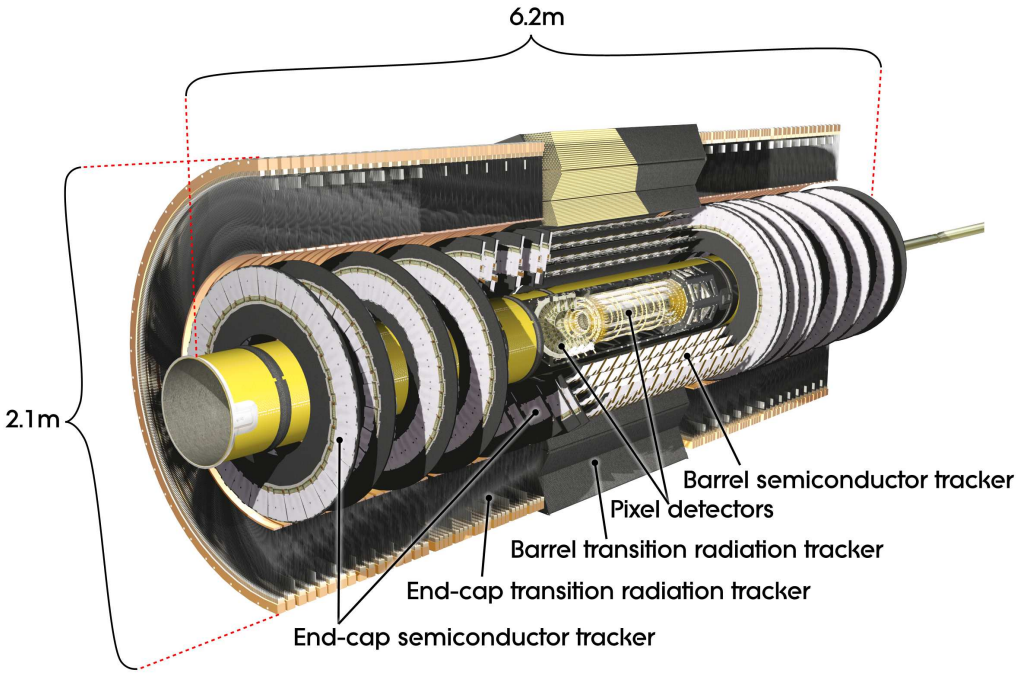


Figure 3.4: Cut-away view of the ATLAS inner detector [28].

multiplicity of the proton-proton collisions at the LHC. To meet these requirements, a combination of three detector technologies is used providing efficient track reconstruction up to $|\eta| < 2.5$. Figure 3.4 shows a cut-away view of the inner detector.
The silicon pixel detector is located closest to the interaction point. More than 80 million pixels are arranged in 3 concentric cylinders around the beam axis in the central part and on 2×3 disks perpendicular to the beam axis in each forward region. The pixel size is $50 \times 400\,\mu\text{m}^2$ allowing for $10\,\mu$m position resolution in $R - \phi$ and $115\,\mu$m in z (R) in the barrel (end-cap) region.

The next layer of the inner detector is the semiconductor tracker (SCT) which consists of almost 16 000 silicon strip sensors with 6.3 million readout channels and is arranged in cylindrical barrel layers and end-cap disks, similar to the pixel detector. Four (2×9) layers of small-angle stereo strip detectors in the barrel (end-caps) measure 4 (9) space points for each track. The mean strip pitch is $80\,\mu$m leading to a position resolution of $17\,\mu$m in $R - \phi$ and $580\,\mu$m in $z(R)$ per layer.

The outermost layer of the Inner Detector is the Transition Radiation Tracker TRT. It consists of more than 300 000 straw drift tubes with a diameter of 4 mm and filled with a $Xe/CO_2/O_2$ gas mixture which allows for the detection of X-rays from transition radiation. The straw tubes are arranged in 73 layers in the barrel and 160 layers in each end-cap. Interleaved between the layers are carbon fibers and foils causing traversing electrons to emit transition radiation which is measured by the neighboring straw tubes. Since the much heavier hadrons emit no transition radiation, the TRT contributes to the discrimination between electrons and hadrons. A charged particle from the interaction point hits on average 36 straw tubes with a single-tube resolution of $130\,\mu$m providing an efficient and accurate track reconstruction in events with high track density.

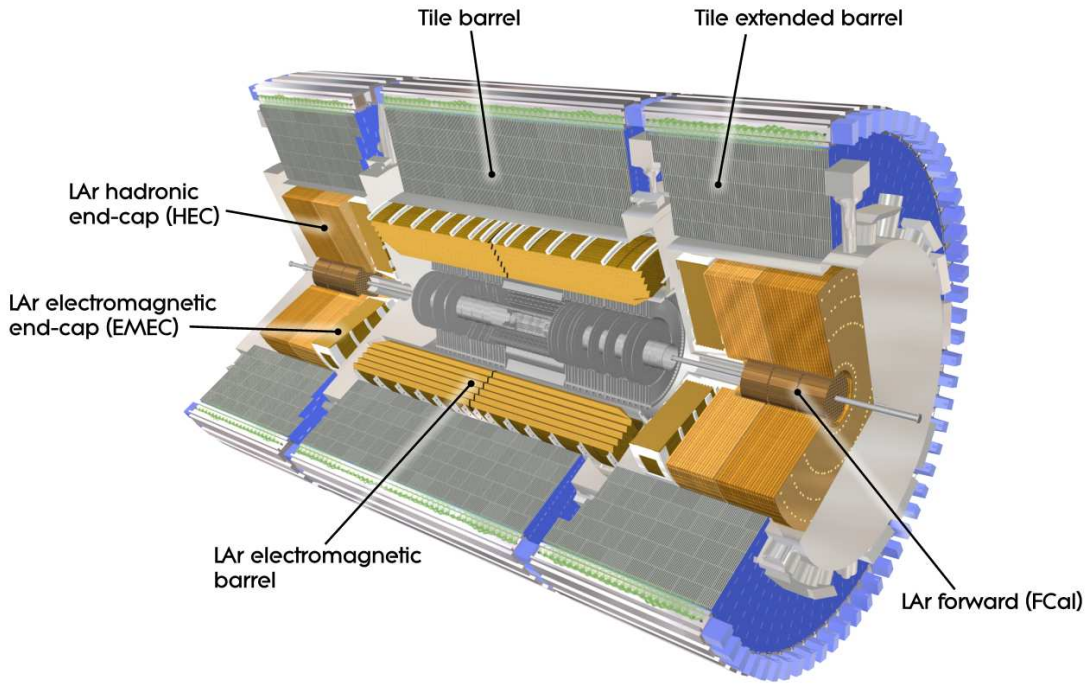


Figure 3.5: Cut-away view of the ATLAS calorimeter system [28].

The Calorimeter System

Like the inner detector, also the calorimeter system consists of several components designed to meet the requirements of measuring electrons, photons and jets with high efficiency and spatial and energy resolution. Figure 3.5 gives an overview of the ATLAS calorimeter system. All calorimeters are realized as sampling calorimeters and provide full solid angle coverage up to $|\eta| < 4.9$.

The innermost layer is the electromagnetic (EM) calorimeter consisting of a barrel part reaching up to $|\eta| < 1.5$ and two end-caps (EMEC) up to $|\eta| < 3.2$ complemented by two forward electromagnetic calorimeters (FCal1) in the region up to $|\eta| < 4.9$. All these calorimeters use liquid argon as active medium. The absorber material of the barrel electromagnetic and end-cap calorimeters are lead plates in accordion shape which allows for fast signals and uniform response. The electromagnetic FCal1 uses copper as absorber. To achieve high spatial and energy resolution for electrons and photons, the granularity in the η-region also covered by the inner detector ($|\eta| < 2.5$) is very fine ($\Delta\eta \times \Delta\phi = 0.025 \times 0.025$). In the first layer, the so-called η-strip layer, the η segmentation is even finer (down to $\Delta\eta = 0.003$). This is, for example, used for the identification of τ jets (see Section 5.2.5). For $|\eta| > 2.5$ the coarser granularity is sufficient for good jet reconstruction.

The EM calorimeters are surrounded by the hadron calorimeters consisting of a barrel tile calorimeter ($|\eta| < 1.0$), two extended barrel tile calorimeters ($0.8 < |\eta| < 1.7$), two hadronic end-cap calorimeters HEC ($1.5 < |\eta| < 3.2$) and the forward hadron calorimeters FCal2 and FCal3 ($3.2 < |\eta| < 4.8$).

The tile calorimeters use plastic scintillator as active material and steel as absorber material arranged in three layers with a granularity of $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$. The HEC calorimeters use liquid argon as active material and copper as absorber material with a granularity of $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ for $|\eta| < 2.5$ and $\Delta\eta \times \Delta\phi = 0.2 \times 0.2$ for $|\eta| > 2.5$. The FCal2 and FCal3 have a readout cell size of $\Delta x \times \Delta y$=3.3 × 4.2 cm^2 and $\Delta x \times \Delta y = 5.4 \times 4.7$ cm^2 respectively. Both use liquid argon as active and tungsten as absorber material. The latter has a shorter interaction length as copper.
The total thickness of the calorimeter system is more than 22 radiation lengths (X_0) and about ten hadronic interaction lengths. This ensures a good energy resolution of highly energetic jets, precise reconstruction of missing energy and minimizes punch-through of particles to the muon spectrometer. The total number of readout channels of the ATLAS calorimeter system is ~260 000.

The Muon Spectrometer

Figures 3.6 and 3.7 show the layout of the muon spectrometer. The outer dimensions of 44 m length and 25 m diameter make the ATLAS detector the largest detector ever built for a collider experiment. Three layers of muon detectors arranged in horizontal concentric cylinders in the barrel region and in vertical discs in the two end-caps measure the deflection of the muon tracks due to the magnetic field provided by the three toroid magnets.
The bending power $\int \mathrm{Bdl}$ along the tracks ranges from 1.5 to 5.5 Tm in the barrel part and from 1 to 7.5 Tm in the end-cap regions. To minimize multiple scattering, the magnetic field of the muon spectrometer is generated in air (air-core toroid magnets) and the support structure is made of aluminum.
To accurately measure muon tracks and provide a muon trigger, the muon spectrometer is instrumented with four detector types:

- Monitored Drift Tube chambers (MDT) are the precision muon tracking detectors over most of the muon spectrometer acceptance. 1 150 chambers contain in total 354 000 drift tubes covering an active area of 5 500 m^2. Each chamber consists of two multilayers of three or four tube layers that are glued on the cross-plates of an aluminum support structure as shown in Figure 3.8. The aluminum tubes of 30 mm diameter and 400 μm wall thickness are filled with an Ar(93%)CO_2(7%) gas mixture at a pressure of 3 bar. The gold-plated tungsten-rhenium anode wires in the tube centers are positioned with respect to the chambers with an accuracy of 20 μm. At a high voltage of 3 080 V, the maximum drift time is about 700 ns. The average spatial resolution of a drift tube is 80 μm. The track position resolution of the MDT chambers is 35 μm.

- Cathode Strip Chambers (CSC) are used as precision muon tracking chambers in the innermost layer of the very forward region ($2.0 < |\eta| < 2.7$). The 32 multi-wire proportional chambers with strip-segmented cathodes have a shorter response time than the MDT chambers to cope with the high background rates in this detector region.

- Resistive Plate Chambers (RPC) are used as trigger detectors in the barrel region ($|\eta| < 1.05$). They consist of two parallel electrode plates with a gap of 2 mm width filled with a $C_2H_2F_4$/Iso-C_4H_{10}/SF_6 gas mixture. The 544 RPCs are operated in avalanche mode with an electric field of ~ 4.9 kV/mm between the electrode plates. Capacitively coupled metallic

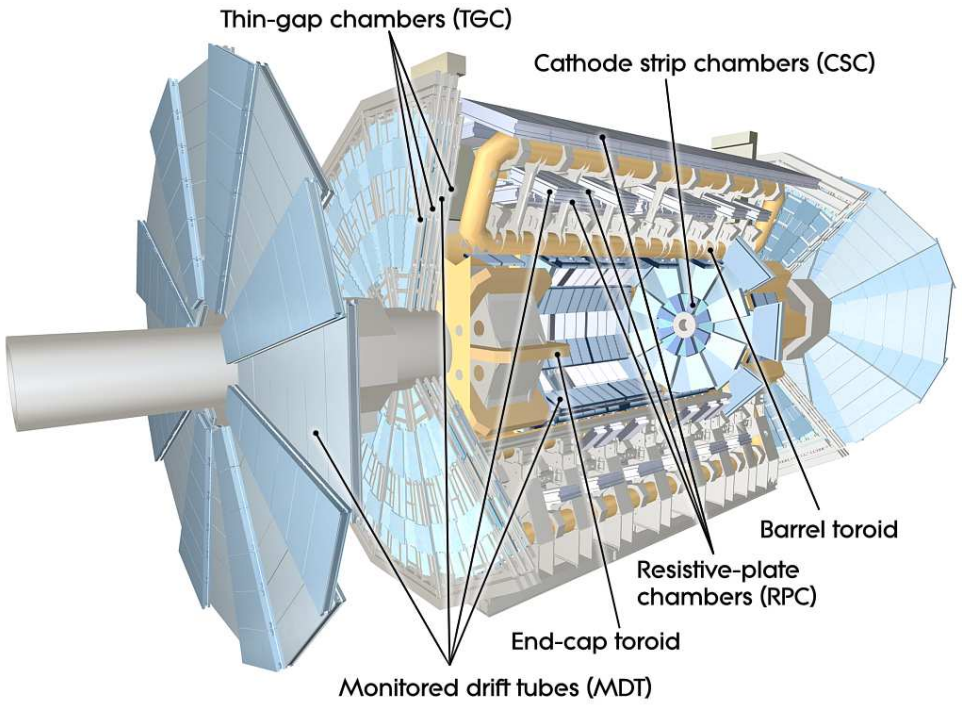


Figure 3.6: Cut-away view of the ATLAS Muon Spectrometer [28].

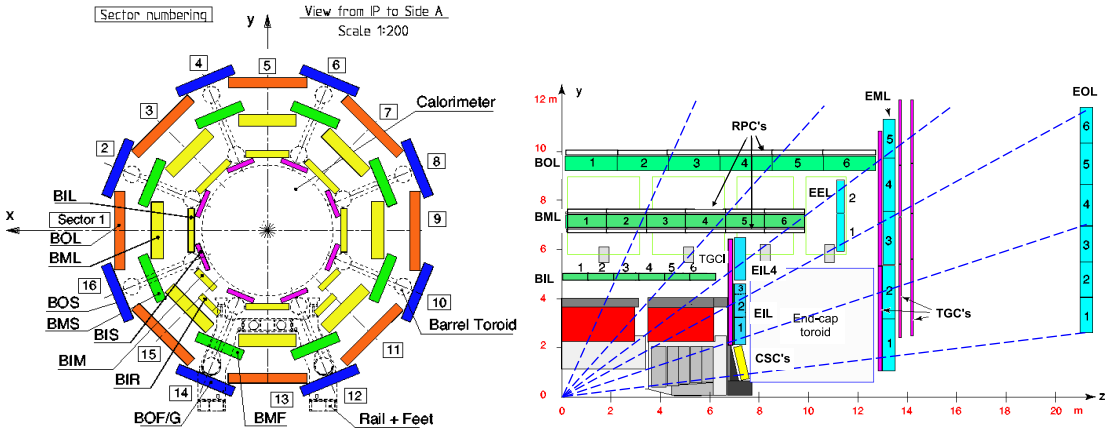


Figure 3.7: Cross-sections of the barrel muon system in transverse plane (left) and in longitudinal plane (right). On the latter, the dashed lines illustrate infinite-momentum muons, propagating along straight trajectories [28].

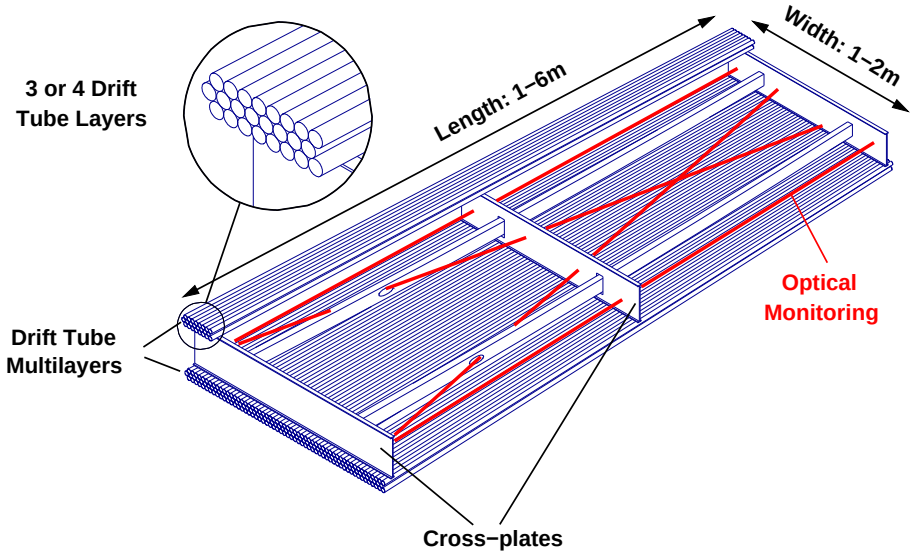


Figure 3.8: Layout of a MDT chamber. Three cross-plates connected by longitudinal beams form an aluminum support structure, carrying two multilayers of three or four drift tube layers. Four optical alignment rays monitor deformations of the chamber.

strips pick up the signals on the outside of the plates. The spatial resolution is about 30 mm and the time resolution is 1 ns.

- Thin Gap Chambers (TGC) are used as trigger detectors in the end-cap region ($1.5 < |\eta| < 2.7$). The 3 588 TGCs are multi-wire proportional chambers with a wire-to-cathode distance of 1.4 mm which is smaller than the wire-to-wire distance of 1.8 mm. They provide a high spatial resolution and a good time resolution with a strongly quenching gas mixture of CO_2 and n-C_5H_{12}.

The RPC and TGC trigger chambers provide bunch crossing identification and measure the coordinate along the drift tubes of the MDT chambers.

To achieve the high momentum resolution of $\Delta p_T / p_T \approx 3\,\%$ at 200 GeV and better than 10 % up to 1 TeV muon energy, a relative alignment of the precision-tracking chambers with an accuracy of 30 μm is necessary, especially for measuring high-p_T muons. More than 12 000 optical alignment sensors [31, 32] monitor the internal deformations of the MDT chambers (therefore the name "Monitored Drift Tube" chambers) as well as the relative positions of the MDT and CSC chambers.

Also the magnetic field map has to be determined very accurately. This is achieved with 1 800 Hall probes distributed throughout the muon spectrometer volume and a very detailed simulation of the magnetic field including perturbations induced by the calorimeters and other metallic structures.

The muon spectrometer provides stand-alone muon momentum measurement. A combination with measurements of the inner detector and the calorimeters improves the efficiency and resolution, especially for low p_T muons.

Trigger and Data Acquisition

A big challenge for every experiment at the LHC is the reduction of the extremely high event rate of approximately 1 GHz to a maximum data taking rate of approximately 200 Hz. The efficient selection of interesting events is the task of the trigger system. In ATLAS the trigger selection is performed in three levels:

1. The level-1 trigger system (L1) uses muon spectrometer (RPC and TGC) and calorimeter information to select events with high-p_{T} muons, electrons, photons and jets. In addition, the total transverse energy and the missing transverse energy are used as trigger criteria. If objects are found that exceed certain configurable p_{T} thresholds, regions of interest (RoI) are passed to the next trigger level with a maximum rate of 75 kHz.

2. The second trigger level (L2) searches again for signs of the above mentioned signals, but uses the full information of all detector components in the RoIs defined by the L1 trigger system. The trigger requirements are chosen in order to reduce the accepted event rate to 3.5 kHz.

3. The third and final stage (L3) of the trigger selection is called Event Filter (EF). While the L1 and L2 triggers are realized by custom made electronic circuits, the EF is implemented on a computer farm. Every processor receives the full detector information of one event selected by the L2 trigger and performs a full reconstruction of the event which takes about 4 seconds. The maximum rate of events accepted by the EF is 200 Hz.

All events accepted by the trigger are recorded on disk and can then be analyzed via the worldwide computing Grid [33]. Every year, several thousand terabyte of data have to be stored.

Chapter 4

Installation and Commissioning of the ATLAS Muon Chambers

The Max-Planck Institut für Physik in Munich together with the Ludwig-Maximilians University have built 88 precision muon tracking (MDT) chambers for the barrel part of the ATLAS muon spectrometer. To ensure reliable performance during the operation of the ATLAS detector, the MDT chambers had to pass several stringent tests during and after their construction as well as before their installation into the ATLAS detector. To simplify the installation, each MDT chamber was assembled with its corresponding RPC trigger chamber in a special frame, the common support. These units are called muon stations. Details about the tests before the installation and the integration of the MDTs to muon stations can be found in [34, 35, 36, 37, 38].
The installation of the muon stations in the ATLAS detector and the commissioning of the muon spectrometer with cosmic muons are described in the following.

4.1 Chamber Installation

The 88 muon detectors built in Munich are called BOS ("barrel outer small") chambers and have been installed during the first half of 2006. They are located in the outermost layer of the muon spectrometer, directly on the toroid coils, see Figure 4.2. Except for special cases, the chambers in the barrel part are installed on rails which are mounted on the cryostats of the toroid coils and are running in z-direction.
Each BOS muon station is equipped with four sliding bearings allowing to move it on the rails. To handle the muon station during installation, a dedicated installation frame was used. The station was mounted on the frame on short rail pieces. The installation frame with each individual muon station was then lowered by crane to the ATLAS cavern about 90 meters underground (see Figure 4.1a).
In the cavern, the installation tool was lifted simultaneously by two cranes and adjusted to the position and angle at which the chamber had to be installed (see Figure 4.1b). The rail pieces of the installation frame were connected to the rails in the detector by short transition rods (see Figure 4.1c). The station was then pulled from the installation frame to its foreseen position on the rails using two winches (Figure 4.1d).

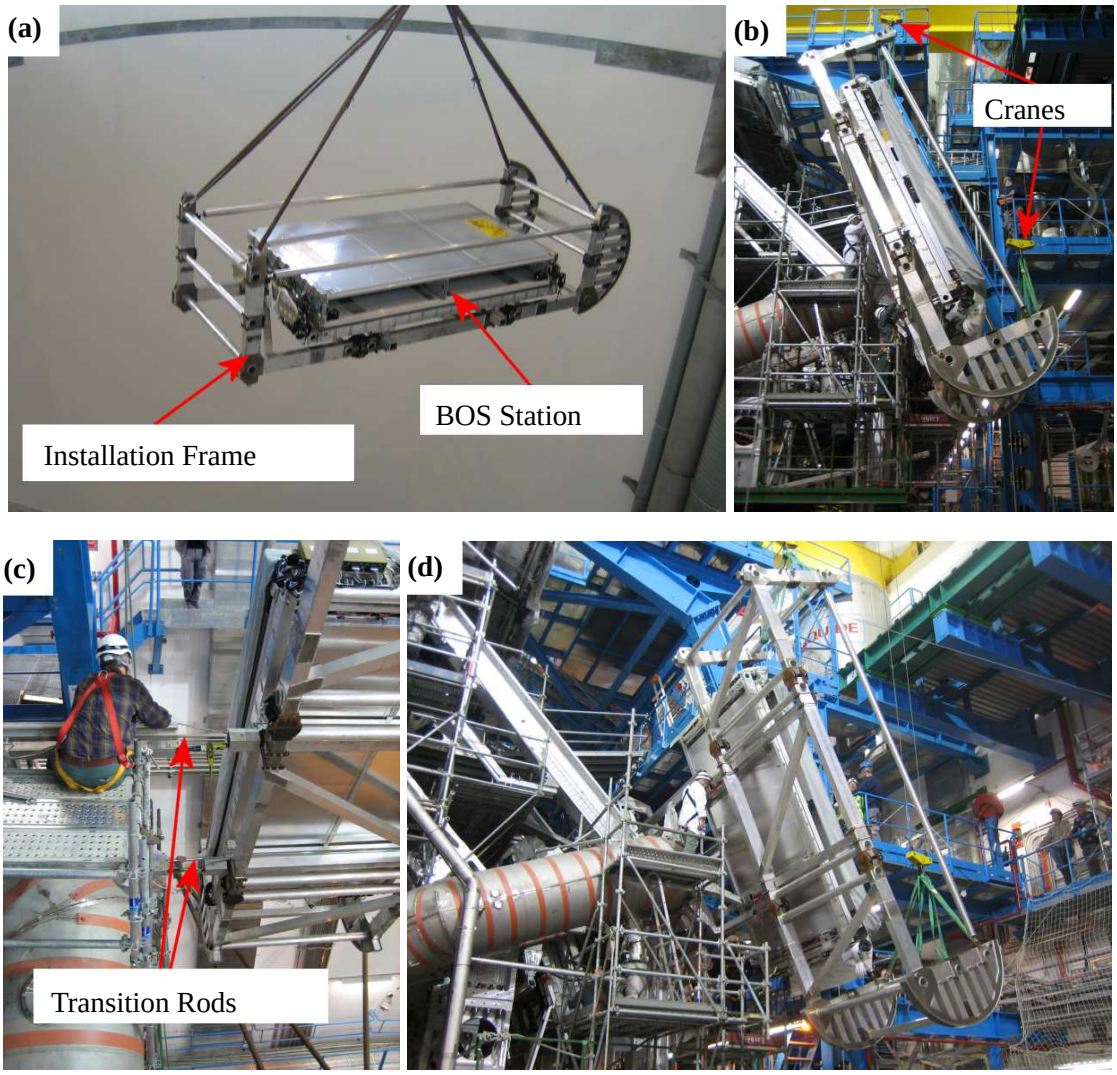


Figure 4.1: Pictures of the installation of a BOS muon station. (a) A BOS station in the installation frame is being lowered to the ATLAS cavern. (b) The chamber is adjusted to its installation position and angle using two independent cranes. (c) The connection between the installation frame and the rails mounted on the toroid coils is established. (d) The station is pulled onto the rails.

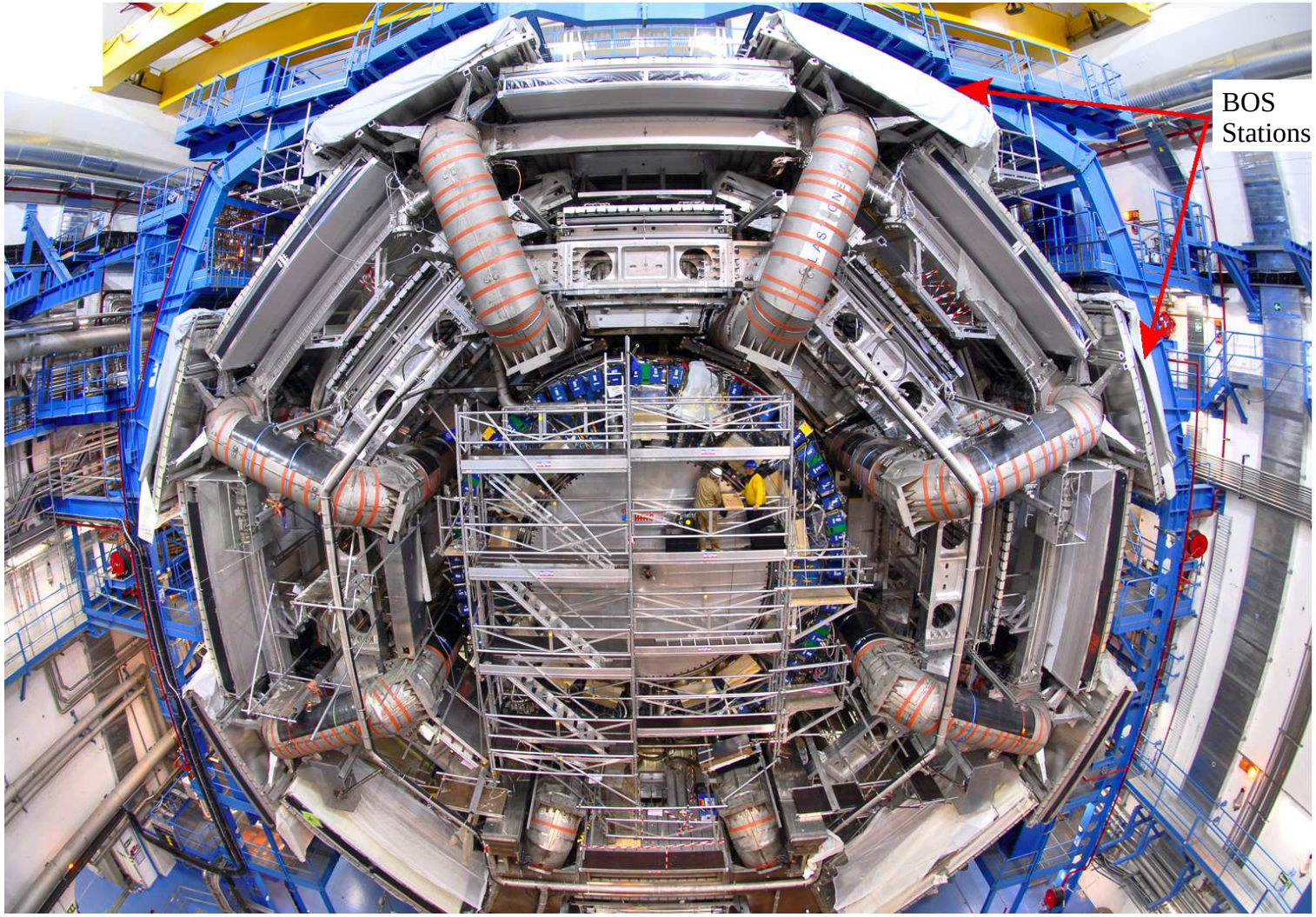


Figure 4.2: Front view of the barrel part of the ATLAS muon spectrometer after the installation of the muon chambers on the eight toroid coils. The BOS MDT chambers built in Munich are mounted in the outer layer of the muon spectrometer, directly on the toroid coils.

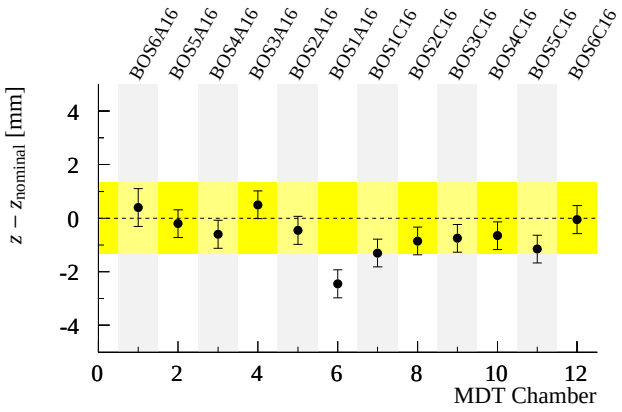


Figure 4.3: Deviation of the BOS MDT chamber positions from the nominal positions after the final positioning on the rails in sector 16 of the ATLAS barrel muon spectrometer. The shaded band indicates one third of the allowed tolerance of 4 mm.

The exact positioning of the muon stations on the rails and the adjustment of the MDT chamber position on the common support was performed by hand with an accuracy of one millimeter. Figure 4.3 shows the deviations of the BOS chamber positions from the nominal positions in the global ATLAS coordinate system in one typical barrel sector after the final positioning on the rails. All chambers are located well within the allowed range of ±4 mm. Special stoppers mounted on the rails are used to fix the chamber positions on the rails. Adjustment screws on the stoppers allow for a fine tuning of the chamber position along the rails. At the final position, each chamber is only fixed at one of the four bearings in order to prevent stress on the common support.

4.1.1 MDT Chamber Sag Adjustment

To avoid distortions of the electrical field inside the drift tubes which could deteriorate their spatial resolution, the anode wires must be centered in the tubes with an accuracy of about 100 μm. The gravitational sag of the drift tubes assembled in the MDT chambers and of the wires held at the end of the tubes differs and depends on the angular orientation of the chambers in ATLAS. To correct for this difference, the chamber sag has to be adjusted to the wire sag of typically 200 μm for the wire tension of 350 g.

For this purpose, the MDT chambers have a sag adjustment mechanism. With two adjustment screws, the middle cross-plate (see Figure 3.8) of the chamber support structure can be moved with respect to the end cross-plates in the direction perpendicular to the tube layers. Figure 4.4a illustrates the principle of the sag adjustment. The sag of the chamber support structure is measured by the in-plane alignment system with an accuracy of about 2 μm.

The sag adjustment of the BOS MDT chambers was already performed during the integration of the chamber into the common support at the BB5 area at CERN, located about 3 km from the ATLAS experiment. Due to vibrations during the transport of the chambers from the BB5 assembly hall to the ATLAS cavern and during the installation, the sag of the chamber could change. Therefore, the sag was measured again after the installation in the detector and deviations

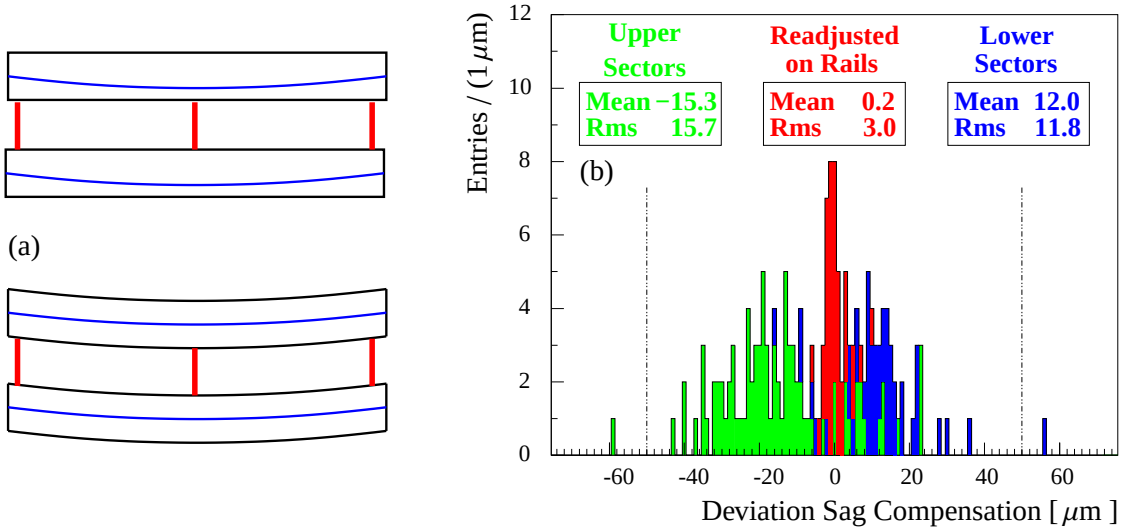


Figure 4.4: Adjustment of the gravitational sag of the MDT chambers to the sag of the anode wires which depends on the installation angle. (a) The principle of how the sag is corrected. (b) The deviation of the achieved sag adjustment from the nominal value for the MDT chambers built in Munich, measured after installation in the ATLAS detector with the chamber internal in-plane alignment sensors. Three categories are shown: chambers that have been readjusted after the installation, chambers installed in the upper sectors of the muon spectrometer and chambers of the lower sectors in which the chambers are mounted in an inverted orientation. The dashed lines indicate half of the allowed tolerance of 100 μm.

from the nominal sag have been corrected for the accessible chambers.
Figure 4.4b shows the results of the sag adjustment of the installed BOS muon chambers. All adjustments are well within the allowed tolerance of $\pm 100\,\mu$m. It can also be seen that the deviations of the chamber sags in the upper and lower sectors are of opposite sign. An explanation for this effect is settling of the sag adjustment screws due to vibrations during the transport of the chambers. In contrast to the upper sector chambers, the chambers for the lower sectors were transported in their installation orientation upside down explaining the opposite sign of the sag deviations.

4.1.2 Tests After Chamber Installation

The following performance tests of the MDT chambers have been performed immediately after their installation in order to detect damages which could have occurred during the installation:

- Measurement of the gas leak rate of each multilayer to detect damages to the gas system. All 88 BOS MDT chambers fulfill the leak rate requirement within the measurement errors. The largest measured leak rates are only a factor 1.5 above the limit.

- Measurement of the high voltage stability in order to detect broken wires causing short-circuits between the wires and the tube walls or other faults in the high voltage distribution to the drift tubes. No problems have been found concerning the stability of the voltage and the currents drawn by the chambers.

After the electrical and gas connections were established, a series of further tests of the MDT chambers has been performed:

- Test of the chamber-internal and the chamber-to-chamber alignment systems. No faults have been detected.

- Test of the temperature and magnetic field sensors which are mounted on the MDT chambers. All magnetic field sensors and $\geq$99.5 % of the temperature sensors are operational.

- Test of the complete chamber read-out chain including the programming of the on-chamber electronics. Less than 1 percent of the electronics showed errors and has been replaced.

- Test of the noise rate of all drift tubes with and without high voltage applied. Several BOS chambers showed a significantly higher noise level as before installation. Further investigation revealed that the noise is due to pick-up on the high voltage cables. Additional low-pass filters have been installed and the chambers are now well below the noise limits [39].

4.2 Commissioning of the Muon Spectrometer with Cosmic Muons

The performance of individual muon detectors as well as of the muon spectrometer as a whole can be studied by means of cosmic ray muons impinging on the ATLAS detector. The measurements of cosmic muons started right after the installation of the muon stations and were extended successively. Since the final gas distribution system was not yet operational and the number of power

supplies was very limited, the number of muon detectors ready for data-taking initially was very small.
At the end of 2005, the first six muon chambers were ready for operation. In November 2006, 13 muon chambers were operated when the barrel toroid was ramped up to nominal field for the first time, allowing for the measurement of muon momenta. Finally in 2008, the whole muon spectrometer (like the other detector components) was ready to take data and more than 200 million cosmic muons were recorded. In the following, the most important results obtained from these data are presented.

4.2.1 Drift Tube Efficiency Measurement

Cosmic muons can be used to determine the detection efficiency of individual drift tubes in MDT chambers by the following algorithm:

- Muon tracks are reconstructed within a MDT chamber using all drift tubes except for the tube under investigation.
- The muon tracks are extrapolated to the tube under investigation.
- The detection efficiency ε of the tube under investigation is given by:

$$\varepsilon = \frac{\text{Number of hits detected by the tube}}{\text{Number of reconstructed tracks traversing the tube}}$$

Figure 4.5 shows a typical result for one tube layer of a MDT chamber. The drift tubes show an average detection efficiency of $97.7 \pm 0.1\,\%$, which is consistent with the expected efficiency loss due to the $400\,\mu$m thick aluminum walls of the 15 mm diameter drift tubes.
Only 1 ‰ of the 338 640 drift tubes of the muon spectrometer had to be permanently disconnected from high voltage and read out due to broken wires or gas leaks.

4.2.2 Reconstruction of Cosmic Muon Tracks

The ATLAS cavern is situated about 90 meters below ground level. Thus, cosmic muons with energies below $\sim$30 GeV at ground level cannot reach the ATLAS cavern but are absorbed in the rock and soil above. However, muons flying through one of the two access shafts with diameters of 13 m and 18 m reach the detector almost unrestrained. Thus, most cosmic muons recorded in ATLAS are constrained into the direction of the access shafts (see Figure 4.6b).
This was confirmed at the end of 2006 when 13 muon stations were operational allowing for the first reconstruction of cosmic muon tracks in the ATLAS muon spectrometer. Figure 4.6a shows the extrapolation of reconstructed cosmic muon tracks back to the surface. A clear accumulation of tracks on the right part of the plot, as well as a small cluster on the left part are visible. These accumulations agree very well with the positions of the two access shafts.

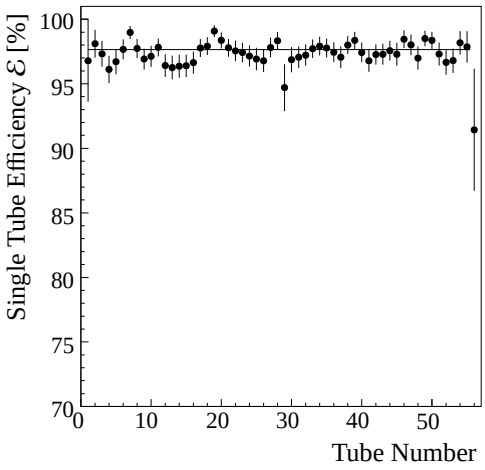


Figure 4.5: Single-tube efficiencies in one tube layer of a MDT chamber installed in the ATLAS detector determined with cosmic muons.

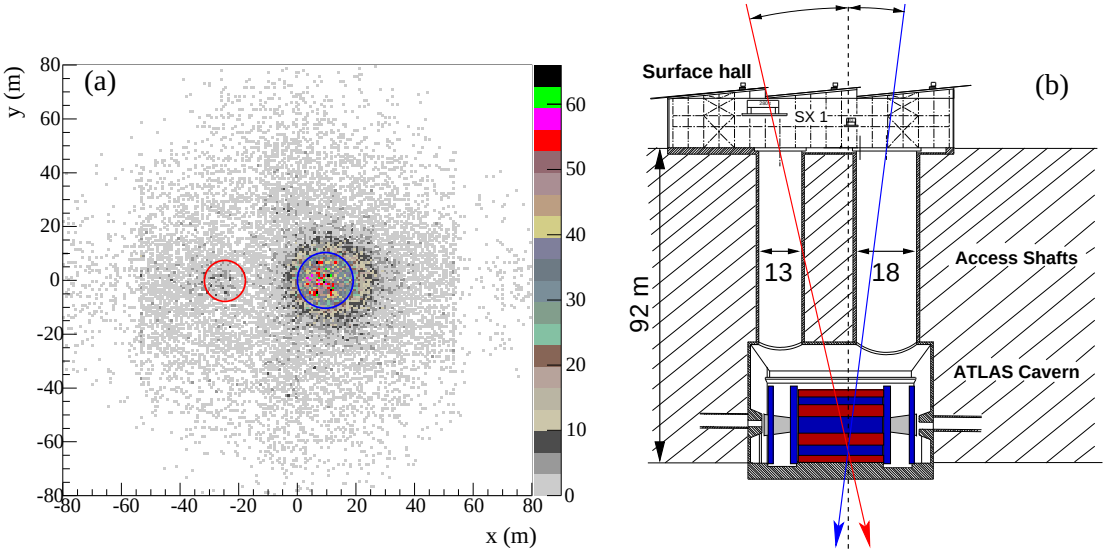


Figure 4.6: Cosmic muons tracks recorded in the ATLAS muon spectrometer. (a) Distribution of the intersections of reconstructed cosmic muon tracks at ground level. Here x and y denote a coordinate system on the surface. The two accumulations of track points are due to the access shafts of 13 and 18 m diameter through which cosmic muons reach the ATLAS cavern (b).

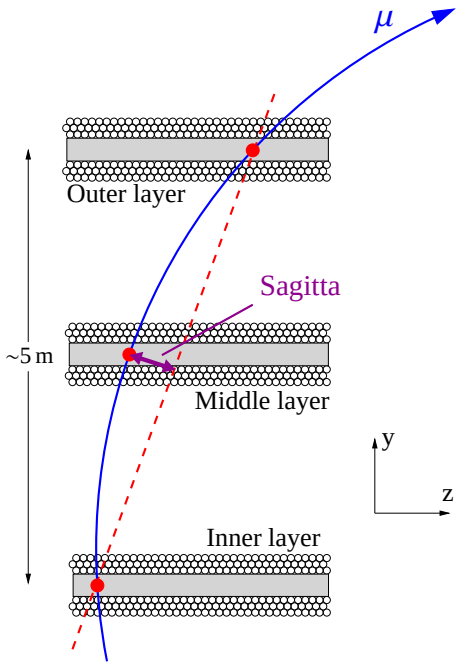


Figure 4.7: Illustration of the track sagitta measurement performed in the ATLAS barrel muon spectrometer.

4.2.3 Alignment with Straight Muon Tracks

The muon momentum measurement in the ATLAS muon spectrometer is based on the measurement of the deflection of a muon track in the magnetic field as illustrated in Figure 4.7. Three layers of MDT chambers measure three track points, one at the entrance of the spectrometer, one in the middle and one at the exit. The deviation of the middle track point from a straight line connecting the inner and outer track points is the track sagitta s. The sagitta is inversely proportional to the muon momentum.

To measure the momentum with the desired resolution of 10 % for 1 TeV muon tracks with a sagitta of $s \approx 500\,\mu$m, the sagitta resolution has to be $\sim 50\,\mu$m. The track point resolution of the MDT chambers is $35\,\mu$m contributing $40\,\mu$m to the sagitta resolution.

Another contribution to the sagitta error is due to misalignment of the MDT chambers with respect to each other which is controlled by the optical alignment monitoring system. Thus the precision of the alignment corrections on the track sagitta has to be $\sim 30\,\mu$m.

The optical alignment system of the ATLAS muon spectrometer monitors *relative* movements of muon chambers with a precision of $\sim 10\,\mu$m. Since the chamber positioning accuracy during installation is in the order of millimeters, the *absolute* positions of the MDT chambers have to be determined initially with high accuracy using straight muon tracks. Dedicated data taking at the beginning of the LHC operation with the toroid magnets switched off will provide enough straight muon tracks to align all muon chambers. However, the upper and lower sectors of the spectrometer

can already be aligned with straight tracks of cosmic muons with switched off magnetic field. For the other sectors horizontal muon tracks are needed, but the rate of horizontal cosmic muons in the ATLAS cavern is too low.
The relative chamber positions within the complete top sector of the barrel muon spectrometer (see Figure 3.7) have been determined with a data sample of 400 000 straight cosmic muons using a linearized χ^2-minimization alignment algorithm [40, 41, 39]. The distances in z between adjacent MDT chambers of the inner and outer layer of the top sector obtained from the track alignment algorithm can be compared to measurements of the distances between the tube walls of adjacent chambers using a feeler gauge which have an accuracy of about 50 μm (see Figure 4.8). A clear correlation between the two measurements can be seen. For the outer chambers, the measurements agree within 85 μm while for the inner chamber layer a relative shift of 190 μm is observed which can be explained by the less accurate method used for the mechanical distance measurement between the inner chambers. The deviations in the order of ±1-2 mm from the nominal distance of 17.5 mm are consistent with the expectation from the chamber positioning accuracy on the rails (see for example Figure 4.3).
Another way to verify the alignment procedure is by looking at the sagitta distribution of straight tracks. In the absence of a magnetic field, the muon tracks are straight and a misalignment of the MDT chambers will cause a deviation of the sagitta measurement from zero (false sagitta). Figure 4.9a shows the distribution of the false sagitta of cosmic muons measured in a chamber triplet in the top sector of the muon spectrometer without magnetic field and assuming nominal detector geometry in the reconstruction. The distribution peaks at −2.6 mm which is consistent with the expected deviations of the actual chamber positions from the nominal geometry. If the chamber displacements measured by the optical alignment system (using a preliminary calibration) are taken into account in the sagitta calculation, the peak of the distribution is shifted to +0.25 mm (see Figure 4.9b). Using the updated information of the chamber positions from the straight track alignment algorithm, the measured sagitta distribution peaks at −10 μm (see Figure 4.9c) which is well within the required accuracy of ±30 μm.
The width of the distributions is getting smaller if the alignment corrections are applied. This can be explained by taking into account rotations of the chambers measured by the optical and straight track alignment algorithm. The remaining width of about half a millimeter and the tails of the sagitta distribution of aligned chambers (Figure 4.9c) can be explained by deviations from straight tracks due to multiple scattering of the muons.

4.2.4 Curved Muon Tracks

An important milestone in the commissioning of the muon spectrometer was reached when the barrel toroid magnet was switched on for the first time after assembly in the underground cavern. On 18 November 2006, the barrel toroid was successfully ramped up to the nominal current of 20.5 kA. Thus, the 13 muon chambers fully operational at this time could record the first curved cosmic muon tracks. Two examples are shown in Figure 4.10 in the ATLAS event display: on the left-hand side a low-momentum muon track with $p_{\mathrm{T}} = 1.6$ GeV and on the right-hand side a high-momentum muon with $p_{\mathrm{T}} \sim 200$ GeV.
In Figure 4.11 the momentum distributions of the recorded cosmic muons and antimuons are shown. The observed ratio of the two charge components $N_{\mu^+}/N_{\mu^-} = 1.48 \pm 0.27$ [42] is com-

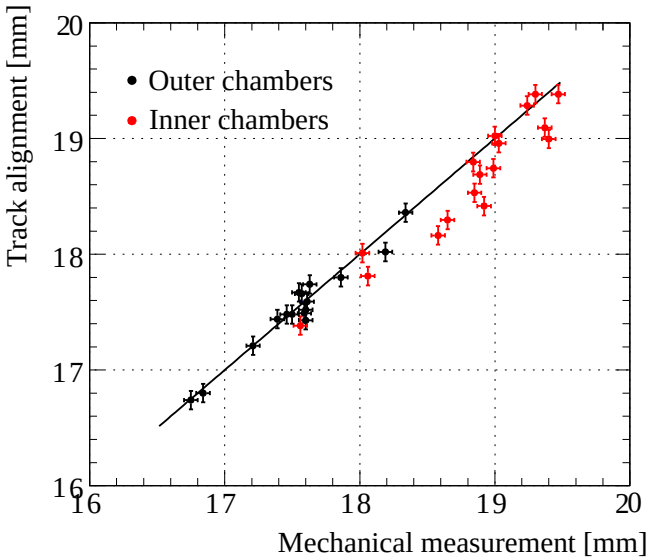


Figure 4.8: Comparison of the distances between adjacent chambers in the inner and outer layer of the top sector of the muon spectrometer as determined by mechanical measurements with a feeler gauge and by the straight track alignment algorithm [41]. The nominal distance is 17.5 mm.

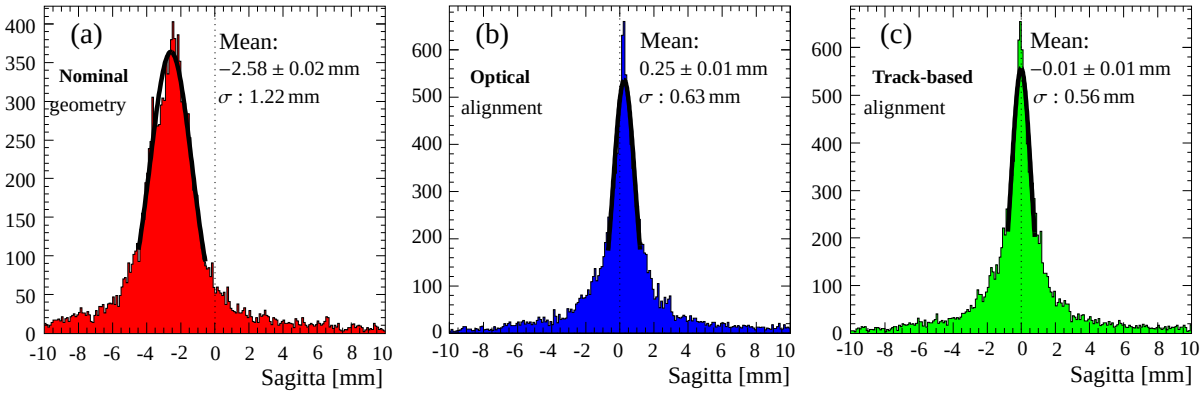


Figure 4.9: Distributions of the measured sagitta of straight cosmic muon tracks in the ATLAS muon spectrometer without magnetic field (a) without applying corrections to the actual muon chamber positions after installation, (b) after taking into account the information of the optical alignment system and (c) after applying misalignment corrections determined with the alignment algorithm using straight cosmic muon tracks [41].

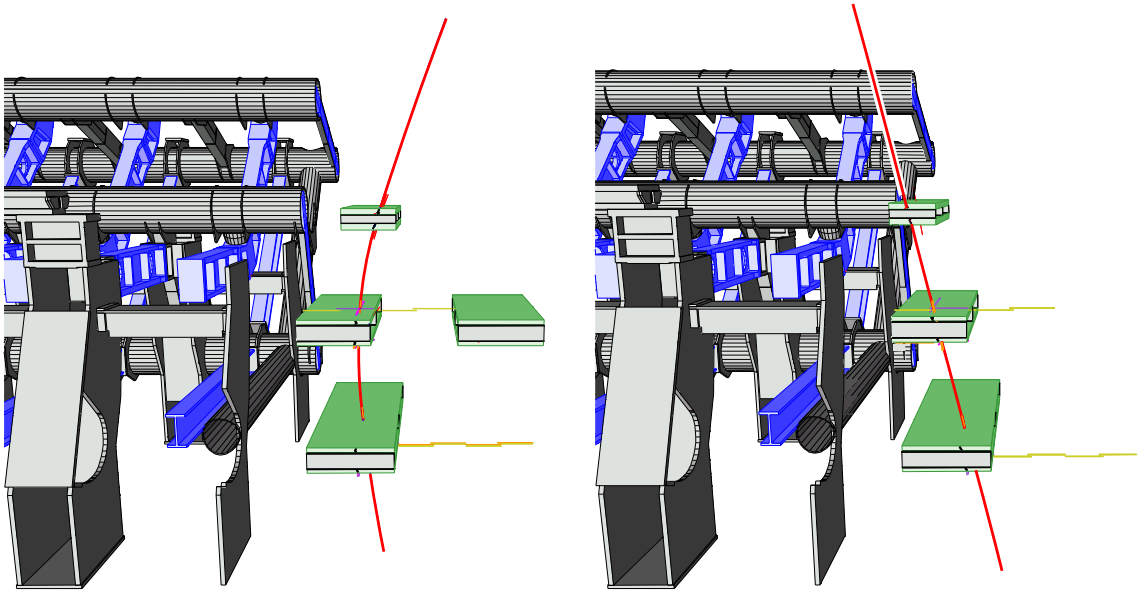


Figure 4.10: Display of events with reconstructed low momentum (left) and high momentum cosmic muons (right). The muons have been recorded by chambers in the bottom sector of the muon spectrometer when the toroid magnet has been switched on for the first time in November 2006. The coils and the support structure of the barrel toroid magnet are shown in gray and blue respectively. The muon chambers traversed by the muon tracks are shown in green and the reconstructed muon tracks are represented by the red curves.

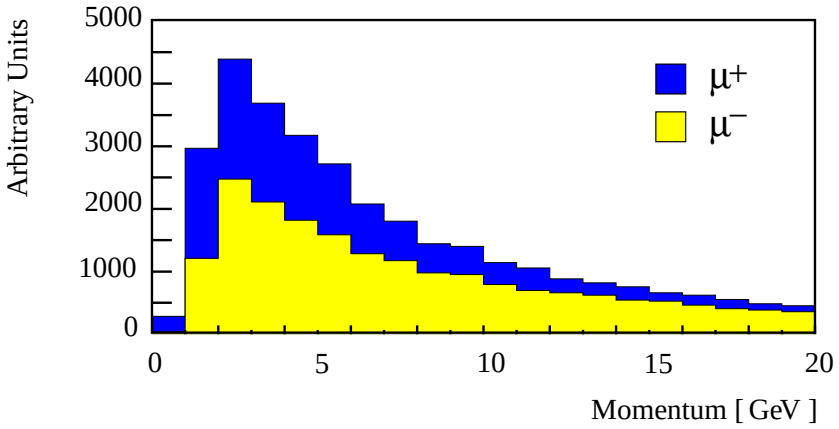


Figure 4.11: Cosmic muon momentum spectrum recorded with 13 muon chambers in the bottom sector of the ATLAS muon spectrometer in November 2006, shown for muons and antimouns. The ratio N_{μ^+}/N_{μ^-} is 1.48 $\pm$ 0.27 [42].

patible with the ratio expected for muons produced in cosmic ray induced air showers [22]. The measured momentum spectrum was not corrected for misalignment effects of the muon chambers.

Chapter 5

The Search for the Higgs Boson

This chapter introduces the ingredients needed for the search for the Higgs boson produced by vector-boson fusion and subsequently decaying into two τ leptons. In this work the fully leptonic as well as the semileptonic final state of the further decaying τ-lepton pair is taken into account. Characteristic features of the studied search channel and the associated background processes as well as the used data samples are described in Section 5.1. The expected reconstruction and identification performance of the objects relevant for the Higgs boson search is discussed in Section 5.2. Although neutrinos from the τ-lepton decays are present in the final state, it is still possible to reconstruct the Higgs boson mass by the collinear approximation which is described in Section 5.3. Section 5.4 finally gives a detailed description of various discriminating variables that can be used in order to distinguish signal and background events.

In the subsequent chapters, two approaches for separating signal and background processes are shown – an analysis based on successive cuts on several discriminating variables (Chapter 6) and an approach based on multivariate techniques that takes into account correlations between different discriminating variables (Chapter 7).
The prospects of discovering the Higgs boson after three years of ATLAS operation are discussed in Chapter 8 which mainly focusses on the influence of systematic uncertainties from theory and the detector performance on the expected signal significance.
Several previous studies have investigated the Higgs boson discovery potential in the vector-boson fusion $H \rightarrow \tau\tau$ channel with a cut-based analysis [43, 44, 45, 46, 47, 48]. The discovery potential of the ATLAS experiment has recently been studied in the context of the "computing system commissioning" (CSC) effort, a test of the worldwide data processing infrastructure which was also used to perform large scale simulations and physics analyses [49]. A cut-based analysis in analogy to this study repeatedly serves as a reference for comparison and is referred to as *baseline analysis* throughout this work.

5.1 Signal and Background Processes

The Higgs bosons produced through vector-boson fusion (see Section 2.2.2) are accompanied by two characteristic jets in the forward region of the detector, the so-called tagging jets. These

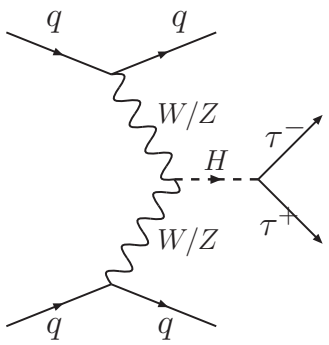


Figure 5.1: Tree-level Feynman diagram of the Higgs-boson production via vector-boson fusion with subsequent decay of the Higgs boson into a τ-lepton pair. The two outgoing quarks form characteristic jets in the forward region of the detector, which allow for a strong suppression of most of the background processes.

Table 5.1: Summary of the τ-lepton decay modes and the corresponding branching ratios. [22]

	τ^- decay mode	Branching ratio [%]
Leptonic decay modes	$e^-\overline{\nu}_e\nu_\tau$	17.84 ± 0.05
	$\mu^-\overline{\nu}_\mu\nu_\tau$	17.36 ± 0.05
Hadronic decay modes	$\pi^-\nu_\tau$	10.90 ± 0.07
	$\pi^-\pi^0\nu_\tau$	25.50 ± 0.10
	$\pi^-\pi^+\pi^-\nu_\tau$	8.99 ± 0.08
	$\pi^-2\pi^0\nu_\tau$	9.25 ± 0.12
	other modes	10.16

forward jets originate from the remnants of two quarks after they have emitted the weak bosons and provide an excellent signature to identify signal events. Furthermore, mostly no additional jet activity but only the Higgs decay products are confined to the central detector region. Solely this characteristic signature of the vector-boson fusion production mode makes it possible to sufficiently suppress the background processes.

A Feynman diagram of the signal process is shown in Figure 5.1. The decay of the Higgs boson into two τ leptons is followed by the τ-leptons decaying to electrons, muons or hadrons. Table 5.1 shows the main decay modes of a τ lepton and the corresponding branching ratios. Correspondingly, three final states can be distinguished in the $H \to \tau\tau$ decay. The purely leptonic ($\ell\ell$) decay channel, where both τ leptons decay into an electron or a muon, with a probability of 12.4 %, the semileptonic (ℓh) decay channel, where one τ lepton decays hadronically and the other one leptonically, with a probability of 45.6 %, and the hadronic (hh) decay mode, where both τ leptons decay to hadrons, with a probability of 42.0 %. In this work, only the leptonic and the semileptonic decay channel have been studied. The fully hadronic mode is experimentally extremely challenging due to the large QCD background and requires an even better understanding of the detector performance.

Due to the neutrinos present in the τ-lepton decays, the signal signatures are characterized by a significant amount of missing transverse energy.

Every process with two or more jets and two τ-lepton decay products in the final state is a potential background. Monte Carlo studies show that the most important background contributions are caused by Z + jets and $t\bar{t}$ production. In addition, W + jets and QCD multijet events can be misidentified as signal. The mentioned processes are discussed in detail in the following:

- Z + jets
  One of the most important backgrounds is the inclusive Z boson production, which can be divided into two mechanisms: QCD production through the interaction of quarks (Figure 5.2c) and electroweak processes (Figures 5.2a and 5.2b).

  The Z boson decays further into a lepton pair (ee, $\mu\mu$ or $\tau\tau$, each with a branching ratio of ~3.4 %), into $\nu\bar{\nu}$ pairs (20.0 %) or into hadrons (69.91 %). Decays of the Z boson into neutrinos or hadrons do not contribute to the background, since no charged leptons are present in these cases. The decays into electrons or muons have no neutrinos in the final state and can thus be suppressed very efficiently by requiring a certain amount of missing transverse energy. Conclusively, only $Z \to \tau\tau$ events significantly contribute to the background.

  The event topology in the QCD production mode and in some electroweak processes (for example Figure 5.2b) differs from vector-boson fusion Higgs production mainly due to the different jet kinematics, providing a good suppression of these backgrounds. Nevertheless, electroweak production processes like in Figure 5.2a are almost indistinguishable from the signal (see Figure 5.1). However, since the cross-section for QCD Z production (2.03 nb) is more than 1 000 times higher than for electroweak production (1.7 pb), QCD Z + jets remains the main background source.

- $t\bar{t}$
  The Feynman diagram of the top-pair production, which is the second most important background, is shown in Figure 5.2d. Each top quark decays into a W boson and a b quark with a branching ratio of almost 100 %. The two b jets from the b quarks can give a similar signature as the tagging jets. The electrons, muons or τ leptons originating from the decays of W bosons can be misidentified as Higgs decay products. The branching ratio of the two W boson decays into leptonic final states ($WW \to (e/\mu/\tau)(e/\mu/\tau) \to ee\nu\nu, e\mu\nu\nu, \mu\mu\nu\nu$) is 6.4 %, while it is only 3.7 % for decays into a semileptonic final state ($WW \to (e/\mu/\tau)(e/\mu/\tau) \to eh\nu, \mu h\nu$, where h denotes a τ-jet). An additional contribution to the semileptonic channel can be present due to jets that are misidentified as τ jets. Therefore, a good rejection of misidentified τ jets is very important in order to suppress $t\bar{t}$ events. In any case, the decay of the W bosons into leptons involves neutrinos, resulting in missing transverse energy. In contrast to $Z \to \ell\ell$ events, the contribution from leptonic W decays can thus not be suppressed by cutting on a minimal missing transverse energy. Due to this, the $t\bar{t}$ process is a more important background contribution to the leptonic decay channel than to the semileptonic decay.
  Even though the kinematics of these events is different to the one of the signal, $t\bar{t}$ is one of the main background sources due to the large cross-section of 833 pb.

- W + jets
  The production of a single W boson with additional jets in the event is another potential background, in particular to the semileptonic decay mode. A signature similar to the signal is given by the W decays into electrons or muons, with an additional jet in the event faking the τ jet. In case two more jets are present in the event, they can be misidentified as tagging jets. Thus, at least three additional jets are needed to contribute to the background. The W + jets background can be significantly reduced by a cut on the transverse mass $m_{\mathrm{T}}^{\ell\nu}$ (see Section 5.4.1) of the lepton and the missing transverse energy $E_{\mathrm{T}}^{\mathrm{miss}}$. A Feynman diagram of the W + jets production is shown in Figure 5.2e.

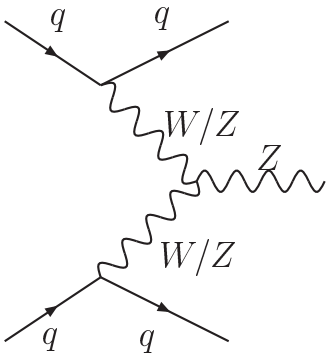
- WW+ jets
  As shown in Figure 5.2f, the production of a WW pair can also provide a signature which is quite similar to the signal. However, the cross-section of this process (112 pb) is ~20 times lower than the one for the Z boson production. Also the kinematic properties differ significantly from the signal, since the neutrinos are not collinearly emitted with the W bosons. Correspondingly, the WW + jets background contribution is much less important than the Z + jets production and has been neglected in this study.

- QCD
  Since the LHC is a proton-proton collider, the cross-section for multijet production is very high. Fortunately, no isolated leptons are present in these events such that the QCD events are strongly suppressed by the requirement of two τ-decay products in the final state. Potential background is caused by events containing misidentified leptons or leptons from b meson decays. Therefore, an excellent suppression of misidentified electrons, muons and τ jets is crucial for a discovery of the Higgs boson in this channel. The remaining $b\bar{b}$ background containing real leptons from b decays can be efficiently suppressed by requiring that the leptons are isolated.

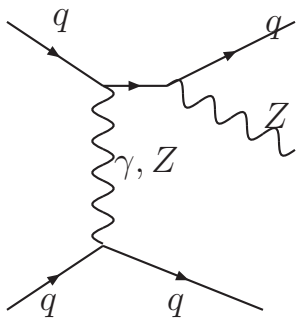
5.1.1 Monte Carlo Simulation

Several different Monte Carlo event generators have been used to simulate the signal and background processes. The simulated data samples and the corresponding cross-sections are summarized in Tables 5.2, 5.3, 5.4 and 5.5. All datasets have been produced in the frame of the central ATLAS Monte Carlo production with the Athena software framework (release 12.0.6) [50].

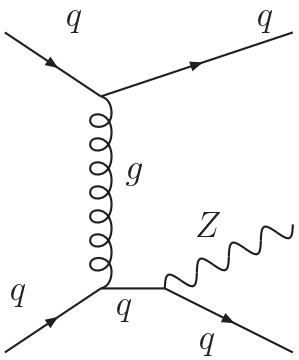
The events generated by the Monte Carlo generators are passed through the detailed simulation of the ATLAS detector based on the GEANT4 [51] package for the description of the detector response. In case of the $t\bar{t}$ background, the detailed detector simulation has been exchanged by the parameterized and thus faster simulation of the detector response by means of the ATLFAST [52] package. This was the only possibility to obtain a sufficient number of events needed for statistically significant analysis results.



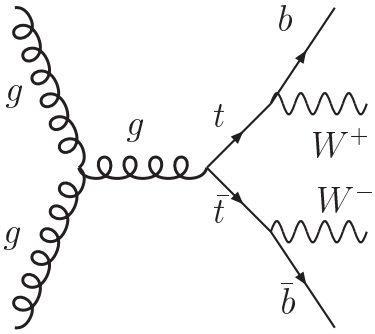
(a) Electroweak Z + jets production via gauge boson fusion.



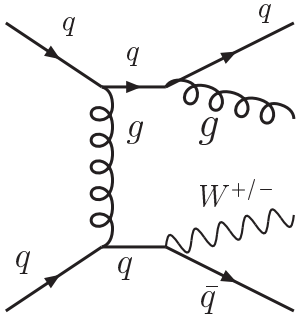
(b) Electroweak Z + jets production by Z boson radiation.



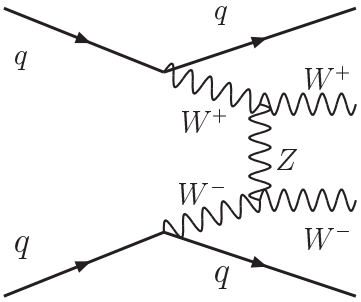
(c) QCD Z + jets production.



(d) $t\bar{t}$ production.



(e) W + jets production by W radiation.



(f) WW + jets production by triple electroweak gauge boson interaction.

Figure 5.2: Tree-level Feynman diagrams of the dominant background processes.

Table 5.2: Overview of the signal Monte Carlo data samples used for this study. All signal datasets include a lepton filter (see text) and are simulated with the detailed detector simulation. The cross–sections include the filter efficiency and are scaled to the next-to-leading order with the specified k-factor and include all branching ratios and filter efficiencies. The last column shows the integrated luminosity the generated numbers of events correspond to.

	k-Factor	Lepton Filter Efficiency [%]	Cross Section [fb]	Events	Integrated Luminosity [fb^{-1}]
$H \to \tau\tau \to \ell\ell$					
$m_H = 105\,\text{GeV}$	1.05	53	26	26 000	1 000
$m_H = 110\,\text{GeV}$	1.04	54	25	27 250	1 090
$m_H = 115\,\text{GeV}$	1.00	57	24	28 250	1 180
$m_H = 120\,\text{GeV}$	1.05	57	22	86 750	3 940
$m_H = 125\,\text{GeV}$	1.04	58	19	25 250	1 330
$m_H = 130\,\text{GeV}$	1.05	57	16	25 750	1 610
$m_H = 135\,\text{GeV}$	1.04	58	13	27 000	2 080
$H \to \tau\tau \to \ell h$					
$m_H = 105\,\text{GeV}$	1.05	46	180	67 500	380
$m_H = 110\,\text{GeV}$	1.04	47	173	84 750	490
$m_H = 115\,\text{GeV}$	1.00	47	160	28 250	180
$m_H = 120\,\text{GeV}$	1.05	47	145	132 250	910
$m_H = 125\,\text{GeV}$	1.04	48	127	86 500	680
$m_H = 130\,\text{GeV}$	1.05	48	108	26 500	250
$m_H = 135\,\text{GeV}$	1.04	48	87	85 250	980

Table 5.3: Overview of the Z + jets Monte Carlo data samples used in this analysis. The second column shows the vector-boson fusion filter efficiency, if applied (see text). The cross-sections include all branching ratios and filter efficiencies and are scaled to the next-to-leading order with a k-factor of 1.24. The last column shows the integrated luminosity the generated numbers of events correspond to. The different Monte Carlo programs are explained in the text.

	VBF Filter Efficiency [%]	Cross-section [pb]	Events	Luminosity [fb^{-1}]
ALPGEN QCD samples:				
$Z \to \tau\tau \to \ell h$				
+ 0 jets	0.3	2.75	11 250	4.1
+ 1 jet	1.0	2.45	92 250	37.7
+ 2 jets	4.0	3.10	515 500	166.0
+ 3 jets	10.6	2.50	356 000	142.0
+ 4 jets	20.7	1.38	218 750	158.0
+ ≥5 jets	29.6	0.73	90 000	122.0
$Z \to \tau\tau \to \ell\ell$				
+ 0 jets	0.04	0.32	9 850	31.0
+ 1 jet	0.13	0.28	161 800	574.0
+ 2 jets	0.43	0.33	439 900	1 340.0
+ 3 jets	1.12	0.26	156 000	590.0
+ 4 jets	2.08	0.14	172 100	1 230.0
+ ≥5 jets	3.30	0.08	48 450	599.0
$Z \to \tau\tau$, AtlFast				
+ 0 jets	–	897	3 760 000	4.2
+ 1 jet	–	226	3 885 000	17.1
+ 2 jets	–	76.5	3 955 000	51.7
+ 3 jets	–	23.6	1 000 000	42.3
+ 4 jets	–	6.68	500 000	74.8
+ ≥5 jets	–	2.48	395 000	159.0
$Z \to ee$				
+ 0 jets	1.0	9.06	13 750	1.5
+ 1 jet	2.8	6.45	7 500	1.2
+ 2 jets	9.9	7.42	252 250	34.0
+ 3 jets	25.8	6.11	103 250	16.9
+ 4 jets	44.8	2.99	23 250	7.8
+ ≥5 jets	64.0	1.58	14 000	8.9
$Z \to \mu\mu$				
+ 0 jets	0.6	4.99	13 400	2.7
+ 1 jet	2.4	5.41	11 250	2.1
+ 2 jets	8.2	6.26	192 250	30.7
+ 3 jets	23.7	5.61	86 500	15.4
+ 4 jets	42.6	2.83	21 500	7.6
+ ≥5 jets	61.9	1.52	15 250	10.0
Sherpa EW samples:				
$Z \to \tau\tau \to \ell h$ + ≤ 3 jets				
No filter	–	1.69	198 500	117.0
VBF filter	12.6	0.214	49 500	230.0

Table 5.4: Overview of the W + jets Monte Carlo data samples used in this analysis and generated with ALPGEN. The second column shows the vector-boson fusion filter efficiency. The cross-sections include all branching ratios and filter efficiencies and are scaled to the next-to-leading order with a k-factor of 1.15. The last column shows the integrated luminosity the generated numbers of events correspond to.

	VBF Filter Efficiency [%]	Cross-section [pb]	Events	Luminosity [fb^{-1}]
$W \to e\nu$				
+ 2 jets	8.2	77.5	94 450	1.22
+ 3 jets	20.7	56.6	95 000	1.68
+ 4 jets	38.5	29.7	98 750	3.32
+ ≥5 jets	55.1	15.3	93 500	6.11
$W \to \mu\nu$				
+ 2 jets	7.5	71.0	97 950	1.38
+ 3 jets	19.4	53.1	77 750	1.46
+ 4 jets	36.7	28.0	58 500	2.09
+ ≥5 jets	54.7	15.3	96 950	6.35

Table 5.5: Overview of the $t\bar{t}$ samples used in this analysis. Since MC@NLO was used as generator for this process, the cross-section is at next-to-leading order accuracy. No filter has been applied in case of the fast detector simulation and for the detailed simulation a lepton filter has been applied (see text). The last column shows the integrated luminosity the generated numbers of events correspond to, where the event weights from MC@NLO have been accounted for.

Detector Simulation	Filter	Cross-section [pb]	Events [$\times 10^3$]	Luminosity [fb^{-1}]
Detailed	1 lepton	461	930	1.5
Fast	–	833	94 300	83.0

Monte Carlo Generators

The Monte Carlo generators used for this study are briefly described in the following:

- Pythia [53]
  A general-purpose generator for hadronic interactions in leading order with parton showering for initial and final state QCD radiation, which leads to relatively soft jets. No spin correlations are taken into account.

- Herwig [54]
  Another leading order general-purpose generator with a slightly different parton showering compared to Pythia. Herwig is particularly sophisticated in treating the decays of unstable particles with full spin correlation. The Jimmy program [55] is used to simulate the underlying event. Most of the signal samples have been simulated with Herwig.

- ALPGEN [56]
  A leading order generator with a different method to match the jets from the parton-showering model to the ones from matrix-element calculations, leading to harder jets. Special emphasis is given to final states with large jet multiplicities, based on exact leading order evaluation of partonic matrix elements. It is also interfaced to Herwig/Jimmy for the hadronization and underlying event simulation. ALPGEN was used to simulate the QCD Z + jets background processes.

- Sherpa [57]
  A relatively new multipurpose generator for hadronic events with the matrix element generator AMEGIC++ [58]. Since Sherpa also includes electroweak processes, it was used to simulate the electroweak $Z \to \tau\tau$ background.

- MC@NLO [59]
  The only next-to-leading order generator for QCD processes. Like for ALPGEN, MC@NLO was also interfaced to Herwig/Jimmy for the hadronization and underlying event simulation. Since higher order effects are particularly important for $t\bar{t}$ production, MC@NLO was used to generate this background.

All τ-lepton decays have not been simulated by the generators themselves, but by Tauola [60].

Event Filter To increase the background statistics in the phase space of interest event filters have been applied at the generator level to save CPU time by eliminating events that would not survive jet related cuts in the analysis.

Lepton filter for signal samples:

- With the number N_ℓ of electrons plus muons with $p_{\mathrm{T}} \geq 5\,\mathrm{GeV}$ and $|\eta| \leq 2.7$:
  $N_\ell \geq 2$ for the samples $H \to \tau\tau \to \ell h$
  $N_\ell \geq 1$ for the $H \to \tau\tau \to \ell h$ samples

VBF filter for Z events:

- With the number N_ℓ of electrons plus muons with $p_T \geq 10\,\text{GeV}$ and $|\eta| \leq 2.7$:
  $N_\ell \geq 2$ for the samples $Z \to ee, \mu\mu$ and $Z \to \tau\tau \to \ell\ell$
  $N_\ell \geq 1$ for the $Z \to \tau\tau \to \ell h$ samples
- At least two jets (cone algorithm with size $\Delta R \leq 0.4$) within $|\eta| \leq 5.0$ and
  $p_T \geq 20\,\text{GeV}$ for the highest p_T jet and
  $p_T \geq 15\,\text{GeV}$ for the second highest p_T jet
- A jet pair with an invariant mass of $m_{jj} \geq 300\,\text{GeV}$ and a pseudorapidity separation of $\Delta\eta_{jj} \geq 2$

Lepton filter for $t\bar{t}$:

- At least one electron, muon or τ lepton within $|\eta| \leq 5.0$ from a W decay

Pile-Up At the LHC design luminosity of $L = 1.0 \times 10^{34}\,\text{cm}^{-2}\text{s}^{-1}$, approximately 23 proton-proton collisions occur every 25 ns. Most of these collisions are elastic and inelastic scattering events (minimum bias) but the remnants of these processes are also recorded together with interesting interactions in the same bunch crossing (pile-up). The additional particles from pile-up events potentially cause difficulties in the reconstruction of the hard parton collision process. For example, the energy determined for a jet from the interesting event could be higher due to energy deposits from pile-up particles or the pile-up events could result in additional jets in the event. If the pile-up originates from the same bunch crossing as the main interaction, it is called in-time pile-up. In addition, there is also an out-of-time pile-up contribution, which means that the contribution is from an earlier bunch crossing. This is the case for several detector components which have a response- and read out time which is much longer than 25 ns, as for example the liquid Argon calorimeters or the drift tube detectors of the muon system. Other pile-up contributions are from the cavern background or from showers induced by particles from cosmic rays.
All these effects are not included in the simulated data used in this study as there have not been enough Monte Carlo datasets including pile-up available.

5.1.2 Detector Simulation

The detector response to the generated particles can be simulated by means of the detailed detector simulation or the fast detector simulation ATLFAST. In the following a brief overview of these two simulations is given – further details can be found in [50, 52].

Full Simulation

The full simulation of the ATLAS detector response is based on a detailed and realistic description of particle interactions with the detector material as provided by the GEANT4 simulation package. A detailed description of the detector geometry and material distribution is used for the simulation, as well as for the exact map of the magnetic field, the realistic response of the read out electronics

and the trigger behavior. The simulation output is then processed by the same reconstruction algorithms which will be used for the real collision data. The described simulation process results in the best possible predictions but is very time consuming. The full simulation of one event typically takes up to 20 minutes on a standard computer. For that reason, only a limited number of events can be fully simulated.

ATLFAST

In order to simulate a large number of events needed for some of the background processes (for example 100 millions of $t\bar{t}$ events), the much faster detector simulation software ATLFAST has been used. It is based on a parameterized description of the detector performance. The momentum vectors of the generated particles are smeared and the identification efficiency is parameterized according to studies with the full detector simulation performed for the most important physics processes. A homogeneous magnetic field is assumed in the inner detector volume and a very simplified calorimeter model is used to simulate the calorimeter response. No difference between electromagnetic and hadronic calorimeter and no longitudinal segmentation of the calorimeters is taken into account. The lateral segmentation is approximated as $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ for $|\eta| < 3.2$ and $\Delta\eta \times \Delta\phi = 0.2 \times 0.2$ for $3.2 < |\eta| < 5.0$. The energy of a generated shower particle is deposited in only one calorimeter cell. A cone cluster algorithm with a cone size of $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2} = 0.4$ is applied to all cells with a deposited energy above 1.5 GeV. The resulting clusters are associated with the originally generated electrons, photons, jets or τ jets.
The fast simulation of the detector takes only a few seconds per event. Clearly, the results obtained by the fast simulation depend on the accuracy of the parameterization of the detector response and are not as realistic as the ones obtained from full simulation. For this analysis, the most important difference between the detailed detector simulation and ATLFAST is the underestimated probability for τ-jet misidentification in the fast simulation. The uncertainties originating from ATLFAST for the background estimation are estimated in Section 6.6.3 and the influence on the discovery potential is discussed in Section 8.2.2.

5.2 Detector Performance

The following section summarizes the expected performance of the ATLAS detector in reconstructing objects used in this analysis (trigger, electrons, muons, jets, τ jets, b jets and the missing transverse energy) is summarized – further details can be found in [28]. One of the most important detector performance measures is the efficiency of reconstructing and identifying the mentioned objects. The efficiency is defined as the fraction of reconstructed and correctly identified objects of the total number of generated objects:

$$\text{Efficiency} = \frac{N_{\text{matched}}}{N_{\text{generated}}} \tag{5.1}$$

where $N_{\text{generated}}$ is the number of generated objects and N_{matched} is the number of objects that have been reconstructed and correctly identified and can be matched to a generated object within a small $\eta - \phi-$cone. The radius R of the cone depends on the type of object: $R = 0.1$ for electrons, muons and τ jets, $R = 0.4$ for jets.

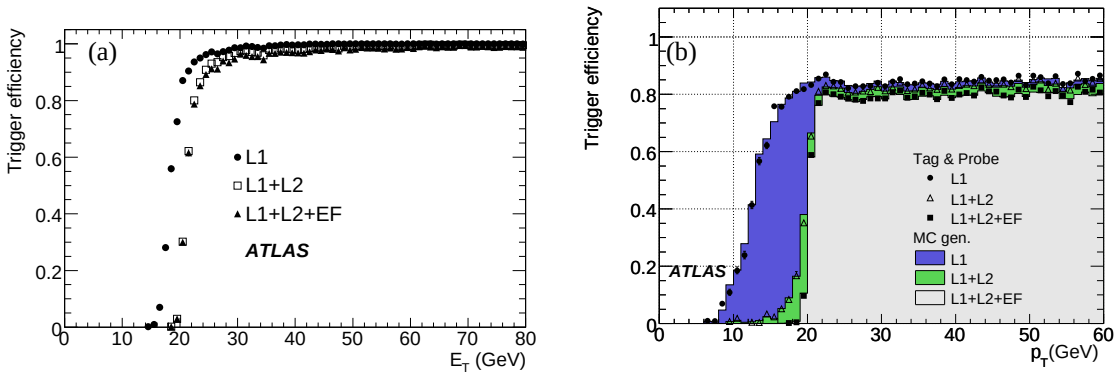


Figure 5.3: Trigger efficiencies (a) for electrons (e22i) and (b) for muons (mu20), depending on the transverse Energy E_T or the transverse momentum p_T. The three levels of the ATLAS trigger system, as described in Section 3.2.3, are denoted as L1, L2 and EF [28].

Another very important performance quantity is the rate of misidentified objects, for example the rate of τ jets that originate from wrongly identified electrons. The rate of misidentified objects is defined as:

$$\text{Misid. rate} = \frac{N_{\text{misidentified}}}{N_{\text{reconstructed}}} \tag{5.2}$$

where $N_{\text{misidentified}}$ is the number of reconstructed objects which cannot be matched to a corresponding generated object and $N_{\text{reconstructed}}$ denotes the total number of reconstructed and identified objects.
Obviously one aims for high efficiency at low misidentification rate.

5.2.1 Trigger Efficiency

As described in Section 3.2.3, the trigger in ATLAS consists of three levels, with the third level (event filter) performing the final decision. For this analysis two different triggers have been used:

- Isolated electrons with $p_T \geq 22\,\text{GeV}$ (denoted as e22i).
- Isolated muons with $p_T \geq 20\,\text{GeV}$ (denoted as mu20).

Figure 5.3 shows the efficiencies of these two trigger requirements depending on the transverse lepton energy or momentum. In the fast detector simulation no trigger simulation is available.
For both studied decay channels, only the mentioned electron and muon trigger has been applied. No τ-jet trigger requirement has been included in the semileptonic channel since reliable simulations of a τ-jet trigger have been included in the ATLAS detector simulation only very recently.

5.2.2 Electron Reconstruction

The reconstruction and identification of electrons [49, 28] in the detailed simulation starts with the information of the electromagnetic calorimeter. An algorithm searches for spatially grouped calorimeter cells with a significant energy deposition (cluster). To suppress photons, each cluster has to be matched to a track of the inner detector. Several additional requirements can be used to further suppress the contribution from photons and jets:

- Calorimeter information:
  Due to the fine granularity of the electromagnetic calorimeter, the lateral and longitudinal shower shape is used to separate electromagnetic from hadronic showers. Also the amount of energy deposited in the hadronic calorimeter, which should be very low in the case of electrons, provides a good suppression of jets.

- Inner detector information:
  A minimum number of hits in the pixel and SCT detectors, a large energy deposition in the transition radiation tracker (TRT) and a transverse impact parameter (minimum distance of the extrapolated track to the beam axis) of $|d_0| < 0.1$ cm is used to identify electrons.

- Cluster–track match criteria:
  The rejection of jets can be significantly improved by ensuring the consistency between the electromagnetic calorimeter and the inner detector information. A good agreement of the extrapolation of the inner detector track to the calorimeter and the cluster position are typical for electrons. The electron energy E measured in the electromagnetic calorimeter should match the momentum p measured in the inner detector. Therefore, requiring the ratio E/p to be close to unity provides additional separation power.

In this analysis, all electron candidates with $p_{\mathrm{T}} > 15\,\mathrm{GeV}$ that fulfill the cuts related to the calorimeter information and have a matched track in the inner detector with at least nine pixel and SCT hits are accepted. The ratio E/p and the TRT requirements are not imposed. This set of requirements is known as “medium” strict electron selection and provides a sufficient rejection of misidentified electrons at adequate efficiency. In addition, an isolation cut is applied, which requires that the energy deposited within a cone of radius $R = 0.2$ around the electron is less than 10 % of the electron energy.
In case of the ATLFAST detector simulation, a track is accepted as electron if

- a calorimeter cluster is found within $\Delta R = 0.15$ and $|\eta| < 2.5$ around a generated electron,

- the sum of all cluster energies within $\Delta R = 0.2$ is close to the generated electron energy and

- no additional cluster is found around the generated electron within $\Delta R = 0.4$.

If an electron fulfills these requirements, its associated energy is the generated energy, smeared by means of a resolution function obtained from test beam studies.
The graphs in Figure 5.4 show the electron efficiency and rate of misidentified electrons as a function of p_{T}, η and ϕ for the signal and the $t\bar{t}$ background from the detailed simulation and for $t\bar{t}$ also from the ATLFAST simulation. It can be seen that for the detailed simulation, the

electron efficiency is lower in the end-cap region ($|\eta| > 1.5$) of the detector than in the barrel region ($|\eta| < 1.5$). In the transition region between barrel and end-cap, the rate of misidentified electrons shows narrow peaks which can be explained by the worse calorimeter resolution in this detector region. Compared to the signal, one observes a lower efficiency for electrons in $t\bar{t}$ events. This is due to the fact that the electrons in $t\bar{t}$ events are less isolated than in the signal.
The less detailed detector description of the ATLFAST simulation leads to a flat efficiency distribution in η and ϕ and thus overestimates the electron efficiency, in particular in the forward regions of the detector. Furthermore, there are almost no misidentified electrons in the fast detector simulation.

5.2.3 Muon Reconstruction

Two algorithms are available to reconstruct and identify muons in ATLAS: STACO and MUID [61]. In this work the STACO algorithm has been used. Nevertheless, the performance of the two algorithms is very similar.
The information of all three main detector components, the inner detector, the calorimeter and the muon spectrometer, are exploited by the muon reconstruction. In the muon spectrometer, the track segments measured by the muon chambers are combined to a muon track, taking into account the material distribution traversed and the magnetic field along the muon trajectory. After that, the track is extrapolated to the interaction point, including a correction of the energy loss in the calorimeters and the inner detector. The energy loss is either estimated from a parameterized probability distribution of the expected energy loss or from the energy measured by the calorimeter, if it exceeds the most probable energy loss. Within the acceptance region of the inner detector ($|\eta| < 2.5$), this extrapolated track is then combined with tracks measured by the inner detector. This combination improves the momentum resolution for muons below 100 GeV and suppresses the misidentification of particles that escape the calorimeter (punch-through) and which are not muons.
The algorithm also reconstructs muons solely from the muon spectrometer information (stand-alone muons) or only from the inner detector measurements with hits in the inner layer of the muon spectrometer (low-p_{T} muons). However, only muons which have been successfully combined from muon spectrometer and inner detector measurements are used in this analysis. In addition, it is required that the muons have a minimum transverse momentum of $p_{\mathrm{T}} \geq 10$ GeV and that the muons are isolated, which means that the amount of energy deposited in the detector within a cone of radius 0.2 around the muon track is less than 10 % of the muon momentum. In this way, the rate of misidentified muons is kept at a low level and the momentum resolution is as good as ~2-3 % for muons up to p_{T}<100 GeV.
In case of the fast detector simulation, the transverse momenta of all generated muons are smeared by a p_{T}-, η- and ϕ-dependent resolution function. All muons with a p_{T}>5 GeV and $|\eta| < 2.5$ are accepted as reconstructed muons. They are further classified as isolated or non-isolated muons depending on the energy deposited in a cone of $\Delta R = 0.4$ around the muon. If a jet is reconstructed within this cone, the muon is added to the jet.
Figure 5.5 shows the efficiency and the ratio of misidentified muons as a function of p_{T}, η and ϕ for the signal and the $t\bar{t}$ background. For the $t\bar{t}$ background also the values for the ATLFAST simulation are shown. The detailed simulation shows regions with lower reconstruction efficiency

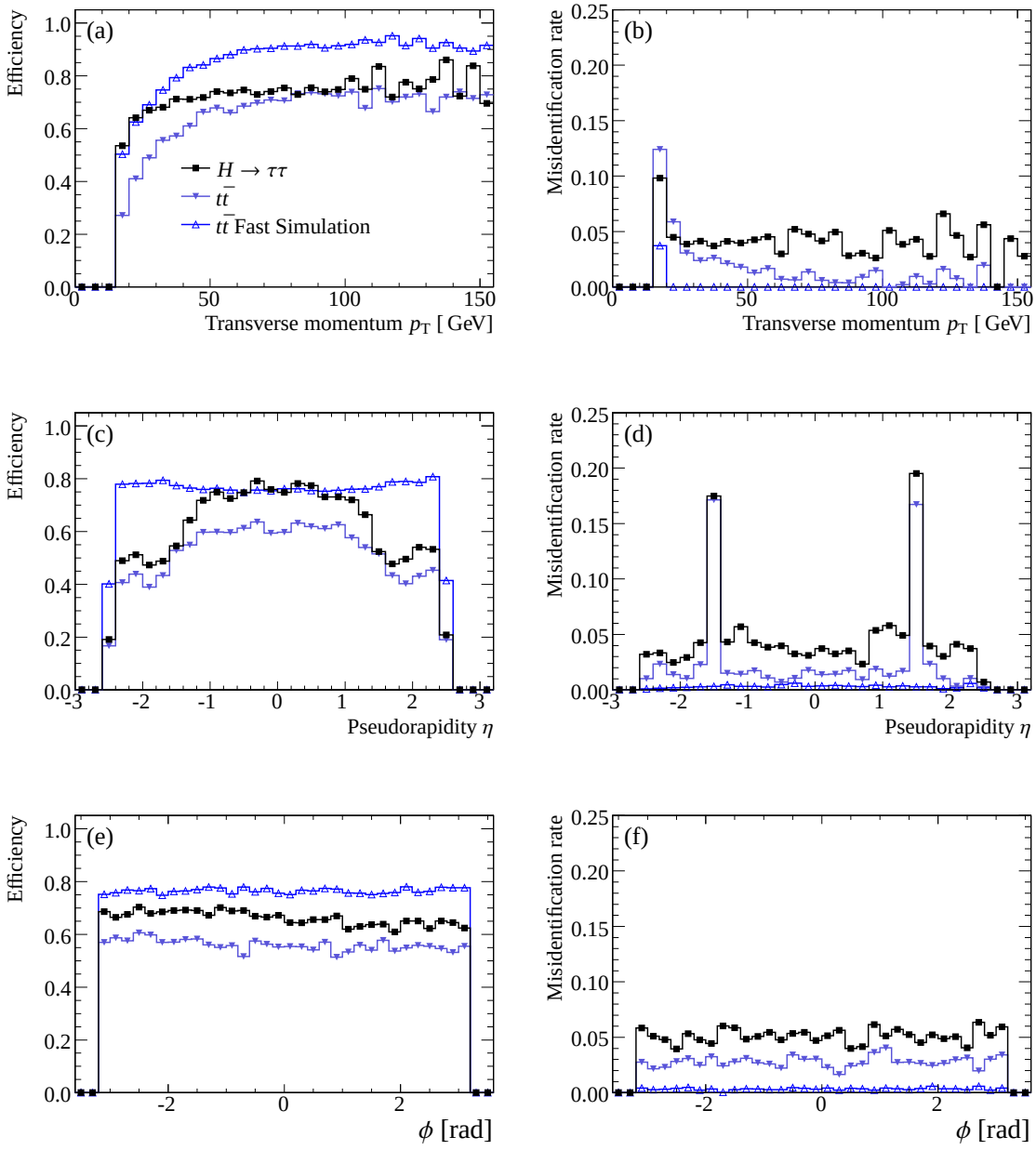


Figure 5.4: Electron reconstruction and identification efficiency as a function of (a) p_T, (c) η and (e) ϕ, for the signal and $t\bar{t}$ background in the detailed detector simulation and for the $t\bar{t}$ process from the fast simulation. The corresponding electron misidentification rates are shown in (b), (d) and (f).

in the η and ϕ distribution compared to the fast simulation. Around $\eta = 0$, the loss of efficiency can be explained by the gaps in the muon spectrometer due to service connections and around $|\eta| = 1.2$ the lower efficiency is caused by the lower coverage with muon chambers in the barrel–end-cap transition region. The small dips at $\phi = -2.1$ and $\phi = -1.0$ are due to the detector feet structure which leads to a higher amount of material the muons have to traverse and a lower coverage with muon chambers.
As not all details about the exact detector geometry are incorporated in the fast detector simulation, it shows a flat efficiency distribution in η and ϕ.

5.2.4 Jet Reconstruction Performance

Jet reconstruction [62] in ATLAS is performed by searching for accumulations of energy deposits in the calorimeter (cluster) which serve as starting points (seeds) for reconstructing jets. Two cluster finding algorithms are available: the "Tower" cluster algorithm which sums up the energy deposits in all calorimeter cells in $\eta - \phi$ bins (two dimensions) and a more sophisticated algorithm, the "Topological" (Topo) cluster algorithm, which reconstructs three-dimensional energy depositions in the calorimeter. The energy clusters found are taken as input to the jet finding algorithms. Again, two jet finding algorithms are available in ATLAS: the so-called "Cone" algorithm and the "k_T" algorithm. For this analysis, the Cone algorithm with a cone of $\Delta R = 0.4$ and Topo cluster seeds has shown the best performance. This is reflected in the higher efficiency for jets from the vector-boson fusion process in the forward region of the detector.
In the fast detector simulation, all clusters with a transverse energy above 10 GeV which could not be assigned to electrons or photons are classified as jets. The jet direction is defined as the cluster direction and the energies are smeared with the parameterized jet energy resolution.
Figure 5.6 shows the jet reconstruction efficiency and the ratio of misidentified jets as a function of p_T, η and ϕ for the signal and the $t\bar{t}$ background. For the $t\bar{t}$ background also the values for the ATLFAST simulation are shown. The jet reconstruction efficiency reaches almost 100 % for jets with a transverse momentum of $p_T \gtrsim 40$ GeV and the efficiency shows a rather flat distribution as a function of η and ϕ. The jet reconstruction efficiencies in the detailed and in the fast detector simulation are very similar to each other.
The jet misidentification rate is different for signal and $t\bar{t}$ background. For the signal the misidentification rate of jets is much higher for low transverse momenta within the pseudorapidity range of $|\eta| < 2.5$. This is due to jets that originate from hadronic τ-lepton decays, which have not been identified as τ jets (see Section 5.2.5).

5.2.5 τ Jet Reconstruction Performance

For the reconstruction and identification of τ jets, again two algorithms are available – a track based approach showing a higher performance for τ jets with low-p_T and the TauRec algorithm [63], which uses calorimeter clusters as τ-jet candidates and has been used in this analysis. The τ identification in TauRec is based on a likelihood analysis with the following discriminating variables against QCD jets:

- The number of tracks associated with the cluster, limited to be either one, two or three. Since in hadronic τ decays there are either one or three charged hadrons, one would ideally

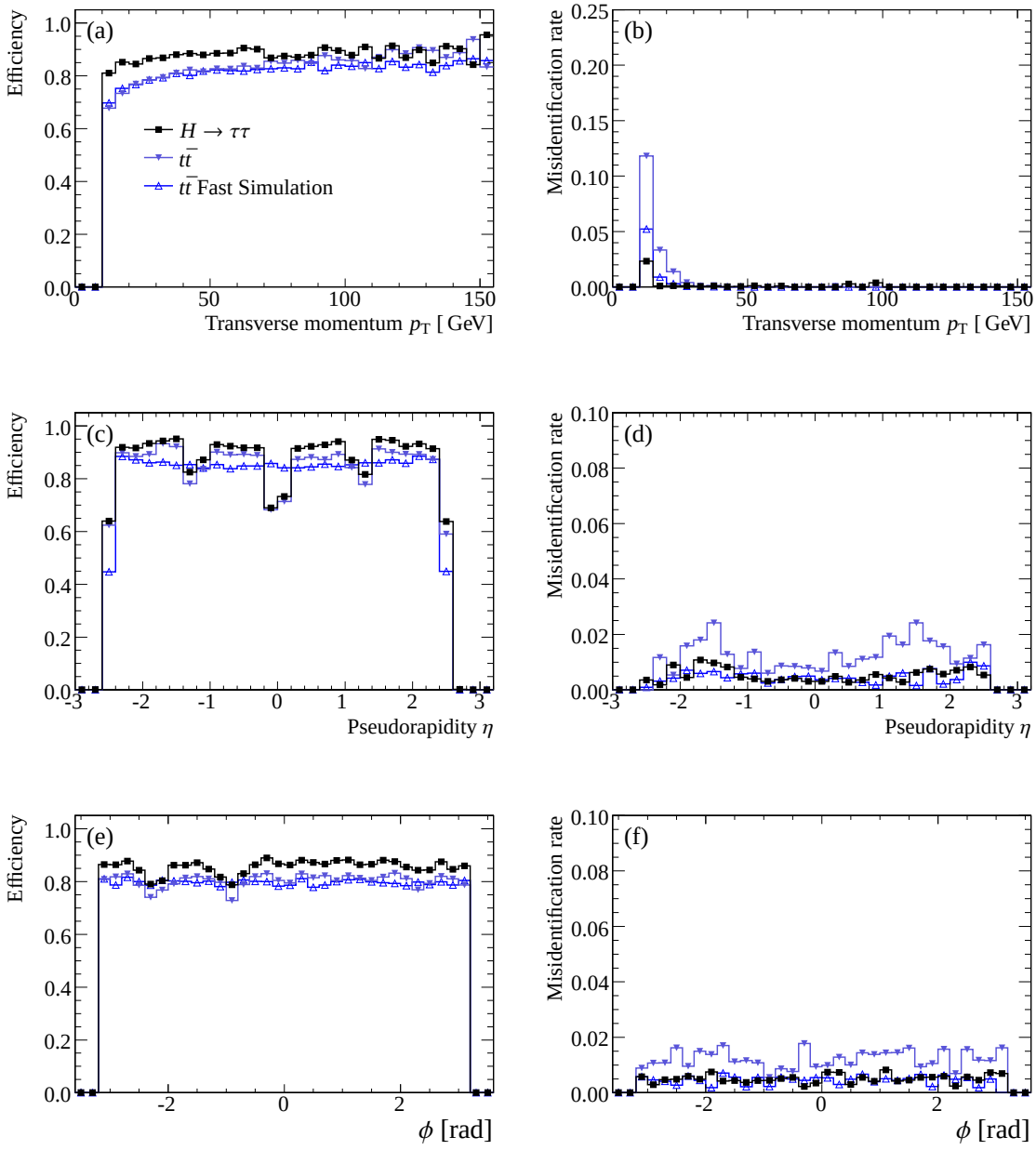


Figure 5.5: Muon reconstruction and identification efficiency as a function of (a) p_T, (c) η and (e) ϕ, for the signal and $t\bar{t}$ production in the detailed detector simulation and for the $t\bar{t}$ process in the fast simulation. The corresponding muon misidentification rates are shown in (b), (d) and (f).

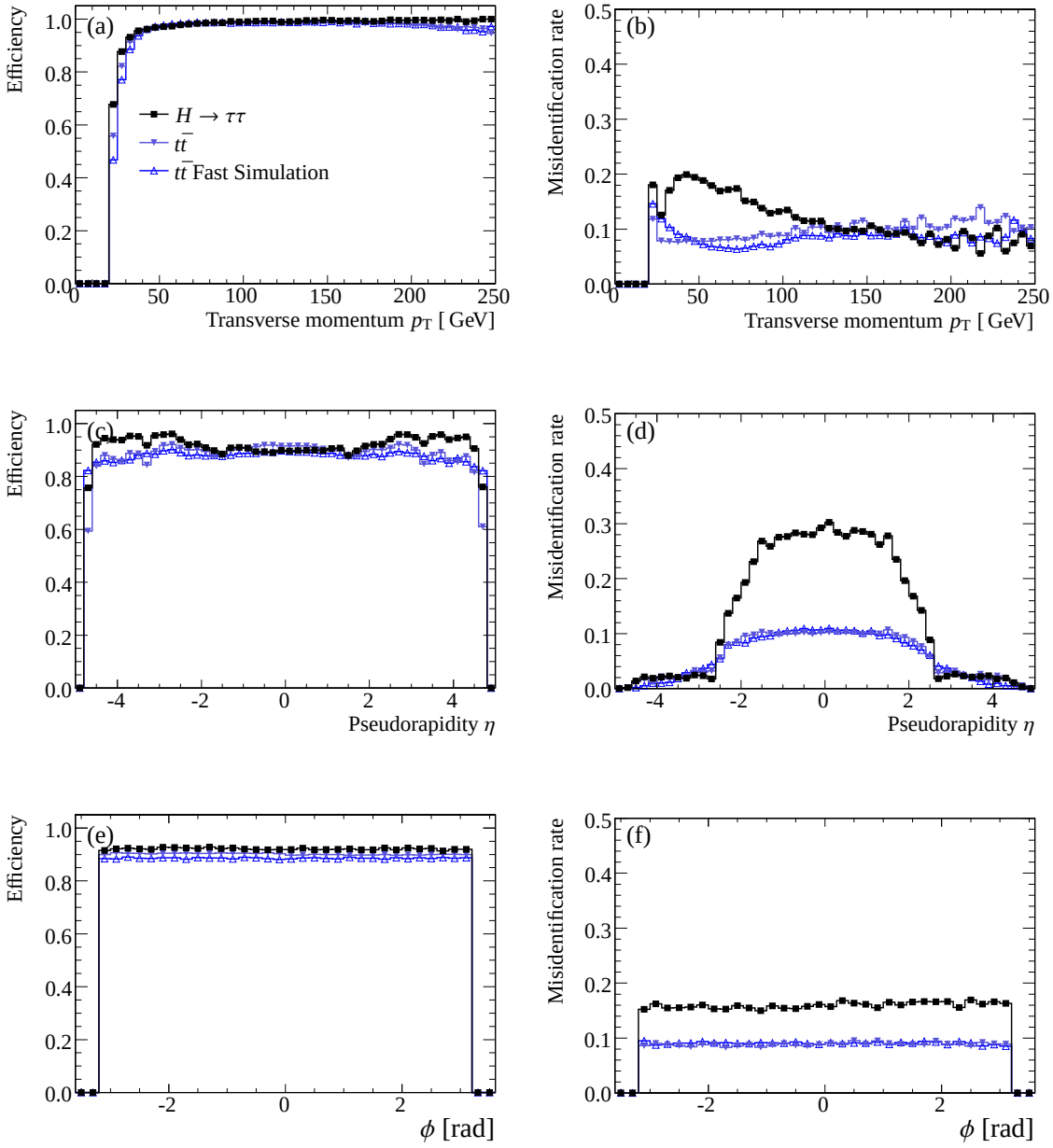


Figure 5.6: Jet reconstruction efficiency as a function of (a) p_T, (c) η and (e) ϕ, for the signal and $t\bar{t}$ production in the detailed detector simulation and for the $t\bar{t}$ process in the fast simulation. The corresponding jet misidentification rates are shown in (b), (d) and (f).

expect only one or three tracks. However, one of the three tracks might have been missed. (In this analysis, *exactly* one or three tracks have been required in addition.)

- The electric charge of the τ jet which is calculated by summing up the charge of all associated tracks should be ± 1.
- The impact parameter (distance to the primary interaction vertex) of the track with the highest p_{T} (leading track) associated with the τ candidate. Due to the relatively long life time of the τ lepton of 2.9×10^{-13} seconds [22], its decay vertex is displaced from the primary vertex.
- The radius of the cluster in the electromagnetic calorimeter. Jets from hadronic τ decays tend to be more collimated than QCD jets.
- The fraction of transverse energy in a concentric ring of $0.1 < \Delta R < 0.2$ around the τ candidate cluster should be small, also because τ jets are well-collimated.
- The number of hits with $E_{\mathrm{T}} > 200$ MeV in the η-strip layer of the electromagnetic calorimeter. Hadronically decaying τ leptons typically have a smaller number of hits compared to QCD jets.
- The width of the shower profile in the η-strip layer (see page 22) of the electromagnetic calorimeter which tends to be relatively small for τ jets.
- The ratio $E_{\mathrm{T}}^{\mathrm{had}}/p_{\mathrm{T}}^{1}$ of the transverse energy $E_{\mathrm{T}}^{\mathrm{had}}$ measured by the hadron calorimeter, and the transverse momentum p_{T}^{1} of the leading track. In contrast to QCD jets, τ jets tend to have a high fraction of their energy associated to the leading track, which leads to a relatively small $E_{\mathrm{T}}^{\mathrm{had}}/p_{\mathrm{T}}^{1}$ ratio.

All τ-jet candidates with a transverse momentum of $p_{\mathrm{T}} > 30\,\mathrm{GeV}$ and a TauRec log-likelihood ratio of larger than 4.0 are accepted as τ jets. For a further suppression of electrons misidentified as τ jets, additional cuts have been applied to the τ-jet candidates:

- A minimum of 0.2 % of the energy of the τ-jet candidate has to be deposited in the first layer of the hadron calorimeter. Electrons usually deposit all their energy in the electromagnetic calorimeter.
- Within the range $|\eta| \leq 1.7$, the ratio of high-threshold (HT) to low-threshold (LT) hits in the TRT has to be smaller than 20 % if there are at least ten LT hits. Due to their higher mass, pions from a hadronic τ decay emit less transition radiation than electrons which results in smaller signals in the TRT.

In case of the fast detector simulation, a jet is considered to be a τ jet if it matches the generated hadronically decaying τ lepton within a cone of $\Delta R < 0.3$ and if the jet energy is comparable with the visible τ energy. Subsequently, a p_{T}- and η-dependent parameterization of the τ-jet identification efficiency is applied to each τ-jet candidate.
The performance of the τ-jet reconstruction and identification is shown in Figure 5.7. It shows the efficiency and misidentification rate as a function of p_{T}, η and ϕ, for the detailed detector

simulation of signal and $t\bar{t}$ background. For the $t\bar{t}$ background also the values for the ATLFAST simulation are shown.
The efficiencies of the three distributions shown are very similar to each other. They are rather flat in p_T, η and ϕ and the average efficiency is around 40 % for the τ jets from $t\bar{t}$ events and around 50 % for the τ jets from the Higgs boson decay. As for the electrons and muons, the slightly lower τ-jet efficiency in $t\bar{t}$ events can be explained by the τ jets being less isolated. Compared to the signal, the rate of misidentified τ jets is much higher for the $t\bar{t}$ process in the detailed simulation. This can be explained by the higher jet multiplicity of the $t\bar{t}$ events in the central detector region. With ~10 %, the rate of misidentified τ jets in the fast detector simulation ATLFAST is much less than in the detailed detector simulation where the rate of misidentified τ jets is as high as ~60 %. This difference causes the biggest uncertainty introduced to the analysis of the Higgs boson discovery potential if ATLFAST is used to estimate the $t\bar{t}$ background. In Section 6.6.3, the impact of using ATLFAST data samples in this analysis is discussed.

5.2.6 *b* Jet Identification

Jets originating from b quarks show characteristics that can be used to separate them from jets originating from lighter quarks (b-tagging):

- Due to the hard fragmentation (the b hadron has a large fraction of the b quark momentum) and the high masses of b hadrons (> 5 GeV), the decay products of the b hadron typically have large transverse momenta with respect to the jet axis and thus, their opening angle is also relatively large.

- The lifetime of b hadrons of around 1.5 ps is rather long leading to a displacement of the b decay vertex from the primary interaction point of a few millimeters depending on the jet energy. Thus, the impact parameters of the tracks associated with the b hadron decay products are rather large.

A likelihood ratio approach [64, 28, 49] is used to estimate the probability that a jet originated from a b quark. In this analysis, a cut of > 1.0 on the output value of this b-jet identification algorithm (b-tag weight) is applied to separate b jets from light quark jets.
In the fast detector simulation, the identification of b jets is done in a similar way as for the τ jets. A jet is labelled as b jet if it can be matched within $\Delta R < 0.2$ to a generated b quark with $p_T > 5$ GeV. In addition, the labelling is stochastically changed for every event using a parameterization of the b-tag efficiency and light quark jet misidentification probability.
The expected performance of identifying b jets with this method is shown in Figure 5.8. The graphs show the b-jet identification efficiency and the rate of misidentified b jets as a function of p_T, η and ϕ for $t\bar{t}$ events in the full and the fast detector simulation. It can be seen that the efficiency as well as the rate of misidentified b jets are higher in case of the full detector simulation.

5.2.7 Missing Energy Reconstruction

Since neutrinos do not interact in the detector, they deposit no energy. However, it is possible to calculate the sum of all neutrino energies in an event in the plane transverse to the beam axis by

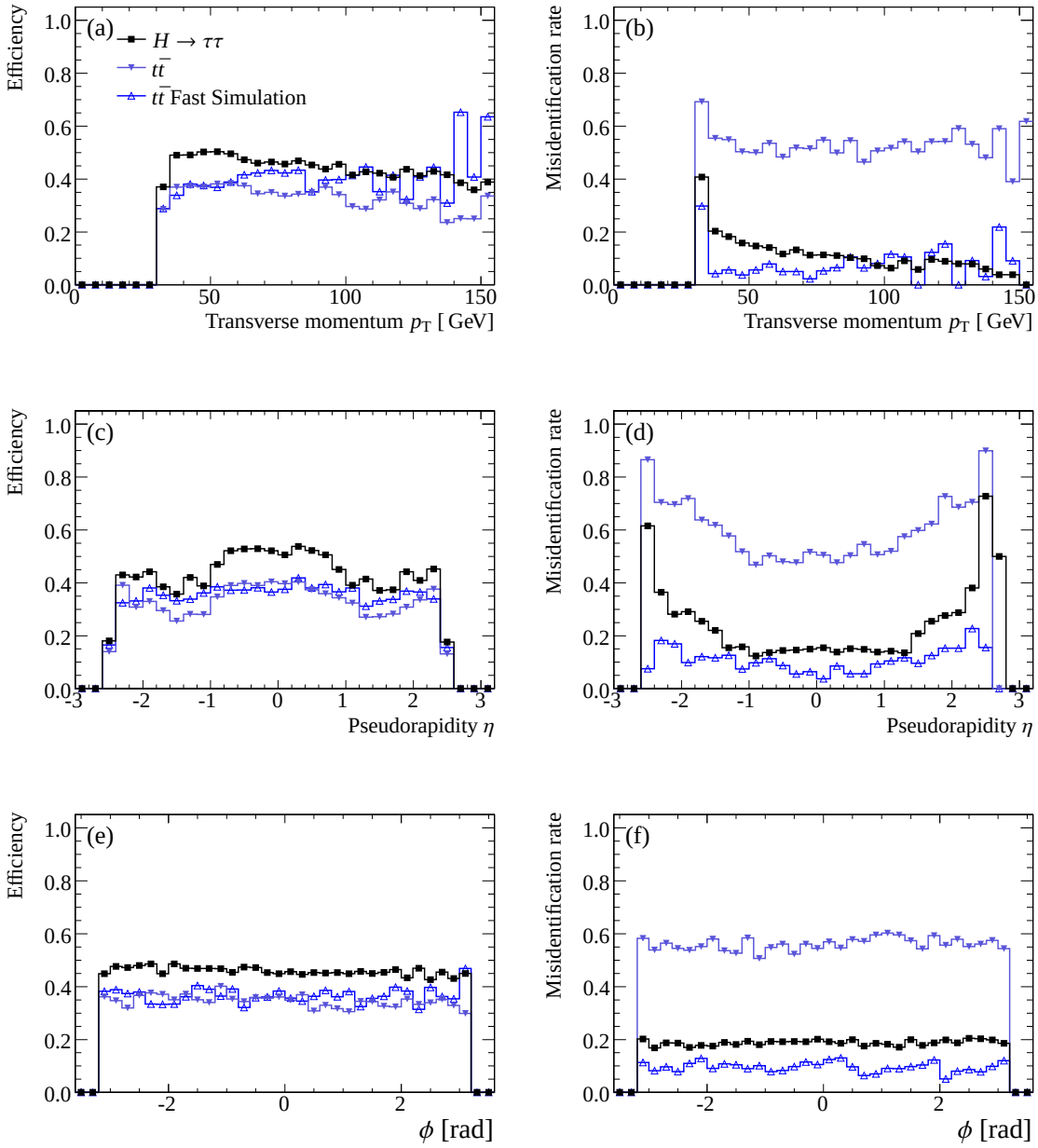


Figure 5.7: Reconstruction and identification efficiency for τ jets as a function of (a) p_T, (c) η and (e) ϕ, for the signal and $t\bar{t}$ production in the detailed detector simulation and for the $t\bar{t}$ process in the fast simulation. The corresponding τ jet misidentification rates are shown in (b), (d) and (f).

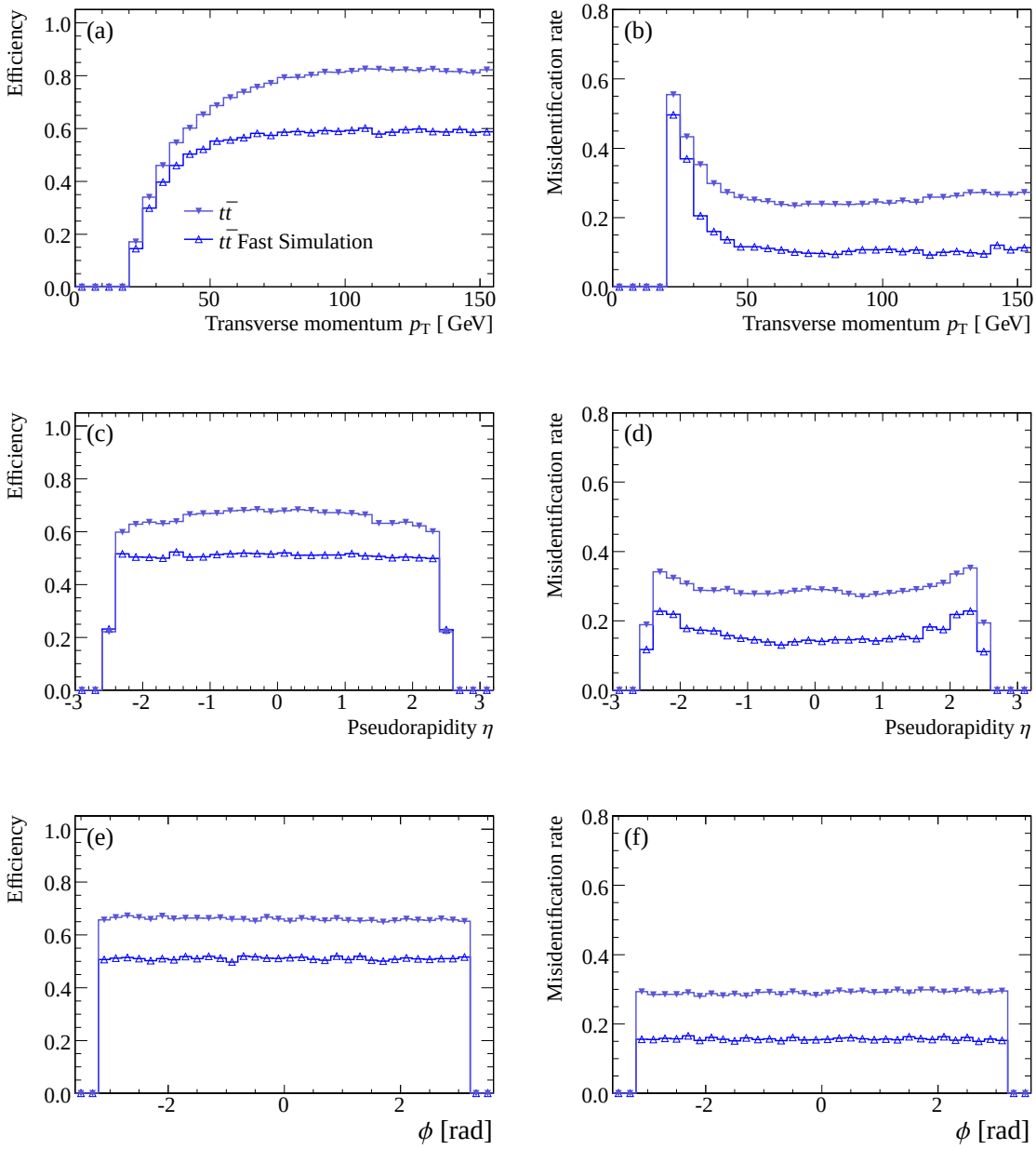


Figure 5.8: Performance of the *b*-jet identification in $t\bar{t}$ events for the detailed and the fast detector simulation. The *b*-tag efficiency is shown as a function of p_T, η and ϕ in (a), (c) and (e) respectively. The corresponding *b* jet misidentification rates are shown in (b), (d) and (f).

using momentum conservation. The colliding protons do not have transverse momentum components and therefore the sum of the transverse momenta of all final state particles has to vanish as well. Thus, the sum of the transverse momenta of all visible final state particles equals the negative sum of the transverse momenta of all neutrinos.
In ATLAS, the calculation of the visible transverse energy is performed in the following way:

- Summation of all calorimeter cell energies, using a calibration independent of the incident particle type.
- Addition of all muons with an associated high quality track in the inner detector in order to reduce the influence of misidentified muons. The muon momentum contributing to the missing energy calculation is measured in the muon spectrometer only, providing the muon energy after traversing the calorimeters. This ensures that the energy loss of the muons in the calorimeter is not double-counted.
- Addition of the estimated energy loss in the inactive material, particularly in the liquid Argon calorimeter cryostat walls.
- Refinement of the energy calibration of the contributing calorimeter cells. Since the calorimeter response depends on the particle type, the energy calibration is also different for different objects like, for example, electrons and jets. Thus, the calibration of all calorimeter cells which can be associated to a close-by reconstructed object, is replaced by the calibration specific for the type of the identified particle.

In the fast detector simulation, all reconstructed objects (photons, electrons, muons, τ jets, QCD jets and unmatched calorimeter clusters) are used to calculate the missing transverse energy $E_{\mathrm{T}}^{\mathrm{miss}}$. Figure 5.9 shows the expected resolution of the missing transverse energy measurement from the detailed detector simulation as a function of the absolute value of the missing transverse energy. For a $E_{\mathrm{T}}^{\mathrm{miss}}$ of $\sim 50\,\mathrm{GeV}$, which is a typical value in the case of $H \rightarrow \tau\tau$ events, the resolution is in the order of 10 %.

5.3 Reconstruction of the Higgs Mass

The decay of each τ lepton in the signal and background processes is always characterized by the emission of one or two neutrinos leading to missing transverse energy in the detector.
By means of the so-called collinear approximation, this overall missing transverse energy can be divided into the two neutrino contributions, one from each τ-lepton decay. Since the invariant mass of a τ lepton (1.777 GeV [22]) is much smaller than the Higgs mass, the τ leptons from the Higgs decay are strongly boosted. In the laboratory frame, the τ-lepton decay products are emitted approximately parallel to the original τ-lepton direction. Thus, the direction of the emitted neutrinos can be assumed to be the same as the direction of the visible τ-decay products (electrons, muons or τ jets). Figure 5.10 demonstrates the small angular distance, as simulated by the Tauola package, between the decaying τ lepton and its visible decay product, justifying the collinear approximation. The splitting up of the total missing transverse energy into the two neutrino contributions from the two τ decays is illustrated in Figure 5.11.

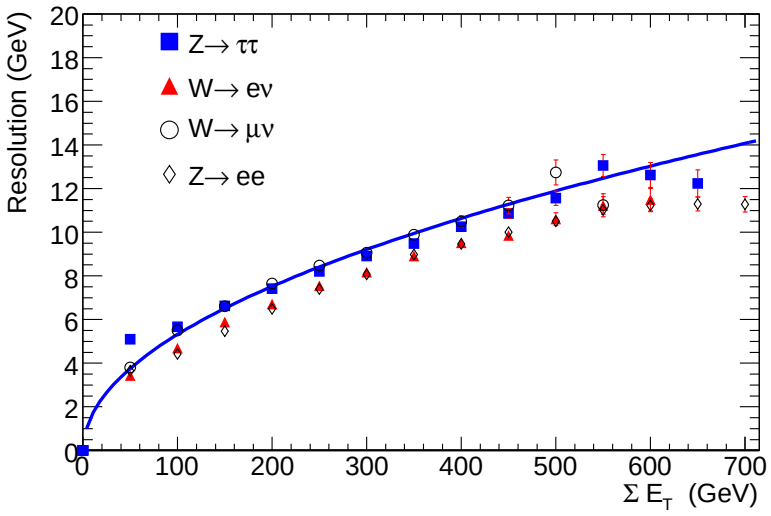


Figure 5.9: Expected resolution of the missing transverse energy $E_{\mathrm{T}}^{\mathrm{miss}}$ measurement from the detailed detector simulation as a function of the absolute value of the missing transverse energy equal to the visible energy $\sum E_{\mathrm{T}}$. [28]

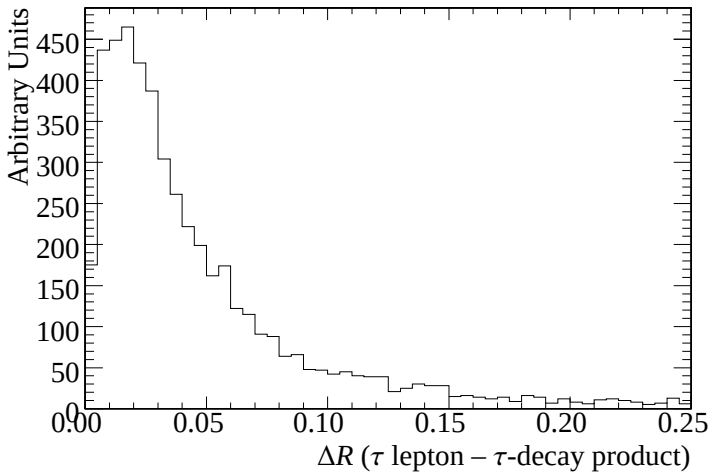


Figure 5.10: Angular distance between the τ lepton and its visible decay product (here: electron or muon). The distance ΔR is defined as $\Delta R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2}$.

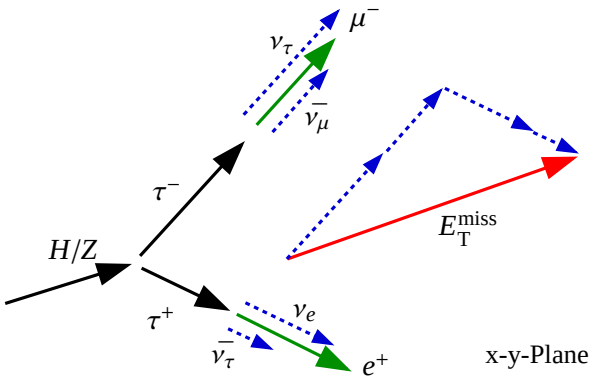


Figure 5.11: Sketch of the collinear approximation and the reconstruction of the invariant τ-pair mass. Due to the strong boost of the τ leptons, it can be assumed that the τ-decay products are emitted in the flight direction of the decaying τ lepton. This direction is known from the visible τ-decay products (here: μ^- and e^+). The measured missing transverse energy ($E_{\mathrm{T}}^{\mathrm{miss}}$) can then be separated into the two neutrino contributions ($\nu_\tau + \bar{\nu}_\mu$ and $\nu_e + \bar{\nu}_\tau$) from the two τ decays. This allows for the complete reconstruction of the τ-lepton four-momenta and thus of the Higgs boson mass. The x-y plane is perpendicular to the beam direction z.

The momentum fractions χ_1 and χ_2 carried by the visible τ-lepton decay products with respect to the initial τ-lepton momenta are defined as:

$$\chi_1 = \frac{p^1}{p^{\tau_1}} = \frac{p_{\mathrm{x}}^1}{p_{\mathrm{x}}^{\tau_1}} = \frac{p_{\mathrm{y}}^1}{p_{\mathrm{y}}^{\tau_1}} = \frac{p_{\mathrm{z}}^1}{p_{\mathrm{z}}^{\tau_1}}, \quad \chi_2 = \frac{p^2}{p^{\tau_2}} = \frac{p_{\mathrm{x}}^2}{p_{\mathrm{x}}^{\tau_2}} = \frac{p_{\mathrm{y}}^2}{p_{\mathrm{y}}^{\tau_2}} = \frac{p_{\mathrm{z}}^2}{p_{\mathrm{z}}^{\tau_2}}. \tag{5.3}$$

with $p^{\tau_{1(2)}}$ representing the momenta of the τ leptons, $p^{1(2)}$ the momenta of the visible τ-lepton decay products (electron, muon or τ jet), and x (y, z) denoting the x (y, z) momentum component. Due to the collinearity the visible momentum fractions are the same in all directions. Therefore, it is sufficient to calculate $\chi_{1,2}$ in the x-y-plane, perpendicular to the beam direction z.
The τ-lepton momenta can be written as:

$$\begin{aligned} p_{\mathrm{x}}^{\tau_1} &= p_{\mathrm{x}}^1 + p_{\mathrm{x}}^{\nu_1}, \quad p_{\mathrm{x}}^{\tau_2} = p_{\mathrm{x}}^2 + p_{\mathrm{x}}^{\nu_2}, \\ p_{\mathrm{y}}^{\tau_1} &= p_{\mathrm{y}}^1 + p_{\mathrm{y}}^{\nu_1}, \quad p_{\mathrm{y}}^{\tau_2} = p_{\mathrm{y}}^2 + p_{\mathrm{y}}^{\nu_2}. \end{aligned} \tag{5.4}$$

where $p^{\nu_{1(2)}}$ are the momenta of the neutrinos originating from the $\tau_{1(2)}$ lepton. In case of the leptonic τ-decay mode, p^ν is given by the sum of the τ neutrino and an electron or muon neutrino, while in case of the hadronic τ decay p^ν corresponds to the momentum of the τ neutrino alone.
Since the missing transverse energy originates from the neutrinos, one can write:

$$\begin{aligned} E_{\mathrm{x}}^{\mathrm{miss}} &= p_{\mathrm{x}}^{\nu_1} + p_{\mathrm{x}}^{\nu_2}, \\ E_{\mathrm{y}}^{\mathrm{miss}} &= p_{\mathrm{y}}^{\nu_1} + p_{\mathrm{y}}^{\nu_2}. \end{aligned} \tag{5.5}$$

Using equations 5.3 and 5.4 this can be written as:

$$\begin{aligned} E_{\mathrm{x}}^{\mathrm{miss}} &= p_{\mathrm{x}}^{1}\frac{1-\chi_1}{\chi_1} + p_{\mathrm{x}}^{2}\frac{1-\chi_2}{\chi_2} \\ E_{\mathrm{y}}^{\mathrm{miss}} &= p_{\mathrm{y}}^{1}\frac{1-\chi_1}{\chi_1} + p_{\mathrm{y}}^{2}\frac{1-\chi_2}{\chi_2} \end{aligned} \tag{5.6}$$

Solving this system of equations one obtains:

$$\begin{aligned} \chi_1 &= \frac{p_{\mathrm{x}}^{1}p_{\mathrm{y}}^{2} - p_{\mathrm{y}}^{1}p_{\mathrm{x}}^{2}}{p_{\mathrm{y}}^{2}(p_{\mathrm{x}}^{1} + E_{\mathrm{x}}^{\mathrm{miss}}) - p_{\mathrm{x}}^{2}(p_{\mathrm{y}}^{1} + E_{\mathrm{y}}^{\mathrm{miss}})} \\ \chi_2 &= \frac{p_{\mathrm{x}}^{1}p_{\mathrm{y}}^{2} - p_{\mathrm{y}}^{1}p_{\mathrm{x}}^{2}}{p_{\mathrm{x}}^{1}(p_{\mathrm{y}}^{2} + E_{\mathrm{y}}^{\mathrm{miss}}) - p_{\mathrm{y}}^{1}(p_{\mathrm{x}}^{2} + E_{\mathrm{x}}^{\mathrm{miss}})} \end{aligned} \tag{5.7}$$

Thus, the momentum fractions χ_1 and χ_2 of the visible τ-decay products can be calculated and the momentum of the decaying τ lepton can be evaluated according to equation 5.3. Knowing the momenta of the τ leptons it is now possible to calculate the Higgs mass m_H by determining the invariant mass $m_{\tau\tau}$ of the τ-lepton pair:

$$m_H = m_{\tau\tau} = \sqrt{(\mathbf{P}_{\tau_1} + \mathbf{P}_{\tau_2})^2} \tag{5.8}$$

The mass of the τ leptons can be neglected because of their high momenta, leading to:

$$m_H = \sqrt{2p_{\tau_1}p_{\tau_2} \cdot (1 - cos(\Delta\phi_{\tau\tau})} \tag{5.9}$$

where $\Delta\phi_{\tau\tau}$ is the angle between the two τ leptons (or, due to the collinear approximation, between the two visible τ-decay products).
Since the Z decay into τ leptons is very similar to the signal and as the Z boson ($m_Z = 91.2$ GeV [22]) is also much heavier than a τ lepton, the collinear approximation can also be used to reconstruct the Z mass.

Limitation of the Collinear Approximation In case of the leptons being emitted back-to-back, the previously mentioned system of equations (5.7) does not have a unique solution and the separation of the total transverse missing energy into the two neutrino contributions is ambiguous.
For leptons that are back-to-back in the r-ϕ plane one can write:

$$\begin{aligned} p_{\mathrm{x}}^{2} &= -a \cdot p_{\mathrm{x}}^{1} \\ p_{\mathrm{y}}^{2} &= -a \cdot p_{\mathrm{y}}^{1}\,, \end{aligned} \tag{5.10}$$

where a is an arbitrary constant greater than zero.
Inserting this into equations 5.7, the numerators of the equations are zero and the calculation of the τ-lepton momentum with equation 5.3 results in a singularity. Therefore, the application of the collinear approximation is limited to the cases where the two τ-decay products are not back-to-back. This is usually ensured by requiring the angular separation of the τ-decay products, $\Delta\phi_{\tau\tau}$ to be smaller than 180°.

Performance of the Collinear Approximation The invariant Higgs mass resulting from the collinear approximation is shown in Figure 5.12a for a simulated Higgs mass of 120 GeV where the intrinsic width of the generated Higgs resonance can be neglected. The distribution is obtained using the generated four-momenta of the visible τ-decay products without additional, detector related effects. The mass peak can be fitted with a Gauss function with a mean of $\mu = (119.5 \pm 0.1)$ GeV and a standard deviation of $\sigma = (3.0 \pm 0.1)$ GeV.

Figure 5.12b shows the impact of the detector performance on the reconstructed Higgs boson mass. The detector response leads to a shift of the mean mass value to $\mu = (118.0 \pm 0.2)$ GeV and to a broadening of the width to $\sigma = (10.3 \pm 0.2)$ GeV. It can be seen that the main source of the degraded mass resolution is the resolution of the missing energy measurement. The finite resolution of the missing transverse energy measurement broadens the mass distribution even more than the collinear approximation itself assuming an exactly known $E_{\mathrm{T}}^{\mathrm{miss}}$. In contrast to this, the lepton momentum resolution has a negligible effect on the reconstructed invariant mass. This can be expected since the resolution of the momentum measurement of electrons and muons is in the order of 2-3 % while it is only ∼10 % for the missing transverse energy measurement. A well understood and precise reconstruction of the missing transverse energy is therefore crucial for the Higgs boson discovery in this channel.
The expected ATLAS detector performance allows for a good separation of the Z resonance and the Higgs signal peak for Higgs masses above $\gtrsim$110 GeV.

5.4 Event Selection Criteria

In order to suppress background processes in the presented Higgs boson search channel, a variety of selection criteria are available which are introduced in the following. The criteria can be divided into three categories: one group related to the τ-decay products, another group concerning the forward jets produced in the vector-boson fusion and a last group related to both parts of an event, the τ-decay products and the forward jets.
For several reasons, the selection cuts on some discriminating variables have been fixed to a certain value or constrained to a subrange of the distribution (preselection). For these variables, the cut value is shown in the following distributions and the motivation for the cut is explained in the text. If no cut is indicated in the figure, no obvious reason for a certain cut value exists and the cut on the variable has been optimized as described in Chapter 6 or has been used as an input variable to multivariate analysis methods (see Chapter 7).

5.4.1 τ-Decay Products Criteria

Trigger

As a first step in each analysis, it is inevitable to require an event to be accepted by the trigger algorithm. In the leptonic and semileptonic $H \to \tau\tau$ channel, there is at least one electron or muon among the decay products in the final state. Therefore, the signal events can be triggered by high-

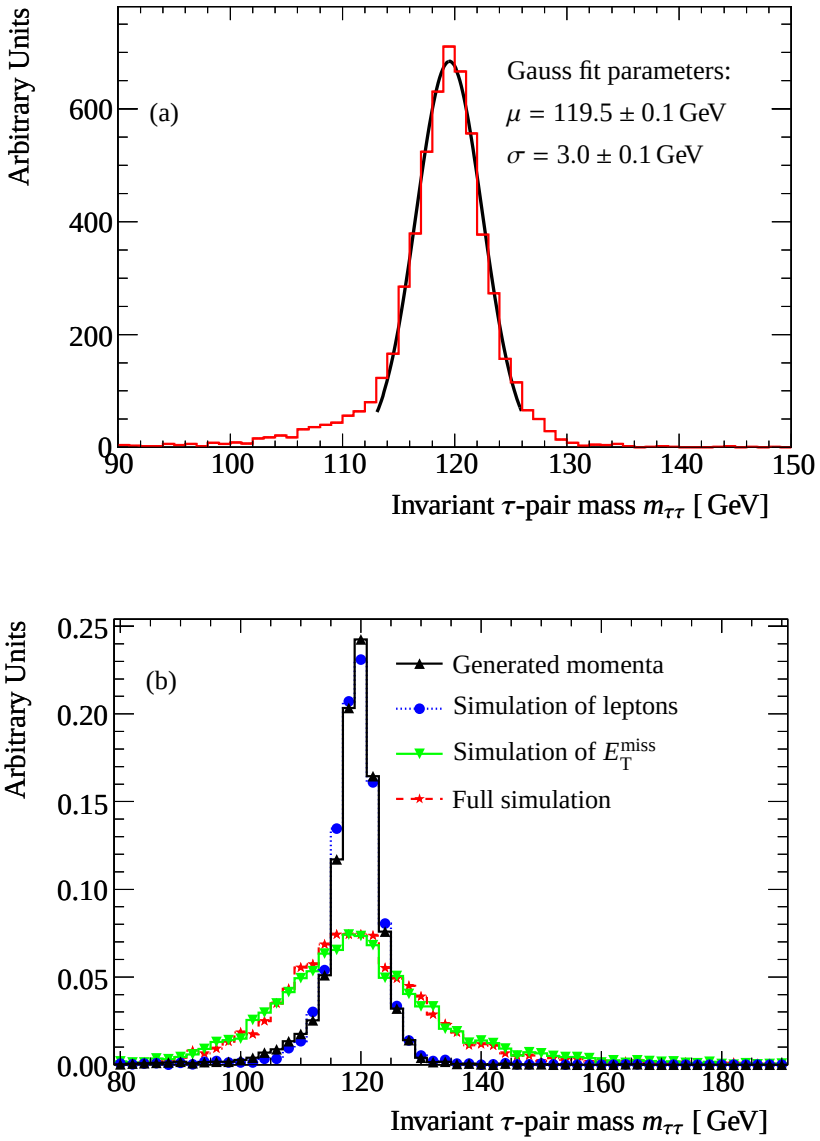


Figure 5.12: Invariant mass distribution of a 120 GeV Higgs boson in the leptonic $H \rightarrow \tau\tau$ decay mode. (a) shows the mass calculated by means of the collinear approximation. The exact, generated momenta of the particles are used for the collinear approximation without detector-related effects. Also shown is a Gauss function fitted to the distribution. In (b) the influence of the expected detector performance is shown. The true momenta of the visible leptons and the true E_T^{miss} are replaced by the reconstructed ones. The missing energy resolution dominates the Higgs mass resolution.

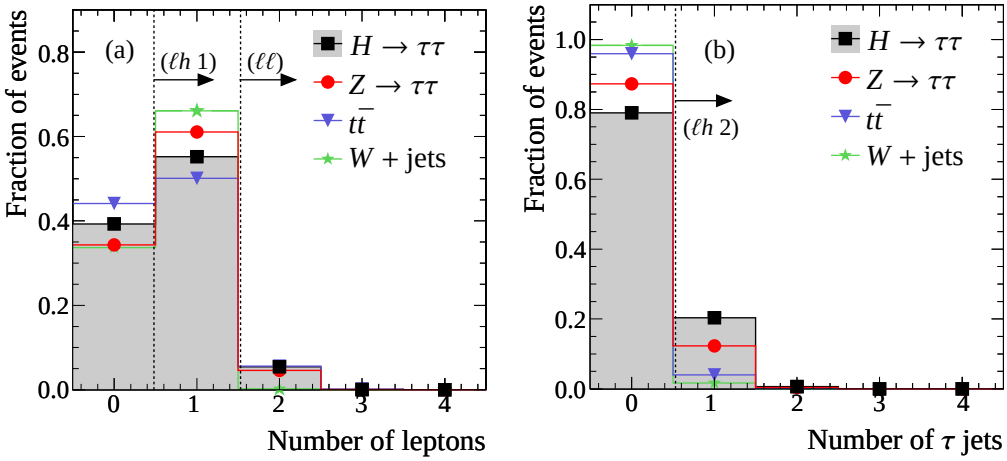


Figure 5.13: (a) Number of reconstructed leptons (electrons and muons) and (b) number of reconstructed τ jets after requiring exactly one reconstructed electron or muon, shown for the signal and the dominant background processes. (The selection cuts for the leptonic and semileptonic decay channels are indicated by the dashed lines.)

p_T lepton triggers: an isolated electron with a transverse momentum of $p_T^e \geq 22\,\text{GeV}$ and/or an isolated muon with $p_T^\mu \geq 20\,\text{GeV}$. The efficiency of these triggers is shown in Section 5.2.1.
To reduce the contribution from wrong triggers, the reconstruction and identification of a lepton of the same type as the triggered one with the corresponding minimum p_T is required.

Reconstructed τ-Decay Products

Obviously, two τ-decay products are required to be reconstructed in each candidate event. For the leptonic channel exactly two leptons are required ($ee, e\mu, \mu\mu$). There is no constraint on the number of reconstructed τ jets in the leptonic channel. For the semileptonic channel, exactly one electron or muon and one τ jet is required ($e\tau, \mu\tau$). Figure 5.13a shows the number of reconstructed leptons for different processes. Due to the τ-decay branching ratios (see Table 5.1) and also the limited detector acceptance $|\eta| < 2.5$ and reconstruction efficiencies for electrons and muons (see Section 5.2), the number of leptons shows a maximum at one. Only for a small fraction of events, two leptons are reconstructed, which is the requirement for an event in the leptonic decay channel (indicated by the dashed line in the plot).
The distribution of the number of τ jets after requiring exactly one electron or muon is shown in Figure 5.13b. To be a candidate for the semileptonic decay channel, at least one τ jet has to be present in such events (which is indicated by the dashed line in the plot).
In addition, it is required that the reconstructed τ-jet products have opposite charge. As shown in Figure 5.14, this requirement suppresses a large fraction of the W + jets events. In these events, a jet is misidentified as one τ-decay product and its charge is therefore not strictly correlated with the charge of the second identified τ-decay product. Around 30 – 40 % of the W + jets events can

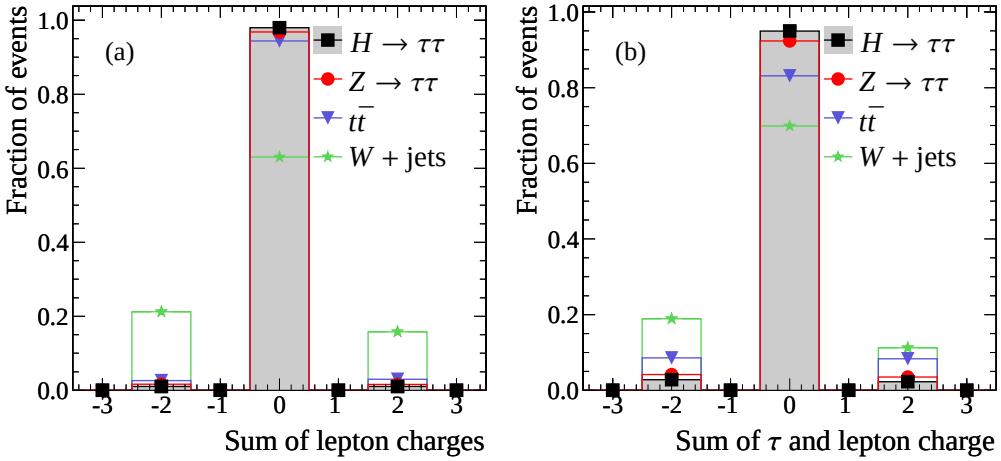


Figure 5.14: Sum of the charges of the two potential τ-lepton decay products. (a) for the leptonic decay mode and (b) for the semileptonic decay channel.

be suppressed by this requirement, while the other background processes and the signal are less sensitive to this selection criterion.
Since one of the τ-decay products in W + jets events is a misidentified object, one would expect a fraction of 50 % of the events not fulfilling the opposite charge cut. However, the charges of the two τ-decay products more often have opposite sign than the same sign. An explanation is charge conservation: the charge of the quark which radiates the W boson is of opposite sign than the W boson. However, the quark charge sign identification via the leading particle charge in the jet (jet charge) is less efficient for QCD jets than for τ jets.

Missing Transverse Energy $E_{\mathrm{T}}^{\mathrm{miss}}$

The neutrinos originating from the decay of the two τ leptons in the Higgs decay result in a significant amount of missing transverse energy. By requiring a minimum value of missing transverse energy in the final state, most of the background processes without neutrinos can be rejected. Figure 5.15 shows the distribution of the reconstructed missing transverse energy for the signal and for dominant background processes. In the $Z \to ee$ and $Z \to \mu\mu$ events, the measured missing transverse energy is expected to be small, solely originating from the finite detector resolution. By requiring $E_{\mathrm{T}}^{\mathrm{miss}} > 40\,\mathrm{GeV}$ in the leptonic Higgs decay mode, a large fraction of the Z background is rejected. For the semileptonic decay channel, the minimal missing transverse energy requirement is lowered to $E_{\mathrm{T}}^{\mathrm{miss}} > 30\,\mathrm{GeV}$ in order to account for the smaller number of neutrinos emitted.

Collinear Approximation

Employing the collinear approximation as described in Section 5.3 it is possible to determine the four-momenta of the two τ leptons provided that the two visible τ-decay products are not emitted

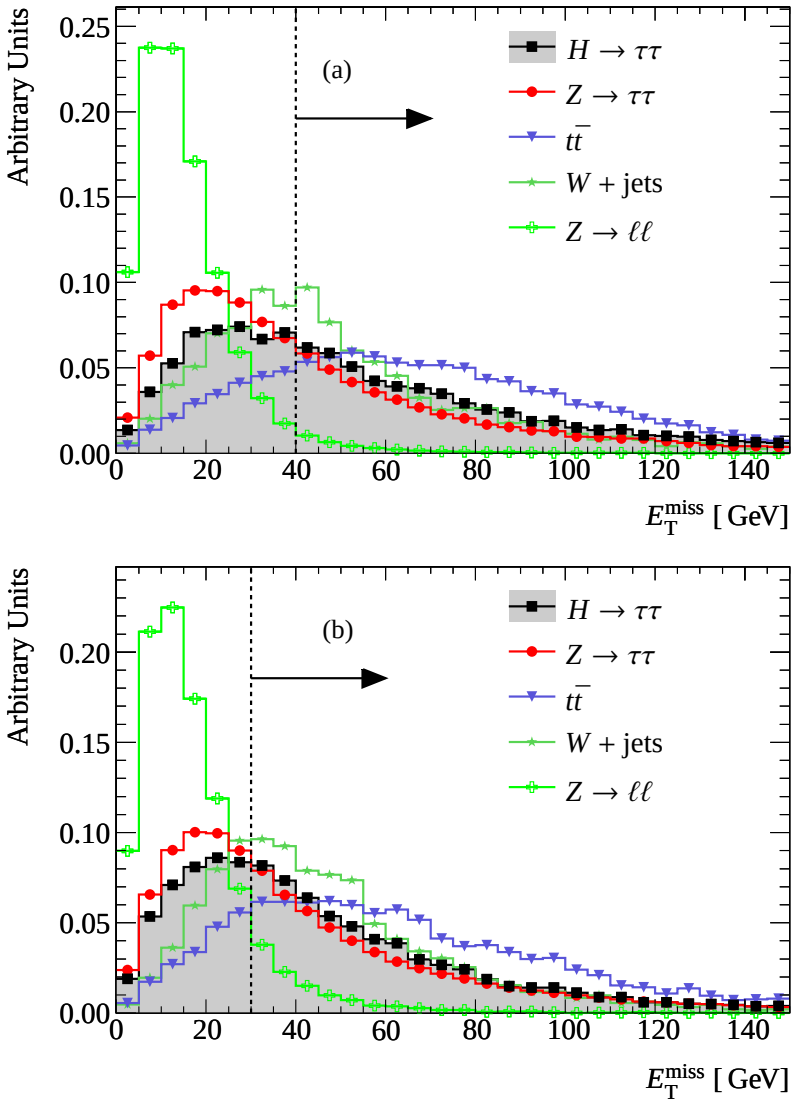


Figure 5.15: Missing transverse energy distribution (a) for the leptonic and (b) the semileptonic Higgs decay channel for $m_H = 120\,\text{GeV}$ and for the dominant background processes. A cut on the missing transverse energy is indicated by the dashed line.

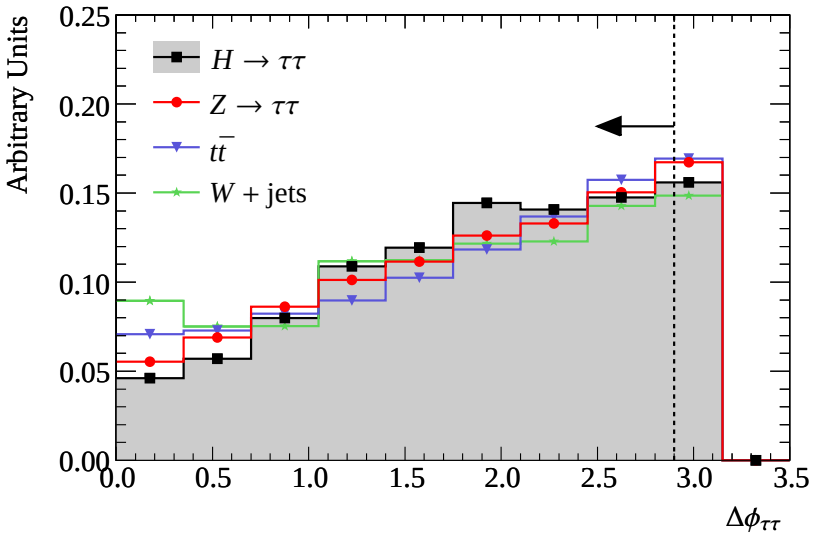


Figure 5.16: Distribution of the azimuthal separation $\Delta\phi_{\tau\tau}$ between the τ-lepton decay products for the signal and dominant background processes. The cut $\Delta\phi_{\tau\tau} \leq 2.9$ indicated by the dashed line ensures the applicability of the collinear approximation.

back-to-back. Therefore, a cut $|\Delta\phi_{\tau\tau}| < 2.9$ is applied to ensure a solvable system of equations. Figure 5.16 shows the distribution of $\Delta\phi_{\tau\tau}$ for the signal and the dominant background processes. All distributions look very similar and thus it is expected that $\Delta\phi_{\tau\tau}$ has no significant separation power.

The visible momentum fractions χ calculated in the collinear approximation (equation 5.3) allow for reducing some of the background contributions. In the leptonic decay mode, χ_1 represents the momentum fraction carried by the lepton with the higher transverse momentum. In the semileptonic case, χ_h represents the momentum fraction carried by the τ jet and χ_ℓ the momentum fraction of the lepton.

The visible τ-decay products cannot have a larger momentum than the τ lepton itself and are assumed to be emitted parallel to the τ lepton. Therefore, the physical values of the variables χ should be in the range $0 < \chi < 1$. By this constraints on the momentum fraction, a large fraction of the $t\bar{t}$, WW and W + jets backgrounds can be rejected. The distribution of the visible momentum fractions χ is shown in Figure 5.17 for the signal and dominant background processes.

$Z \to \tau\tau$ events have a similar topology as the signal events and thus also have similar χ distributions. For the $Z \to ee$ and $Z \to \mu\mu$ events, the χ distribution peaks around one as expected since no neutrinos are present in those events. The very low $E_{\mathrm{T}}^{\mathrm{miss}}$ results in low neutrino momenta obtained from the collinear approximation and thus in visible momentum fractions close to one. For other backgrounds, the visible momentum fraction has a broader distribution than for the signal with many unphysical solutions below zero or above one. In case of the semileptonic channel the visible momentum fraction χ_h of the τ jet in the signal events shows a higher mean value than for the leptonic final state which is expected since in the hadronic τ decay only one neutrino is

emitted.
The differences between the signal and the different background processes are even more pronounced in the two-dimensional $\chi_{1,2}$ distributions as shown in Figure 5.19.

Transverse Mass

The invariant mass of the transverse components of one lepton momentum and of the missing transverse energy, the transverse mass $m_{\mathrm{T}}^{\ell\nu}$, is given by:

$$m_{\mathrm{T}}^{\ell\nu} = \sqrt{2 {p_{\mathrm{T}}}^{\ell} E_{\mathrm{T}}^{\mathrm{miss}} \cdot (1 - cos\Delta\phi)}. \tag{5.11}$$

In the leptonic decay mode, ${p_{\mathrm{T}}}^{\ell}$ is from the lepton (electron or muon) with the higher transverse momentum p_{T}. In the semileptonic decay mode, ${p_{\mathrm{T}}}^{\ell}$ is the transverse momentum of the electron or muon. The angle $\Delta\phi$ is the azimuthal angle between the lepton and the missing transverse energy. The distribution of the transverse mass is shown in Figure 5.20. For the W + jets events, the calculated mass value corresponds to the transverse component of the W boson mass. One can see that the W + jets events have a higher transverse mass than the signal or the other background processes except $t\bar{t}$, with the trailing edge of the distribution close to the W boson mass (80.4 GeV [22]).
Therefore, requiring the transverse mass to be less than 30 GeV helps to suppress mainly the W + jets background, but also a large fraction of the $t\bar{t}$ events is discarded. This discriminating variable is particularly important for the semileptonic decay mode where it helps to suppress W + jets events in which the W boson decays leptonically with additional QCD radiation faking a τ jet. In the leptonic decay mode, the W + jets background is much less important and, thus, the cut on $m_{\mathrm{T}}^{\ell\nu}$ can be looser.

5.4.2 Tagging Jets Criteria

Tagging Jet Candidates

The presence of at least two reconstructed jets in the final state is required to account for the vector-boson fusion signature with two jets in the forward region of the detector. The number of reconstructed jets with $p_{\mathrm{T}} > 20$ GeV is shown in Figure 5.21.
In contrast to the high p_{T} jets of the vector-boson fusion process, jets produced by QCD processes have a steeply falling p_{T} distribution. In order to suppress W and Z events which contain QCD jets it is required that at least one of the two jets has a minimum transverse momentum of $p_{\mathrm{T}} > 40$ GeV. The two jets with the highest p_{T} are considered as candidates for the tagging jets. Since in the vector-boson fusion process the jets are mainly produced in opposite hemispheres of the detector, all events containing these two jets in the same hemisphere are rejected.

Tagging Jet Kinematics

The two jets produced in the vector-boson fusion process typically have a large rapidity gap ($\Delta\eta_{jj}$) and high energy and thus a large invariant mass (m_{jj}). In contrast, jets from QCD processes are more located in the central region of the detector. Figure 5.22a shows the distribution of the

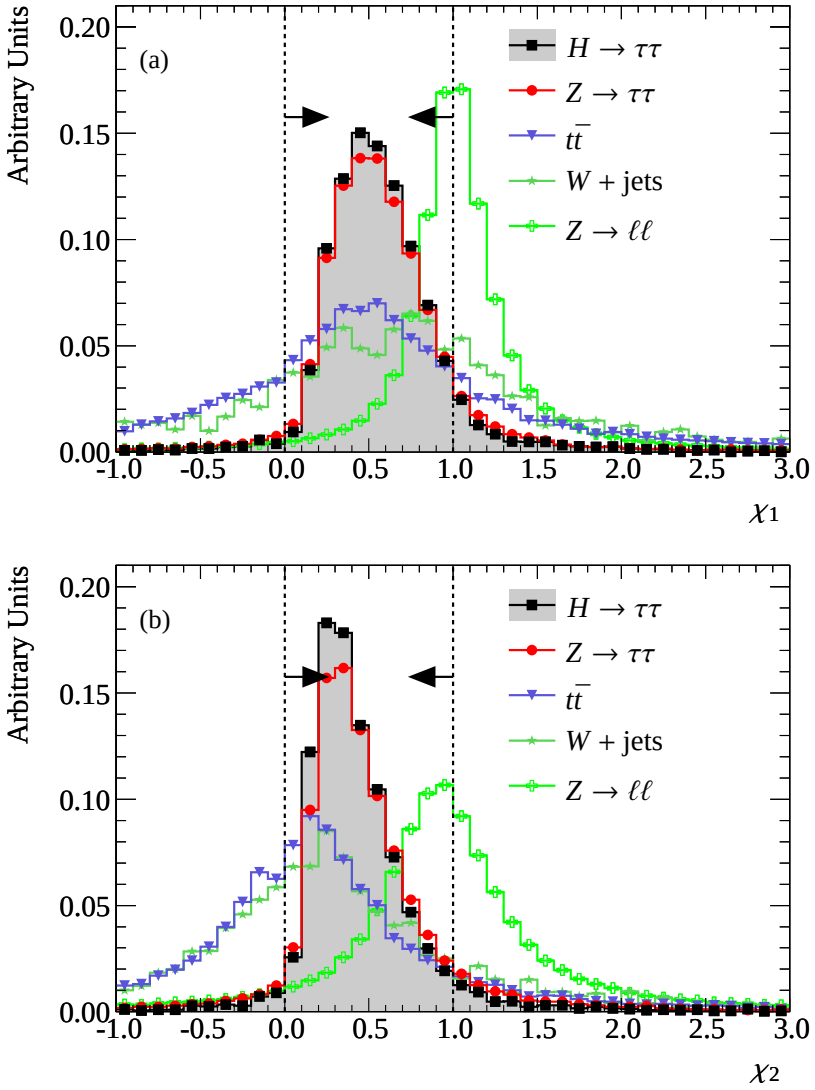


Figure 5.17: Distributions of the visible momentum fractions $\chi_{1,2}$ in the collinear approximation for the leptonic decay mode. (a) χ_1 is the momentum fraction of the lepton with the higher transverse momentum, (b) χ_2 the one for the lepton with the lower transverse momentum. The applied cuts $0 < \chi_{1,2} < 1$ are indicated as dashed lines.

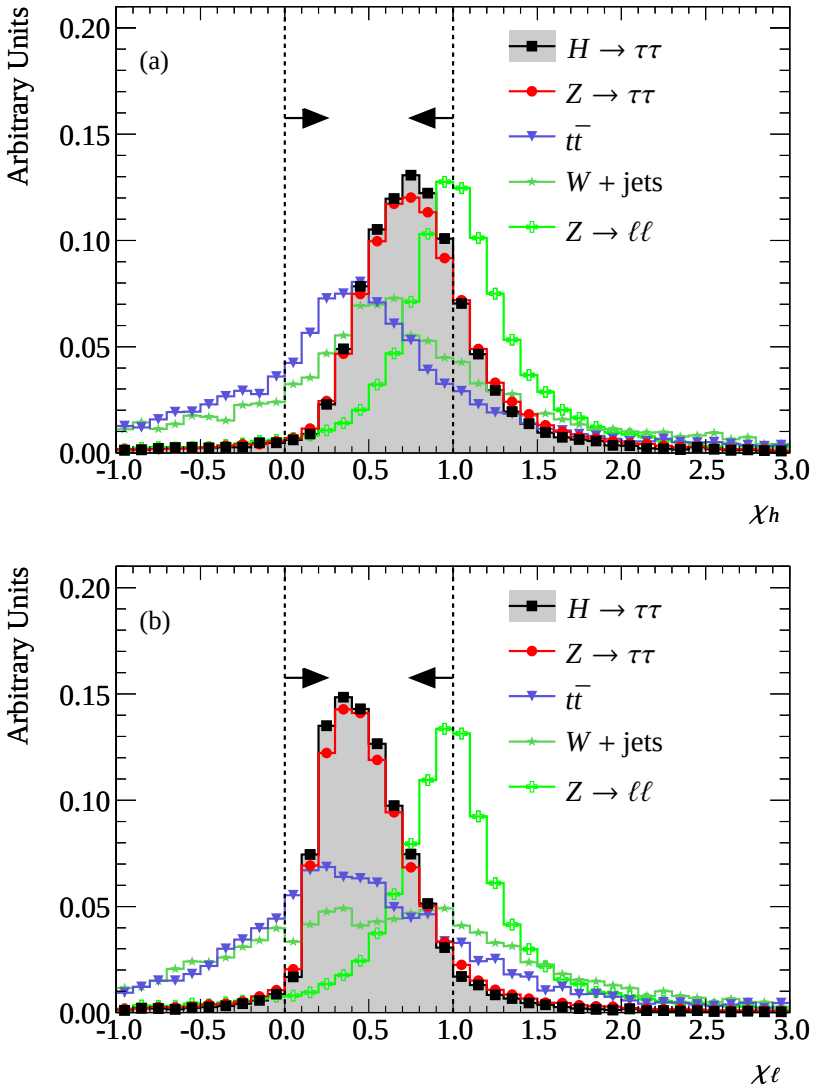


Figure 5.18: Distribution of the visible momentum fractions $\chi_{\ell,h}$ in the collinear approximation for the semileptonic decay mode. (a) χ_h is the momentum fraction of the τ jet and (b) χ_ℓ the momentum fraction of the lepton. The applied cuts $0 < \chi_{\ell,h} < 1$ are indicated as dashed lines.

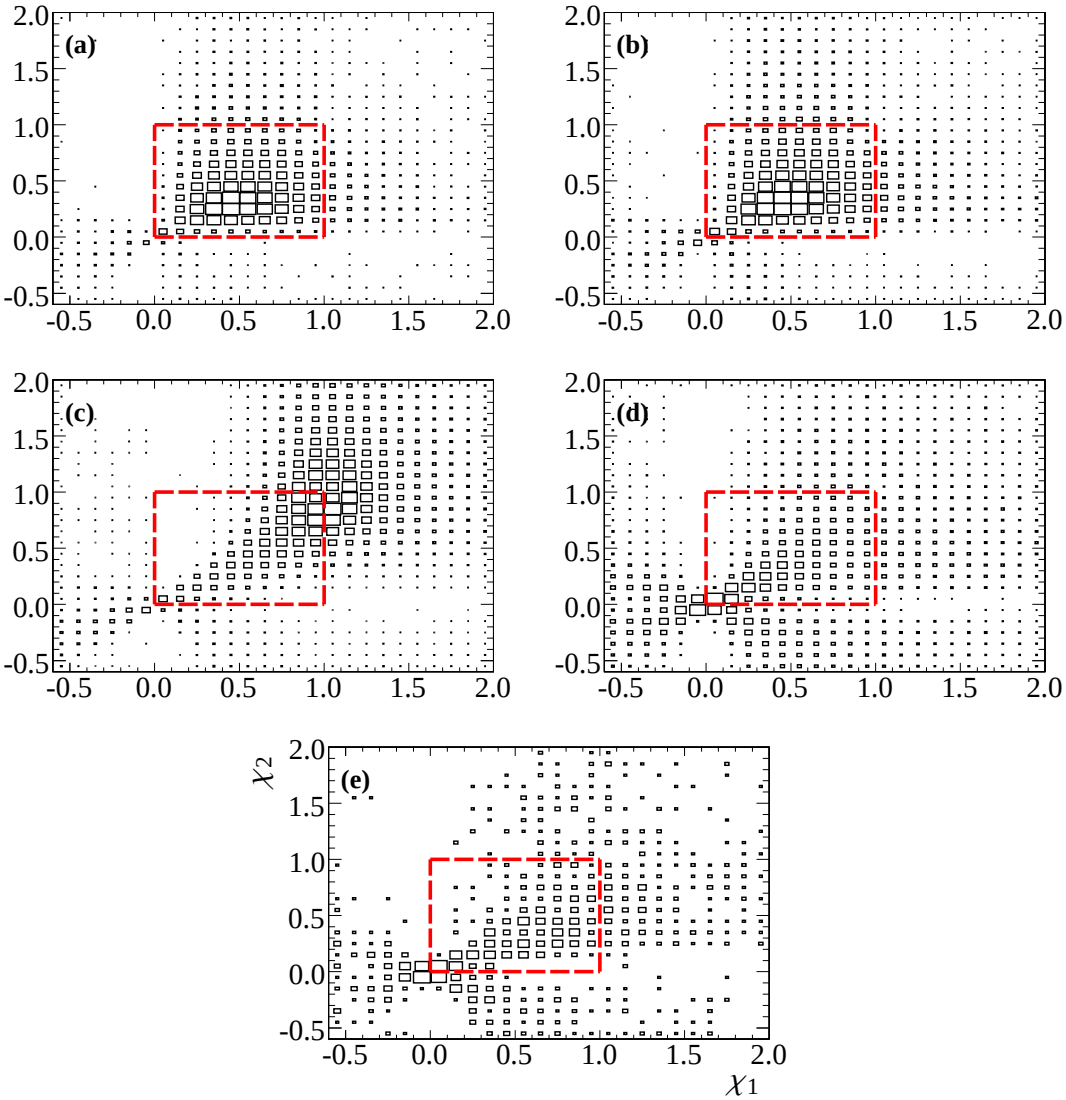


Figure 5.19: Correlations between the visible momentum fractions χ_1 and χ_2 for the leptonic final state shown for (a) $H \to \tau\tau$ with m_H=120 GeV, (b) $Z \to \tau\tau$, (c) $Z \to \ell\ell$, (d) $t\bar{t}$ and (e) W + jets events. The dashed box indicates the physical range $0 < \chi_{1,2} < 1$ in the collinear approximation.

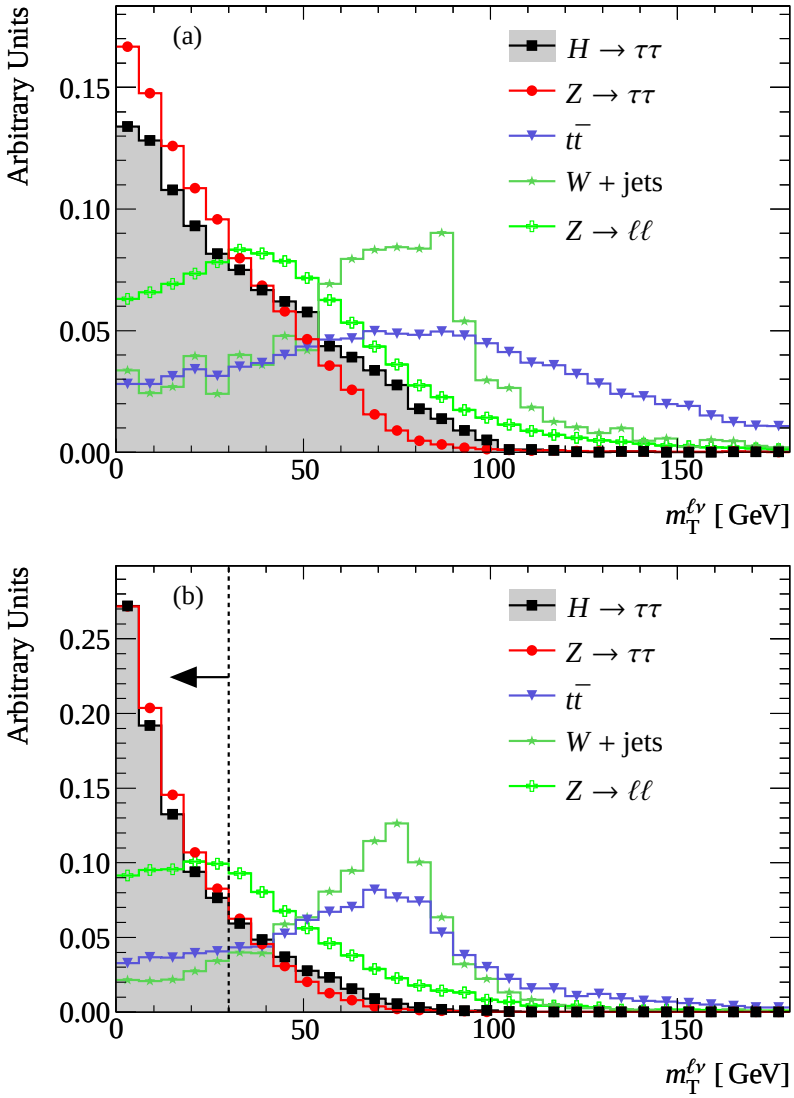


Figure 5.20: Distribution of the transverse mass $m_T^{\ell\nu}$ (a) for the leptonic and (b) the semileptonic Higgs boson decay. The applied cut in the semileptonic decay mode is indicated by the dashed line.

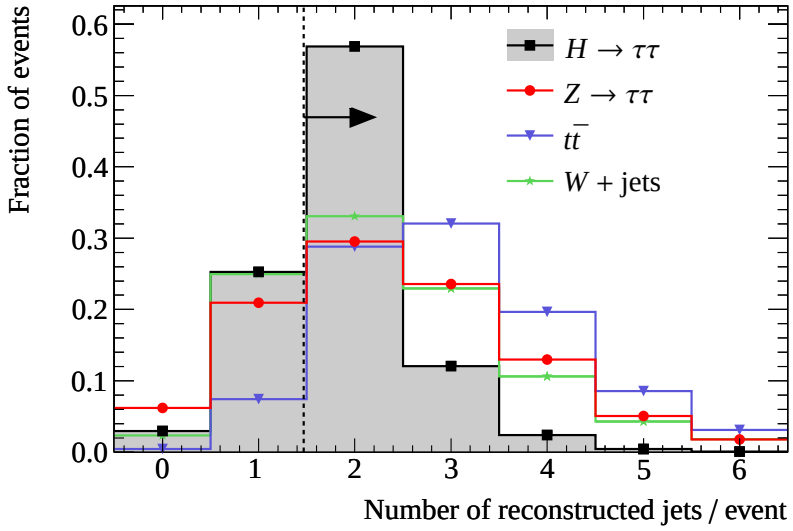


Figure 5.21: Number of reconstructed jets with $p_T > 20\,\text{GeV}$ per event for the signal and dominant backgrounds. At least two jets are required in each event as indicated by the dashed line in order to fulfill the vector-boson fusion signature.

pseudorapidity gap $\Delta\eta_{jj}$ and Figure 5.22b the invariant mass m_{jj} distribution of the tagging jet pair for the signal and the background processes. Requiring a minimum pseudorapidity gap $\Delta\eta_{jj}$ and a minimum invariant jet pair mass are the most important cuts to suppress all background contributions.

A potential additional discriminating variable is the ϕ-separation of the tagging jets $\Delta\phi_{jj}$, which is shown in Figure 5.23 for the signal and the dominant background processes. A cut on $\Delta\phi_{jj}$ increases the signal-to-background ratio only by a small amount. However, this cut allows for a discrimination against the otherwise very signal-like electroweak $Z \to \tau\tau$ production.

In addition, requiring $\Delta\phi_{jj}$ to be less than some value can also help to suppress events from double parton scattering, where one parton from each colliding proton participates in the hard process (e. g. Z boson production) and the interaction of a second parton pair from the same protons creates a jet pair. The jets from the second interaction tend to be emitted back-to-back. First studies show that the additional contribution to the $Z \to \tau\tau$ background from double parton scattering is in the order of 15 % [65, 66]. Since no simulations of double parton scattering events have been available for this analysis, the suppression of this background by cutting on $\Delta\phi_{jj}$ could not be studied.

It was also proposed that the $\Delta\phi_{jj}$ distribution can provide information about the structure of the coupling of the Higgs boson to weak gauge bosons [67, 68].

b-Jet Veto

In the signal events, no b quarks and therefore no b jets are present. In contrast, $t\bar{t}$ events are characterized by two b quarks originating from the top decays. Thus, by vetoing against events

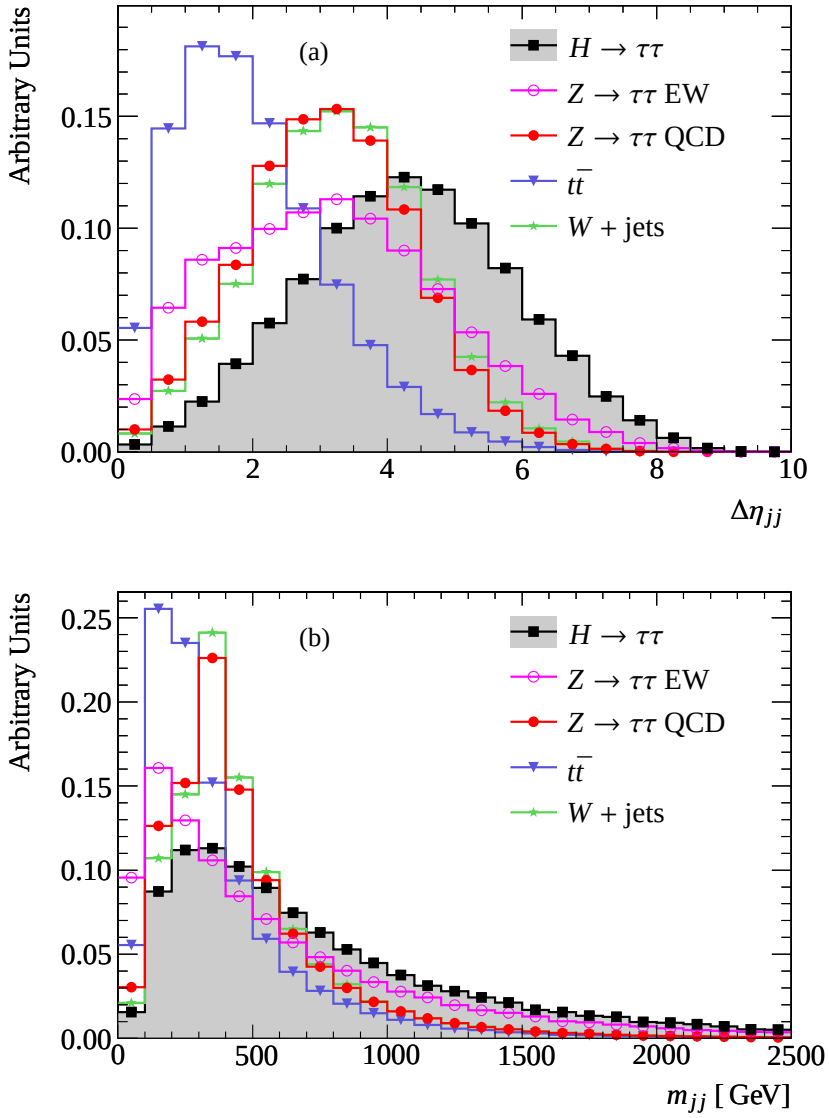


Figure 5.22: Distributions of (a) the pseudorapidity gap $\Delta\eta_{jj}$ and (b) the invariant mass m_{jj} of the two tagging jets for the signal and dominant background processes.

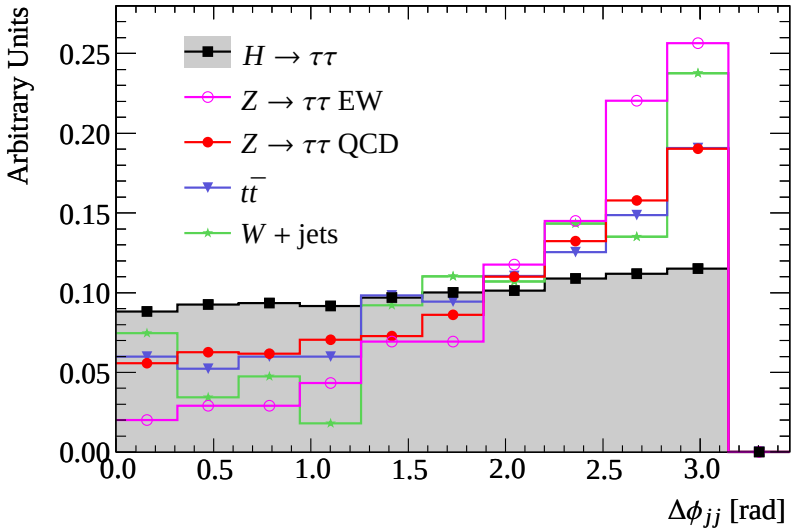


Figure 5.23: Azimuthal angle separation $\Delta\phi_{jj}$ between the two tagging jets.

in which a jet is tagged as a b jet helps to suppress the $t\bar{t}$ background. The b-jet identification performance of the ATLAS detector is described in Chapter 5.2.6. Figure 5.24 shows the b-tagging weight distribution of the most likely b jet in an event for the signal and background processes. A large fraction of the $t\bar{t}$ events can be rejected by requiring for all jets a b-tagging weight smaller than one.

5.4.3 Overall Event Topology (Jets and τ-Decay Products) Criteria

Jet – τ-Decay Product Separation $\Delta\eta_{j\tau}$

The Higgs production via vector-boson fusion is characterized by the Higgs decay products being emitted within the rapidity gap spanned by the two tagging jets. There are two visible τ-decay products and two tagging jets. Two parameters are defined to describe the η separation between the jets and the τ-decay products: the distance $\Delta\eta^1_{j\tau}$ between the jet with the smaller η and the τ-decay product with the smaller η and the distance $\Delta\eta^2_{j\tau}$ between the jet and the τ-decay product with the higher η values:

$$\begin{aligned} \Delta\eta^1_{j\tau} &= min\{\eta_{\tau 1}, \eta_{\tau 2}\} - min\{\eta_{j1}, \eta_{j2}\}, \\ \Delta\eta^2_{j\tau} &= max\{\eta_{j1}, \eta_{j2}\} - max\{\eta_{\tau 1}, \eta_{\tau 2}\}. \end{aligned}$$

The parameter with the smaller value is called $\Delta\eta^{\text{min}}_{j\tau}$, the other $\Delta\eta^{\text{max}}_{j\tau}$:

$$\begin{aligned} \Delta\eta^{\text{min}}_{j\tau} &= min\{\Delta\eta^1_{j\tau}, \Delta\eta^2_{j\tau}\}, \\ \Delta\eta^{\text{max}}_{j\tau} &= max\{\Delta\eta^1_{j\tau}, \Delta\eta^2_{j\tau}\}. \end{aligned}$$

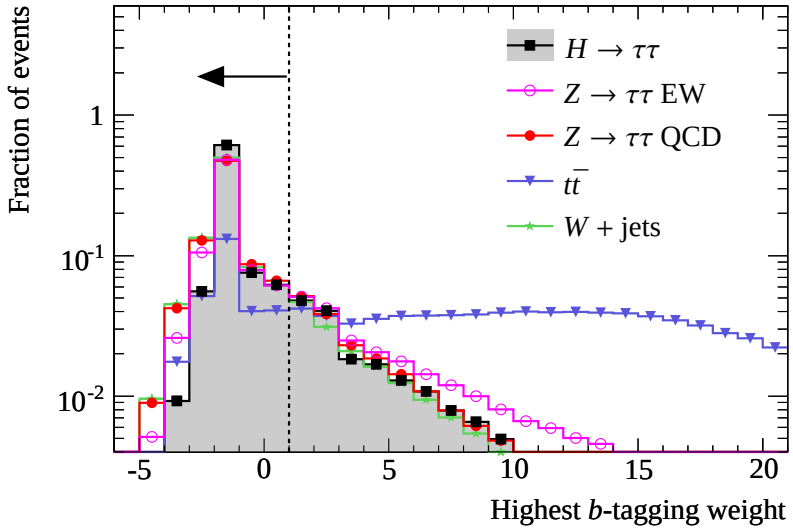


Figure 5.24: Distribution of the largest b-tagging weight in the event. Jets in $t\bar{t}$ events, the only source of real b jets, tend to have much higher values than the jets in the signal and other backgrounds events. The b-jet veto cut (b-tagging weight ≤ 1) is indicated by the dashed line.

Figure 5.25 shows the $\Delta\eta_{j\tau}^{\text{min}}$ and the $\Delta\eta_{j\tau}^{\text{max}}$ distribution for the signal and dominant backgrounds. Clearly, these variables can help to distinguish signal and background processes.
Obviously, $\Delta\eta_{j\tau}^{\text{min}}$ and $\Delta\eta_{j\tau}^{\text{max}}$ are correlated with the pseudorapidity distance $\Delta\eta_{jj}$ between the tagging jets as shown in Figure 5.26.

Central Jet Veto

One of the most important analysis cuts is related to a feature of the vector-boson fusion processes, namely that no color flow exists between the quarks. This manifests itself in the suppression of jet activity in the central region of the detector. A veto against events with additional central jets therefore strongly improves the signal-to-background ratio. All jets in addition to the tagging jets with a transverse momentum of $p_{\text{T}} > 20\,\text{GeV}$ and a pseudorapidity of $|\eta| < 3.2$ are considered as central jets. Figure 5.27 shows the number of central jets.

Transverse Momentum Balance

Another way to account for the low central jet activity in the vector-boson fusion process is to consider the total transverse momentum $p_{\text{T}}^{\text{tot}}$ in the event. This is the sum of the transverse momenta of the reconstructed objects assumed to result from the hard scattering: the two tagging jets, the two τ-decay products and the missing transverse energy.

$$p_{\text{T}}^{\text{tot}} = p_{\text{T}}^{j1} + p_{\text{T}}^{j2} + p_{\text{T}}^{\ell 1} + p_{\text{T}}^{\ell 2} + E_{\text{T}}^{\text{miss}} \tag{5.12}$$

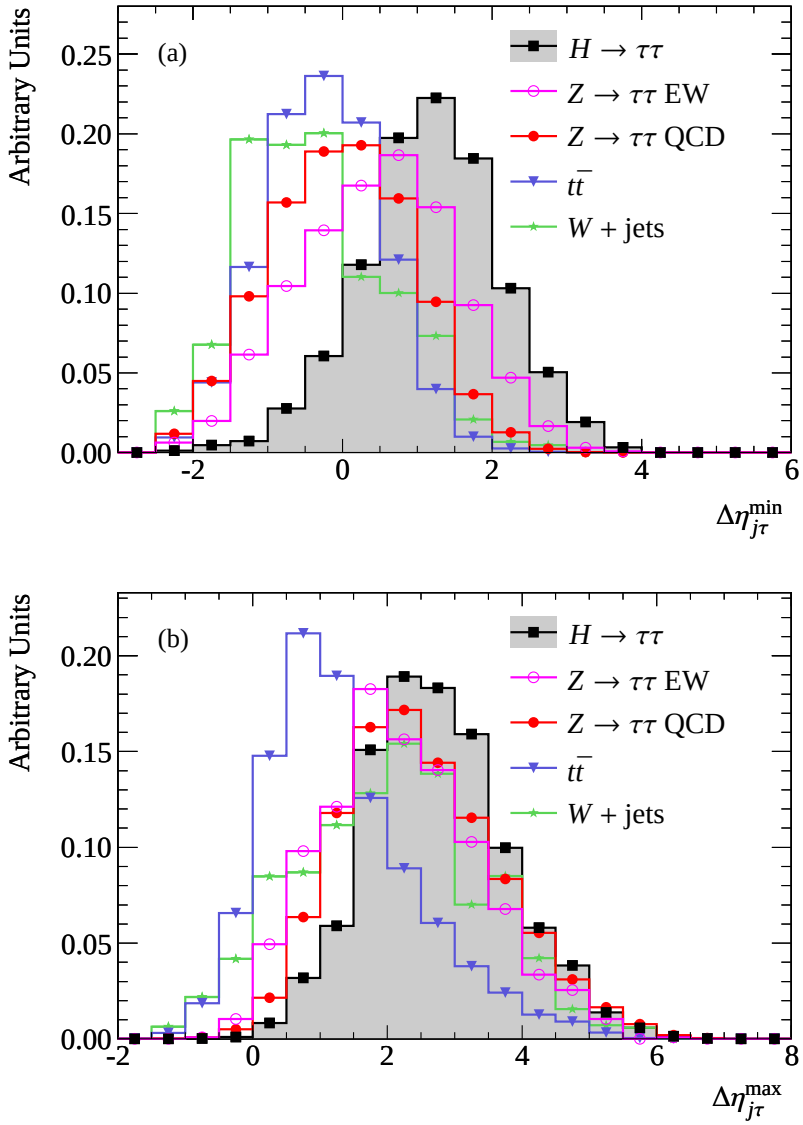


Figure 5.25: Pseudorapidity separation between the two tagging jets and the τ-decay products. In (a) the smaller separation $\Delta\eta_{j\tau}^{\text{min}}$ and in (b) the larger separation $\Delta\eta_{j\tau}^{\text{max}}$ is shown.

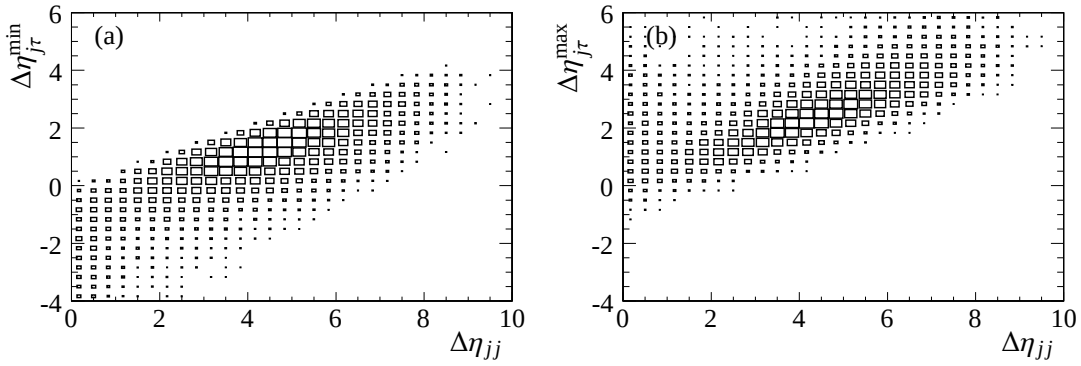


Figure 5.26: Pseudorapidity separation between the two tagging jets and the τ-decay products as a function of $\Delta\eta_{jj}$. In (a) the $\Delta\eta_{j\tau}^{\min}$- and in (b) the $\Delta\eta_{j\tau}^{\max}$-correlation with $\Delta\eta_{jj}$ is shown for the signal.

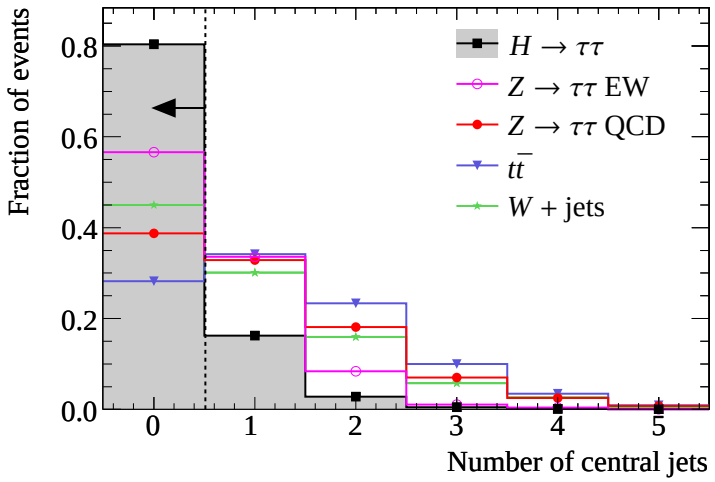


Figure 5.27: Number of central jets (jets with $|\eta| < 3.2$, $p_{\mathrm{T}} > 20\,\mathrm{GeV}$ and not tagging jets) for signal and background processes. The dashed line indicates the central jet veto cut.

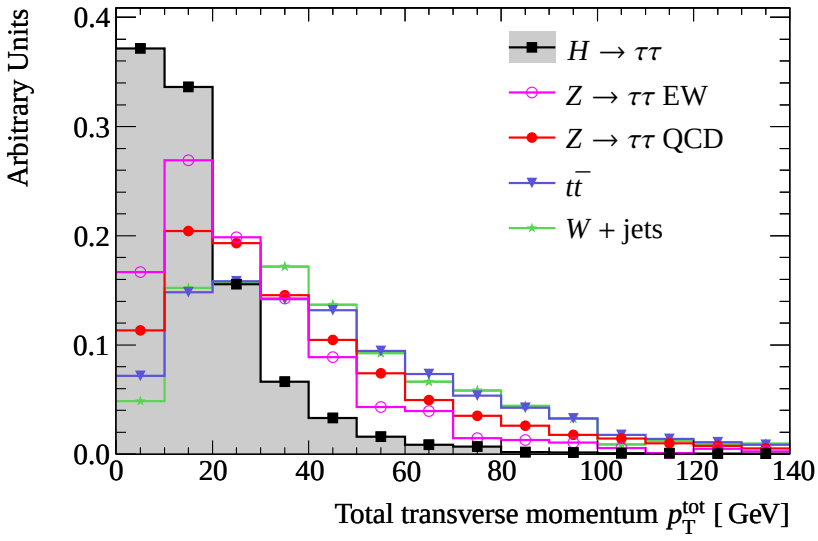


Figure 5.28: Distribution of the total transverse momentum p_T^{tot}, the sum of the transverse momenta of the two tagging jets, the two τ-decay products and the missing transverse energy.

The distribution of p_T^{tot} is shown in Figure 5.28.
If no other objects are present in the event, p_T^{tot} should be small, ideally zero. On the other hand, p_T^{tot} is higher in events with additional jet activity. Clearly, p_T^{tot} is highly correlated with the previously mentioned central jet multiplicity as illustrated in Figure 5.29. The figure shows the p_T^{tot} distribution separately for events with and without central jets in the final state.
In former analyses, the central jet veto has been applied to suppress backgrounds with QCD radiation. However, an additional pile-up jet in a high luminosity environment which does not originate from the hard-scattering process of the Higgs boson production could trigger the central jet veto while the p_T^{tot} value remains unchanged. Thus, the p_T^{tot} variable may prove to be more robust against pile-up effects.

Higgs Mass Window

As previously discussed, the invariant mass of the Higgs boson in the $H \to \tau\tau$ decay channel can be reconstructed using the collinear approximation. This allows for a further background suppression. As shown in Figure 5.12b, the expected standard deviation of the Higgs mass reconstruction is approximately $\sigma \sim 10\,\text{GeV}$. Rejecting all events outside a mass window of $\pm 15\,\text{GeV}$ around the expected Higgs mass significantly reduces all backgrounds while keeping $\sim 80\,\%$ of the signal events.
The invariant $\tau\tau$ mass distribution of the signal and backgrounds is shown in Figure 5.30. The τ candidate pairs $t\bar{t}$ and W + jets events are produced non-resonant and therefore show a flat distribution. The signal distribution is shown for a Higgs mass of $m_H = 120\,\text{GeV}$. The smaller the Higgs boson mass, the more difficult it becomes to separate the signal from the $Z \to \tau\tau$ background.

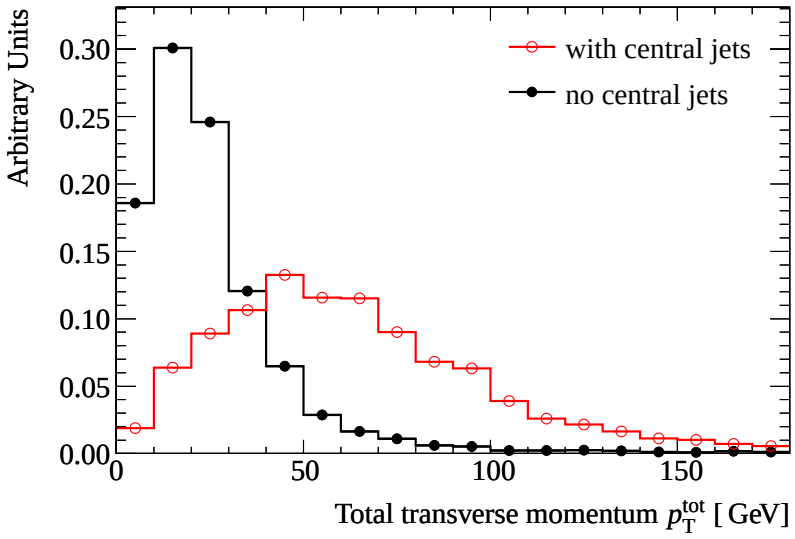


Figure 5.29: Distribution of the total transverse momentum p_T^{tot}, shown separately for $t\bar{t}$ events with and without an additional central jet.

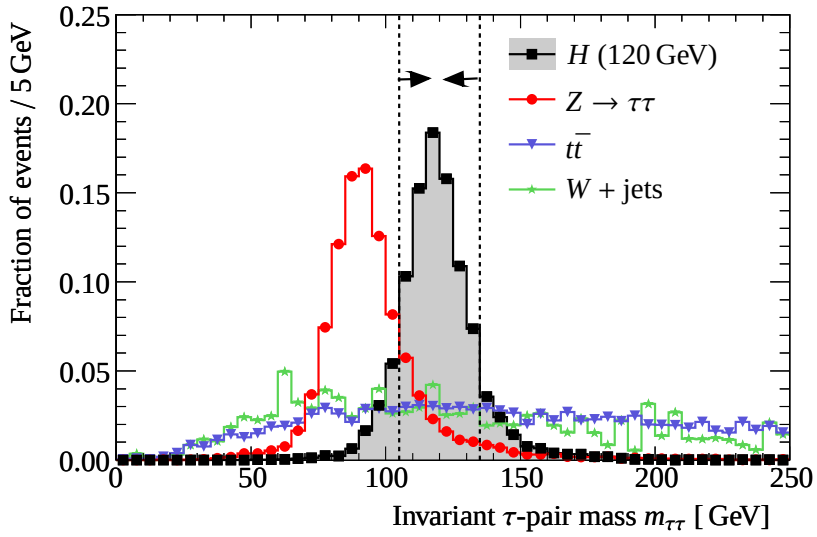


Figure 5.30: Invariant τ pair mass distribution (normalized to unity) obtained by using the collinear approximation, where the Higgs boson was simulated with a mass of m_H = 120 GeV. The mass window used ($\Delta m_{\tau\tau} = m_H \pm 15$ GeV) for the event selection is indicated by the dashed lines.

Chapter 6

Optimization of the Signal Selection

The variables described in Section 5.4 can be used to separate signal and background events. An optimization of the cut values was performed in order to achieve a maximum signal significance. The cut optimization was done by scanning over the range of possible cut values for all variables. The set of cut values that results in a maximum of a predefined measure of the signal versus background separation is considered to be the optimal one.
Several issues have to be addressed in order to obtain reliable cut values: the composition of the background processes, the preselection criteria, the definition of the measure of the signal-to-background separation to be maximized, the set of variables used in the optimization and the methods for treating correlated variables. These aspects are discussed in the following.

6.1 Composition of the Background

The background consists of all processes that can be potentially misidentified as signal events. However, some of these processes could either not be simulated at all or with very low statistics only. For example, due to the low number of simulated events, including the W + jets process in the background calculation would result in a high statistical uncertainty in the background prediction. Therefore, only the two most important background processes which have been simulated with sufficient statistics have been taken into account: $Z \rightarrow \tau\tau$ and $t\bar{t}$. For the $Z \rightarrow \tau\tau$ background, the QCD as well as the electroweak production have been taken into account. The background samples have been combined by weighting them according to their cross-sections.

6.2 Preselection

In order to be able to calculate all discriminating variables, a few requirements are necessary. In addition, several cuts have been fixed or the range of the possible cut values has been limited before the optimization procedure for reasons discussed below. The preselection cuts are:

- Two visible τ-decay products with opposite charge.

- The cut on the missing transverse energy is limited to the range $E_{\mathrm{T}}^{\mathrm{miss}} > 40\,\mathrm{GeV}$ in the leptonic channel and to $E_{\mathrm{T}}^{\mathrm{miss}} > 30\,\mathrm{GeV}$ in the semileptonic channel. Hence, the strong suppression of all processes with no or only very little missing transverse energy (like $Z \to ee, \mu\mu$) is ensured.

  Due to the small amount of simulated data available for these processes, it has not been possible to reliably determine the rejection of these events by all the other cuts. To be conservative, a minimum missing transverse energy was required. The lower minimum missing transverse energy cut in the semileptonic analysis reflects the fact that in the hadronic decay of a τ lepton only one neutrino is involved leading to a lower missing energy in the event.

- The range for the lepton momentum fractions $\chi_{1,2}$ is limited to the physical range from zero to one. To allow for the collinear approximation, events with back-to-back leptons are rejected by limiting the $\Delta\phi_{\tau\tau}$ cut value to the range $\Delta\phi_{\tau\tau} < 2.9$.

- In the semileptonic analysis, W + jets is an important background while it is suppressed in the leptonic analysis due to the low electron and muon misidentification rate. Also the number of simulated W + jets events was not sufficient to be included in the cut optimization. Therefore, a preselection cut $m_{\mathrm{T}}^{\ell\nu} < 30\,\mathrm{GeV}$ is introduced to guarantee a strong suppression of these events and the cut optimization is limited to the range below $m_{\mathrm{T}}^{\ell\nu} < 30\,\mathrm{GeV}$.

- To account for the vector-boson fusion signature, two tagging jets in opposite hemispheres of the detector are required and both visible τ-decay products have to be located within the rapidity gap of these two jets.

- Since the central jet veto highly depends on the level of pile-up while data samples including pile-up were not available with sufficient statistics for all the background processes, the central jet veto is not included in the optimization and has therefore been applied at preselection level.

  However, a possibility to take the impact of pile-up into account is to use inner detector information to distinguish jets originating from separated vertices.

  This approach has not been followed in the frame of this thesis but first studies [69] indicate that this method allows for a central jet veto independent of the level of pile-up.

- The b-jet veto is not included in the optimization as too little $t\bar{t}$ events with detailed detector simulation have been available. For the cut optimization, the fast detector simulation was used for the $t\bar{t}$ background. The fast simulation does not include a b-tagging algorithm providing a continuously distributed variable for identification of b jets (see Figure 5.24), instead it assigns b jets to b quarks on a statistical basis. Therefore, there is no possibility to optimize the b-jet veto and thus, it has been applied on the preselection level. Since the $t\bar{t}$ background contribution is less important in the semileptonic than in the leptonic channel, the b-jet veto cut has only been applied in the leptonic analysis.

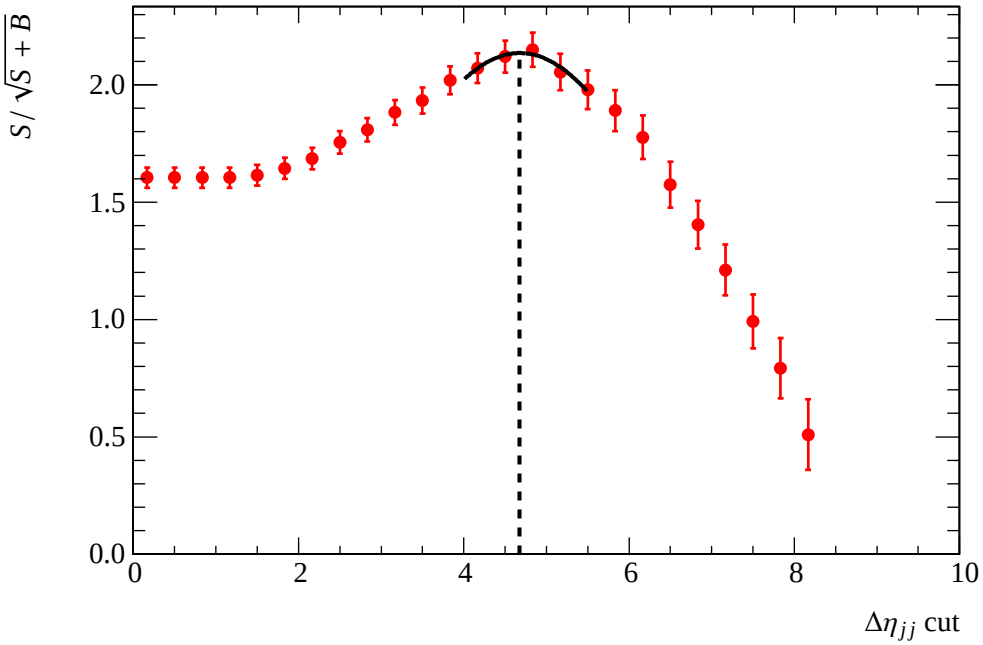


Figure 6.1: Optimization curve for the $\Delta\eta_{jj}$ variable: the ratio $S/\sqrt{S+B}$ as a function of the cut value on $\Delta\eta_{jj}$. Around the maximum at $\Delta\eta_{jj} = 4.7$ (dashed line) a Gauss function was fitted and the mean of the Gauss function was taken as the optimal cut value.

6.3 Signal-to-Background Ratio

The optimal cut value for a discriminating variable has been obtained by maximizing the ratio $S/\sqrt{S+B}$ [1], with S representing the number of signal events and B the number of background events. S and B have been calculated in a τ-pair mass window around the simulated Higgs mass m_H and correspond to an integrated luminosity of 30 fb^{-1}. In order to obtain robust optimization results the events have been counted within a loose mass window cut of $m_{\tau\tau} = m_H \pm 30$ GeV.
The optimal cut value for a variable is obtained by plotting the $S/\sqrt{S+B}$ ratio as a function of the cut value (see Figure 6.1 as example). A Gauss-function has been fitted around the maximum of the curve and the mean of the Gaussian has been taken as the optimal cut value.
The described cut optimization has been performed for each of the previously described discriminating variables. In Figure 6.2, the optimization curves are shown for the variables which lead to an increase of the $S/\sqrt{S+B}$ ratio, while for the variables in Figure 6.3 the $S/\sqrt{S+B}$ ratio does not improve by applying a cut on them. Each variable has been tested without additional requirements other than the mentioned preselection cuts. Therefore, possible correlations between cut variables are not taken into account in this procedure. The treatment of correlations is discussed in the following.

[1] This ratio is one possible definition of the signal significance (Section 8.1).

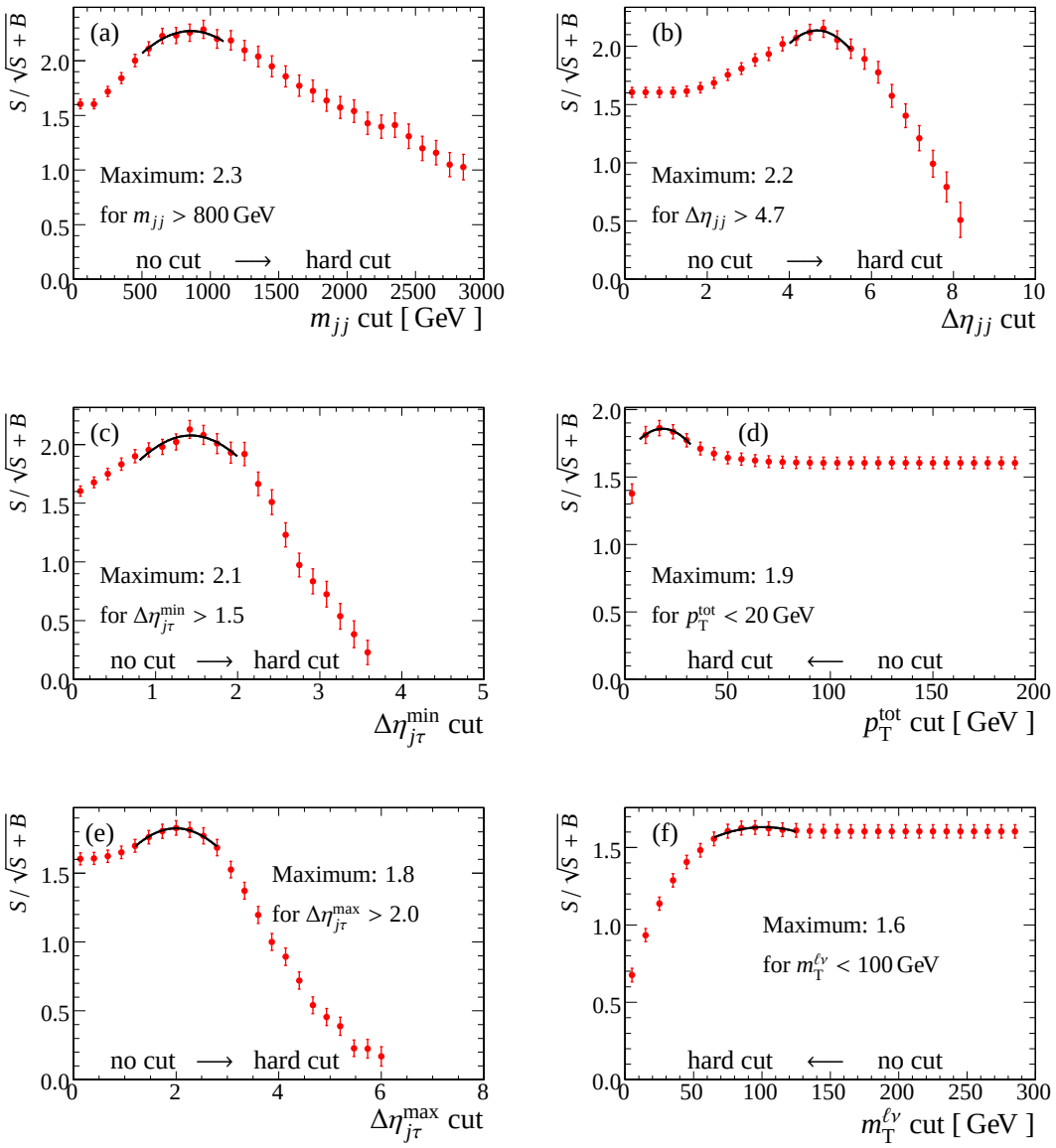


Figure 6.2: The ratio $S/\sqrt{S+B}$ as a function of the cut value for selected discriminating variables (see text). Applying cuts on these variables improves $S/\sqrt{S+B}$. The Gaussian fits around the maxima of the functions are indicated.

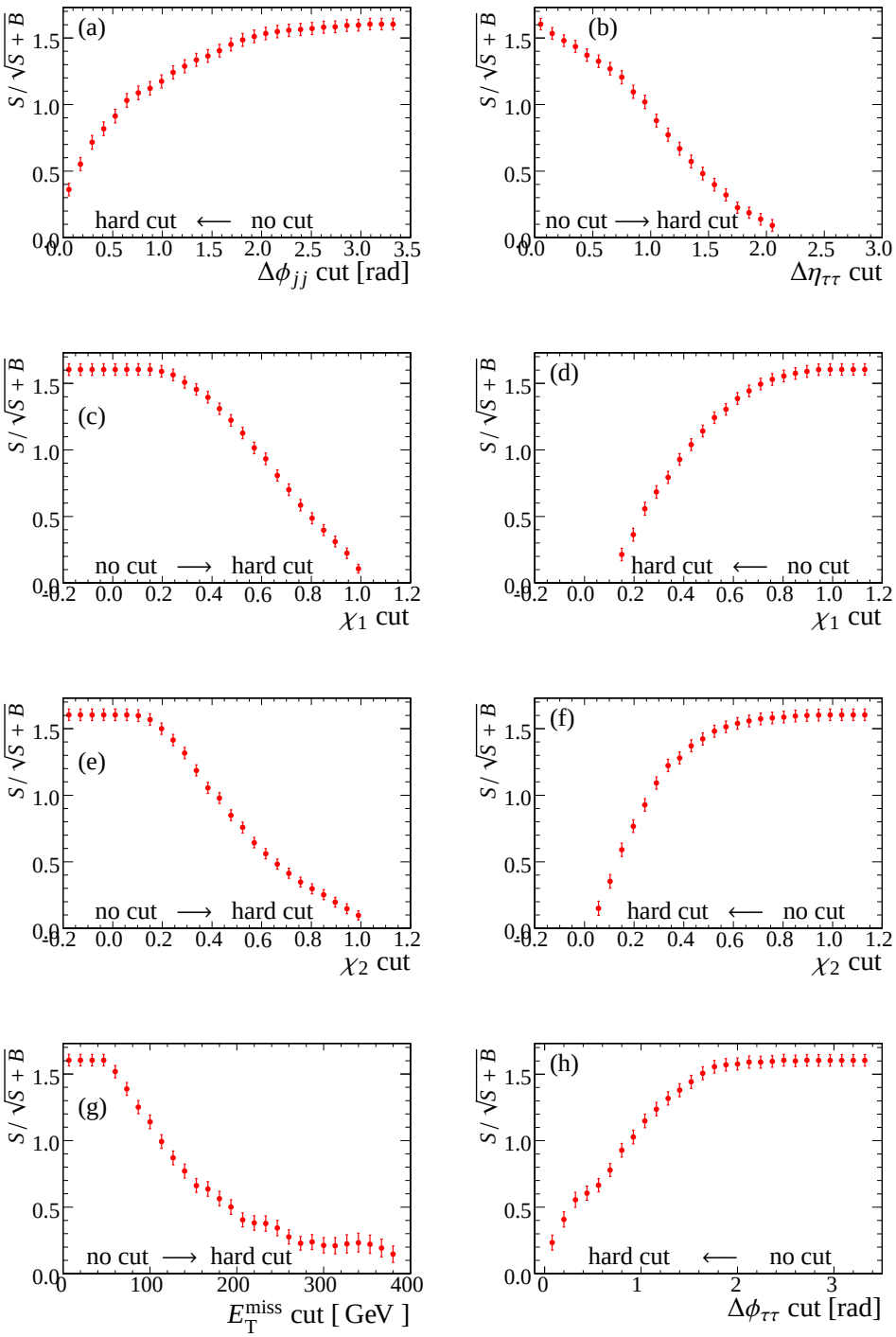


Figure 6.3: The ratio $S/\sqrt{S+B}$ as a function of the cut value for selected discriminating variables (see text). Cuts on these variables do not improve $S/\sqrt{S+B}$.

6.4 Treatment of Correlations

Two different approaches have been considered:

6.4.1 Parallel Cut Optimization

The most simple approach is to apply cuts only on the discriminating variables for which the cut improves $S/\sqrt{S+B}$ (the variables shown in Figure 6.2: m_{jj}, $\Delta\eta_{jj}$, $\Delta\eta_{j\tau}^{\text{min}}$, $\Delta\eta_{j\tau}^{\text{max}}$, $p_{\text{T}}^{\text{tot}}$ and $m_{\text{T}}^{\ell\nu}$). These variables are optimized in parallel, independently of each other. The cut values obtained do not depend on the selection cuts on the other discriminating variables. The results of this optimization procedure are shown in Section 6.6.

However, this method has two drawbacks:
First, it does not take into account the effect of the other cuts on the signal-to-background ratio. Even if the shape of the distribution of a discriminating variable is not modified by a cut on another variable, the numbers of signal and background events will change. Depending on the cut optimization criterion, this could result in a different optimal cut value. For example, if a cut on another variable increased the S/B ratio such that $S > B$, the ratio $S/\sqrt{S+B}$ approaches $S/\sqrt{S}$ which has its maximum at the maximum number of signal events S. Thus, the cut on this variable would get very soft even if the ratio S/B would be increased by a harder cut. Therefore, the optimal cut value on a variable depends on the cuts on the other variables. If the cuts are optimized independently, the cut values usually become too hard and the number of signal and background events is reduced too much compared to the global optimum. It may be useful to omit cuts in order not to loose too many signal events, for example, using only the variables with high separation power. The result of this approach is also shown in Section 6.6.

Second, not only the absolute number of events but also the shapes of distributions of discriminating variables may change if a cut on a correlated variable is applied. Obviously, this will result in a different optimal cut value. As an example, Figure 6.4 shows the distribution of the η-separation of the two tagging jets $\Delta\eta_{jj}$, before and after a cut on the invariant jet pair mass m_{jj} has been applied.

Since there is a correlation between the two variables, the shapes of the distributions depend on cuts which have been applied on them. Figure 6.5 shows the optimization of the cut on $\Delta\eta_{jj}$ with and without the prior cut on the invariant jet pair mass m_{jj}. The resulting optimal cut values of $\Delta\eta_{jj} > 4.7$ (no m_{jj} cut applied) and $\Delta\eta_{jj} > 3.8$ (with $m_{jj} > 700\,\text{GeV}$ applied) differ considerably.

Notice that with the $m_{jj} > 700\,\text{GeV}$ cut applied the optimal cut in $\Delta\eta_{jj}$ is lower than without the cut on m_{jj}. Looking at the normalized $\Delta\eta_{jj}$ distributions (see Figure 6.4) one would expect the contrary. But, in this case, the effect of the reduced number of events exceeds the effect of the shape change of the $\Delta\eta_{jj}$ distribution.

For these reasons, the parallel cut optimization is quite unstable and the result depends on the preselection cuts and the set of cuts optimized. This can be improved by iterating the optimization procedure, described in the following.

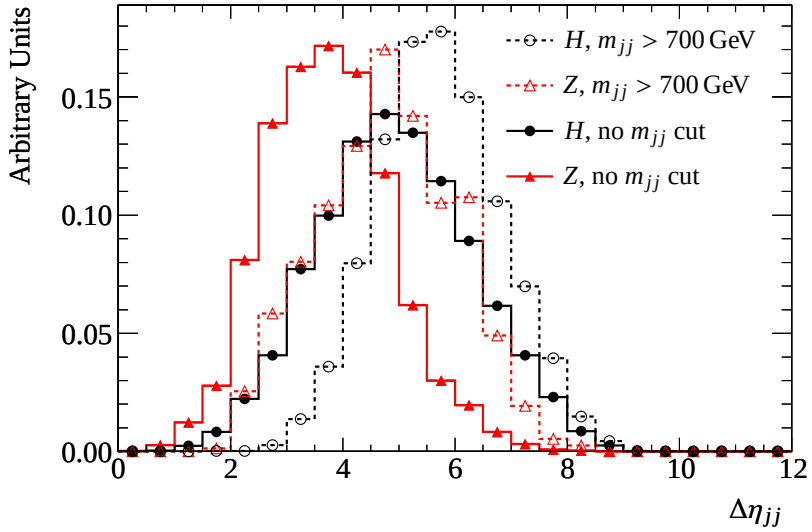


Figure 6.4: Distributions of the pseudorapidity separation $\Delta\eta_{jj}$ of the tagging jets for signal and Z + jets background before (full lines) and after (dashed lines) the cut on the invariant jet pair mass $m_{jj} > 700\,\text{GeV}$.

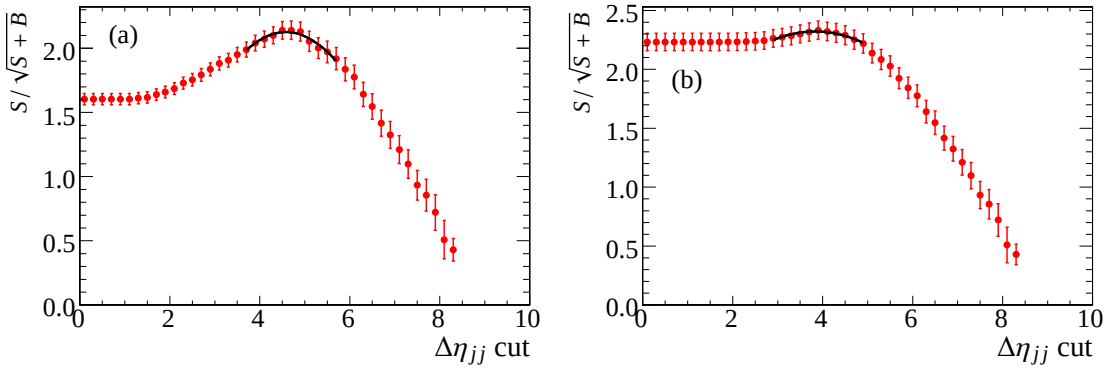


Figure 6.5: Comparison of the cut optimization on $\Delta\eta_{jj}$ (a) without other cuts applied before and (b) with a prior cut on $m_{jj} > 700\,\text{GeV}$ applied. While the optimal cut value for (a) is 4.7, it is 3.8 for (b).

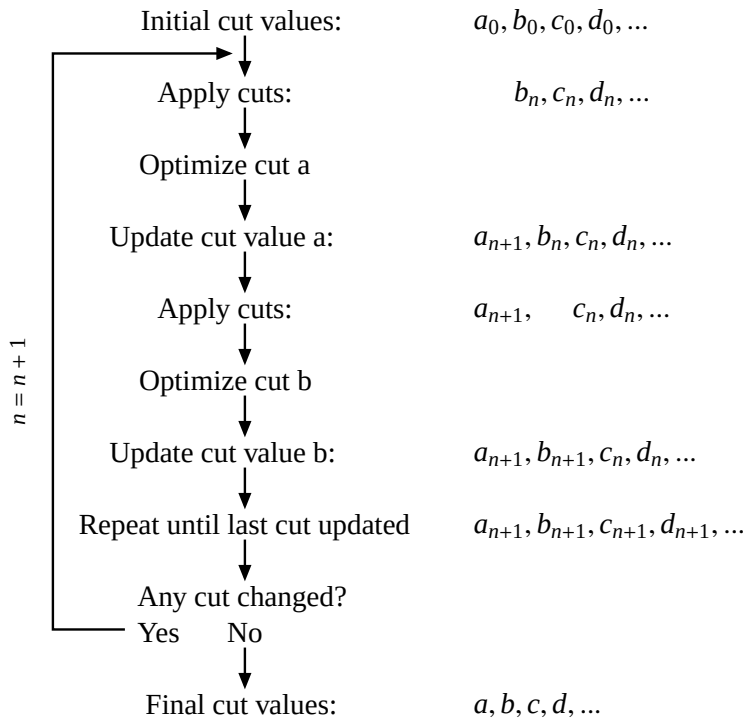


Figure 6.6: Sketch of the iterative cut optimization procedure.

6.4.2 Iterative Cut Optimization

In the iterative cut optimization approach, one starts with an arbitrary set of cut values. The cut on one variable is optimized after applying all other cuts. The set of cuts is then updated with the new optimized cut. Next, the cut on a second variable is optimized after applying the updated set of cuts. This procedure is repeated for every variable. Then the whole procedure is repeated. The iteration is stopped when the cut values do not significantly change anymore from one iteration to the next. Figure 6.6 illustrates this procedure.

By this approach, the correlations between the discriminating variables are taken into account and the final cut values depend neither on the initial cut values, nor on the order in which the cuts are applied. In addition, including weak variables does not change the optimized cut values since the cuts on weak variables usually become very soft.

Nevertheless, the following problems could occur:

- diverging cut values when starting the optimization with weakly discriminating variables,
- oscillations between several cut values without reaching a stable maximum,
- selection of a local maximum of the optimization function.

The results of both cut optimization methods are compared and summarized in Section 6.6.

6.5 Cut Efficiency Factorization for the $t\bar{t}$ Background

Due to the limited number of simulated events, even in case of the $t\bar{t}$ dataset of the fast detector simulation only very few $t\bar{t}$ events survive all analysis selection criteria resulting in a large statistical uncertainty on the prediction of the $t\bar{t}$ background.
In order to estimate the $t\bar{t}$ contribution more reliably, the selection efficiency $\varepsilon_{\mathrm{mw}}$ of the mass window cut has been determined with high statistics at an earlier stage of the analysis:

$$\varepsilon_{\mathrm{mw}} = \frac{n_i(\text{mass window})}{n_i(\text{whole mass range})}, \tag{6.1}$$

where n_i(mass window) is the number of events in a mass window around the Higgs peak and n_i(whole mass range) is the total number of events remaining after the i-th cut.
Instead of applying the mass window cut at the end of the analysis the efficiency $\varepsilon_{\mathrm{mw}}$ can be used to calculate the number of events in the mass window,

$$n_f(\text{mass window}) = \varepsilon_{\mathrm{mw}} \cdot n_f(\text{whole mass range}), \tag{6.2}$$

where n_f indicates the final number of events surviving all analysis cuts. With this mass window factorization no events are rejected by the mass window cut and the statistical uncertainty is decreased.
It is assumed here that the efficiency factor $\varepsilon_{\mathrm{mw}}$ is independent of the subsequent cuts. However, the shape of the invariant $\tau\tau$-mass distribution (and thus the value of $\varepsilon_{\mathrm{mw}}$) *does* change with some of the selection cuts, for example by applying a cut on $\Delta\phi_{\tau\tau}$ as shown in Figure 6.7. Nevertheless, up to $m_{\tau\tau} \lesssim 200\,\mathrm{GeV}$ the $m_{\tau\tau}$-shape is rather independent of the analysis cuts. Thus, $\varepsilon_{\mathrm{mw}}$ calculated with respect to the number of $t\bar{t}$ events in a fixed subrange between 0 and 200 GeV,

$$\varepsilon_{\mathrm{mw}} = \frac{n_i(\text{mass window})}{n_i(0 - 200\,\mathrm{GeV})} \tag{6.3}$$

can be assumed to be independent of the analysis cuts. Figure 6.8 shows the resulting value of $\varepsilon_{\mathrm{mw}}$ calculated in a mass window of $105 - 135\,\mathrm{GeV}$ after applying one of the analysis cuts. Within the statistical uncertainty, the mass window factorization for the mass window $m_{\tau\tau} = 105 - 135\,\mathrm{GeV}$ results in $\varepsilon_{\mathrm{mw}} \sim 20\,\%$, independent of all other analysis cuts. The described factorization increases the $t\bar{t}$ statistics at the end of the analysis by a factor of 5 and thus allows for a more reliable prediction of the $t\bar{t}$ background contribution.
Since the mass window factorization is also applied during the cut optimization procedure (see Section 6), the influence of this factorization on the shape of the distributions of discriminating variables has been tested in addition. Figures 6.9 and 6.10 show the distributions of all discriminating variables before and after the mass window factorization for a window of $m_{\tau\tau} = 90 - 150\,\mathrm{GeV}$. The factorization has only a minor influence on the $\Delta\phi_{\tau\tau}$ distribution and there is no significant impact on the shape of all other distributions. Thus, it can be assumed that the mass window factorization does not introduce any large bias on the final analysis result.

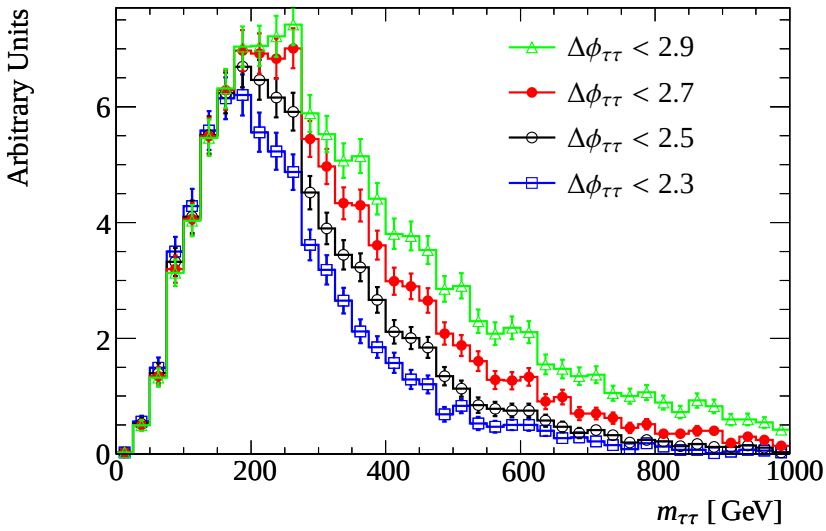


Figure 6.7: Invariant $\tau\tau$-mass distribution of the $t\bar{t}$ process, shown for several different cut values on the ϕ separation $\Delta\phi_{\tau\tau}$ between the visible τ-decay products. Compared to the other discriminating variables, the cut on $\Delta\phi_{\tau\tau}$ has the largest impact on the $m_{\tau\tau}$-shape. The distributions are normalized to have the same integral in the range from 0 – 200 GeV. In this range, the shapes of the different spectra agree within the statistical uncertainties, while the shape above 200 GeV differs significantly.

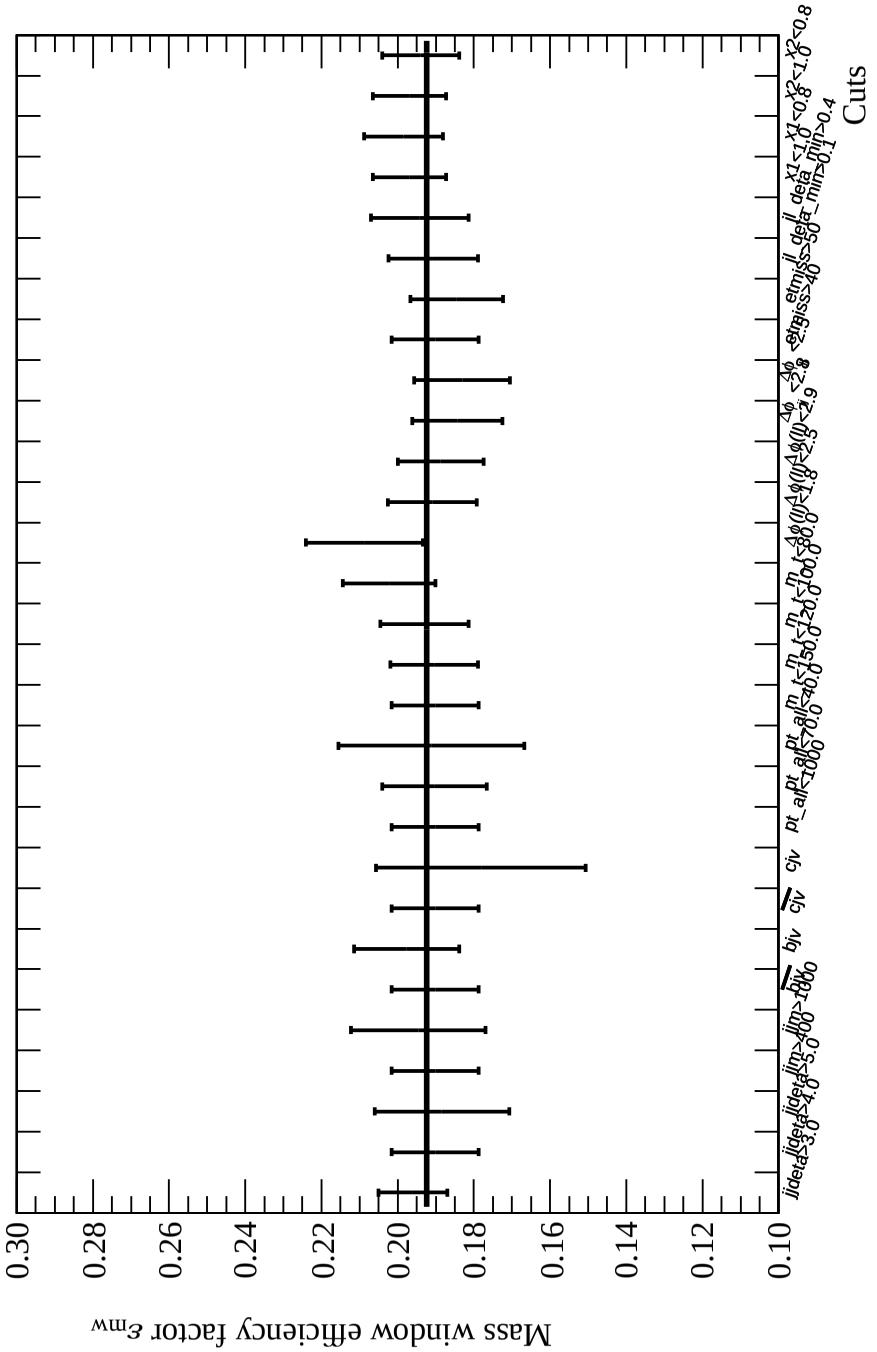


Figure 6.8: Efficiency factor $\varepsilon_{\mathrm{mw}}$ obtained from the mass window $\varepsilon_{\mathrm{mw}} = \frac{n(105-135\,\mathrm{GeV})}{n(0-200\,\mathrm{GeV})}$ as a function of different selection cuts applied before determining $\varepsilon_{\mathrm{mw}}$.

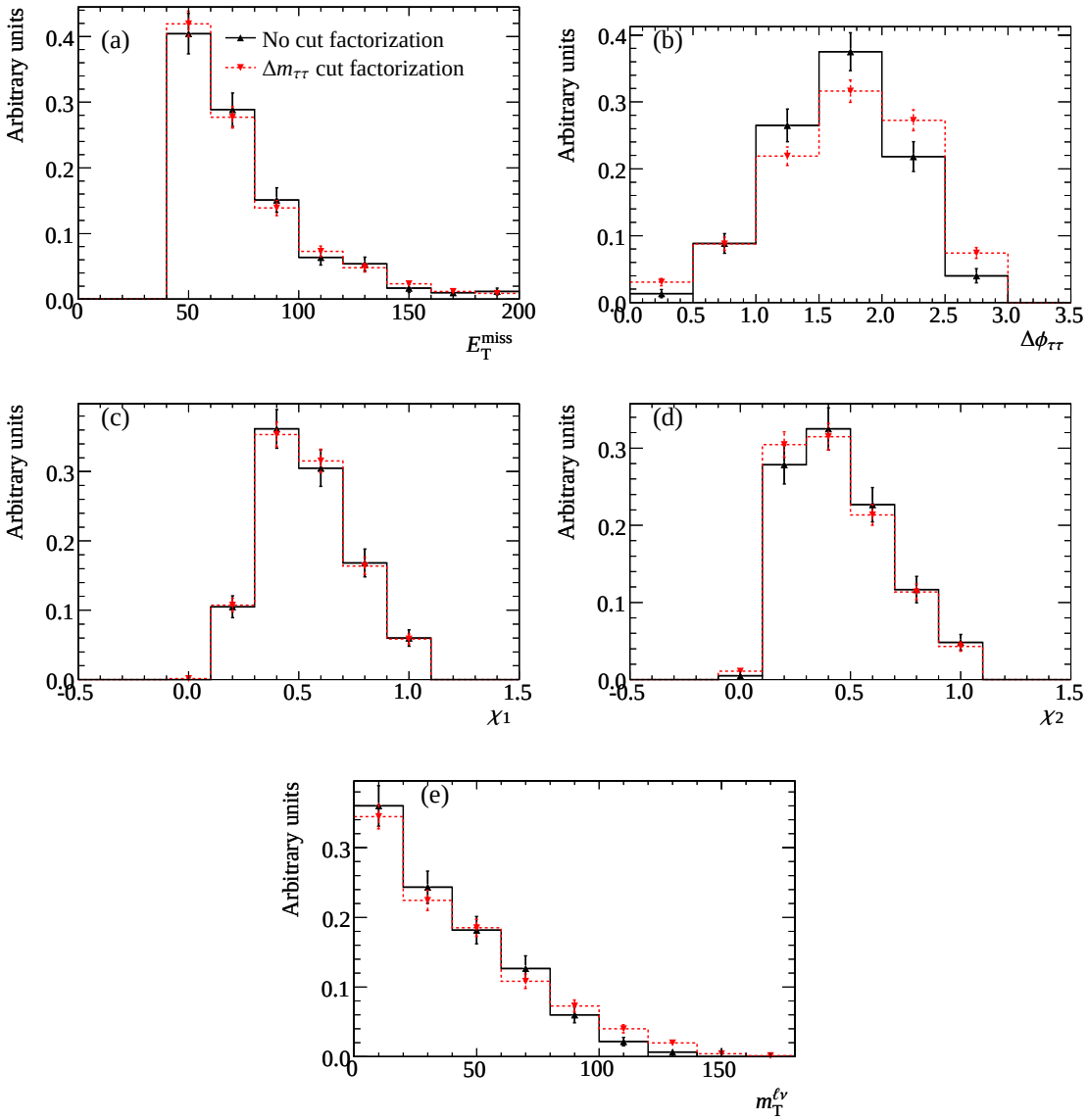


Figure 6.9: Distributions of different discriminating variables, before (full line) and after (dashed line) applying the mass window factorization, shown for the $t\bar{t}$ process. In (a) the missing transverse energy $E_{\mathrm{T}}^{\mathrm{miss}}$, in (b) the ϕ separation between the two τ-decay products $\Delta\phi_{\tau\tau}$, in (c) and (d) the visible momentum fractions χ_1 and χ_2 and in (e) the transverse mass $m_{\mathrm{T}}^{\ell\nu}$ distribution is shown.

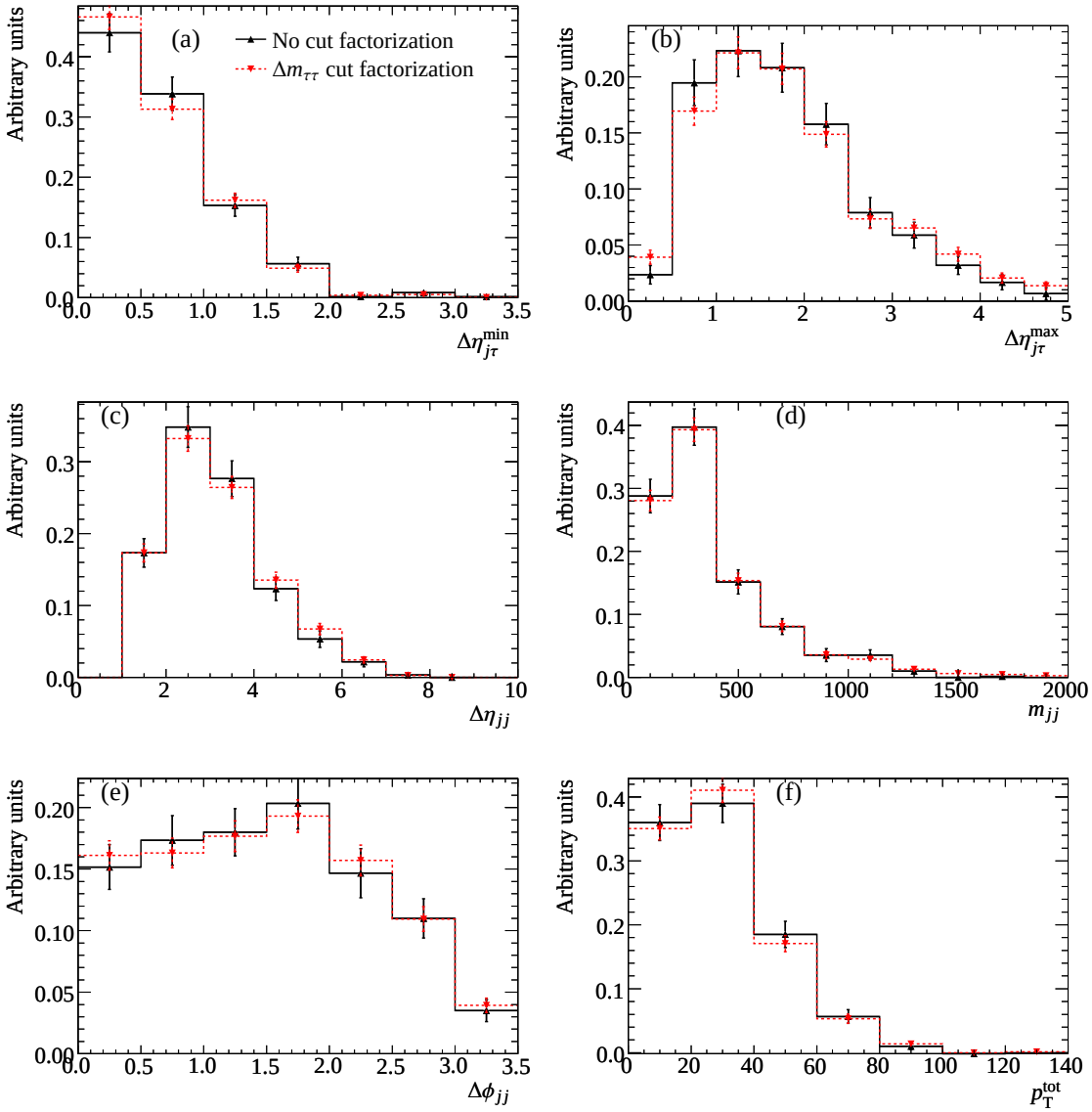


Figure 6.10: Distributions of different discriminating variables, before (full line) and after (dashed line) applying the mass window factorization, shown for the $t\bar{t}$ process. In (a) and (b) the η separation of tagging jet and τ-decay product $\Delta\eta_{j\tau}^{\min}$ and $\Delta\eta_{j\tau}^{\max}$, in (c) the η separation of the tagging jets $\Delta\eta_{jj}$, in (d) the invariant mass of the tagging jets m_{jj}, in (e) the ϕ separation of the tagging jets $\Delta\phi_{jj}$ and in (f) the p_{T} balance $p_{\mathrm{T}}^{\mathrm{tot}}$ distribution is shown.

Table 6.1: Cut values obtained with different cut optimization strategies (see text), shown for the leptonic decay channel. Brackets indicate that the cut has not been optimized, but fixed (see text). For comparison, the last column shows the cut values of the baseline analysis.

Discriminating variable		Optimization scenario				Baseline analysis
		parallel		iterative		
		A	**B**	**C**	**D**	
$E_{\mathrm{T}}^{\mathrm{miss}}$	$>$	(40 GeV)	(40 GeV)	(40 GeV)	40 GeV	40 GeV
$\Delta\phi_{\tau\tau}$	$<$	(2.9)	(2.9)	(2.9)	2.5	2.7
χ_1		(0.0 – 1.0)	(0.0 – 1.0)	(0.0 – 1.0)	0.0 – 1.0	0.0 – 0.75
χ_2		(0.0 – 1.0)	(0.0 – 1.0)	(0.0 – 1.0)	0.0 – 1.0	0.0 – 0.75
$m_{\mathrm{T}}^{\ell\nu}$	$<$	100 GeV	—	—	100 GeV	—
$\Delta\eta_{j\tau}^{\mathrm{min}}$	$>$	1.5	1.5	0.7	0.6	—
$\Delta\eta_{j\tau}^{\mathrm{max}}$	$>$	2.0	—	—	1.5	—
m_{jj}	$>$	800 GeV	800 GeV	700 GeV	600 GeV	700 GeV
$\Delta\eta_{jj}$	$>$	4.7	4.7	3.6	3.5	4.4
$\Delta\phi_{jj}$	$<$	—	—	—	2.2	—
$p_{\mathrm{T}}^{\mathrm{tot}}$	$<$	20 GeV	—	—	30 GeV	—

6.6 Results of the Cut-Based Analysis

6.6.1 Results for the Leptonic Decay Channel

Table 6.1 gives an overview of the cut values obtained for different sets of discriminating variables for the parallel as well as for the iterative cut optimization procedure. In Table 6.2 the corresponding cross-sections after applying all analysis cuts and the signal significances are shown for an integrated luminosity of 30 fb^{-1}.

In the first scenario (**A**), the parallel optimization was performed using all variables shown in Figure 6.2. These are the variables allowing for an improvement of $S/\sqrt{S+B}$, if one performs a cut on only this variable without applying a cut on the other variables. The set of cuts obtained by this scenario leads to a $S/\sqrt{S+B}$ ratio of 2.5 ± 0.11 with the uncertainty being due to the statistical uncertainty of the simulated data.

The parallel optimization method has also been used in the second scenario (**B**). However, only the three most important variables (m_{jj}, $\Delta\eta_{jj}$ and $\Delta\eta_{j\tau}^{\mathrm{min}}$) have been included. Compared to scenario **A**, this results in a higher efficiency for the signal as well as for the background, slightly increasing the $S/\sqrt{S+B}$ ratio to 2.6 ± 0.12.

In the third scenario (**C**), the same three variables as in scenario **B** have been used in the iterative cut optimization resulting in a higher $S/\sqrt{S+B}$ ratio of 2.8 ± 0.11 since correlations are taken into account.

It is worth mentioning that the iterative procedure is very stable: The final cut values are independent of the ordering of the cuts during the optimization and of the initial cut values, indicating that the cut optimization finds the global maximum of $S/\sqrt{S+B}$. The optimization procedure

Table 6.2: Results of the different cut optimization strategies. The first lines show the cross-sections times efficiency [fb] of the signal and the dominant background processes and the signal-to-background ratio after applying all cuts except for the mass window cut. The last line shows the expected signal significance $S/\sqrt{S+B}$ within the mass window of $m_{\tau\tau} = m_H \pm 15$ GeV for an integrated luminosity of 30 fb^{-1}. For comparison, the last column shows the results obtained for the baseline analysis.

	Cut Optimization Scenario				Baseline
	A	**B**	**C**	**D**	Analysis[2]
Cross-section · efficiency [fb]					
$H \to \tau\tau$ (120 GeV)	0.32 ± 0.02	0.43 ± 0.02	0.69 ± 0.02	0.58 ± 0.02	0.54 ± 0.02
$Z \to \tau\tau$ QCD	0.17 ± 0.01	0.42 ± 0.02	1.3 ± 0.04	1.0 ± 0.03	1.3 ± 0.04
$Z \to \tau\tau$ EW	0.18 ± 0.03	0.36 ± 0.04	0.60 ± 0.05	0.23 ± 0.03	0.43 ± 0.04
$t\bar{t}$	0.32 ± 0.08	1.8 ± 0.18	5.9 ± 0.32	1.5 ± 0.16	4.7 ± 0.30
B total	0.67 ± 0.09	2.6 ± 0.19	7.8 ± 0.33	2.7 ± 0.17	6.4 ± 0.29
S/B	0.47 ± 0.064	0.17 ± 0.014	0.09 ± 0.005	0.21 ± 0.015	0.09 ± 0.005
$S/\sqrt{S+B}$ for a ± 15 GeV mass window around $m_H = 120$ GeV and 30 fb^{-1}:					
	2.5 ± 0.11	2.6 ± 0.12	2.8 ± 0.11	3.0 ± 0.11	2.5 ± 0.10

converges quickly within four to six iterations and no oscillation or divergence of the cut values occurs.

In the fourth scenario (**D**), all available variables have been taken into account in the iterative cut optimization.

The cuts on variables with only small impact on the $S/\sqrt{S+B}$ ratio have been set to very loose values rejecting no signal events. In this way, these cuts still can suppress background processes which have not been simulated. With a signal significance of 3.0 ± 0.11, the fourth scenario performs best.

Tables 6.3 and A.1 (Appendix A) show the cross-sections times efficiency for the background processes and several signal mass points after each analysis cut from the fourth cut optimization scenario **D**. Figure 6.11 illustrates this evolution of the cross-sections for the most important processes. As shown, jet-related variables ($\eta_{j1} \times \eta_{j2} < 0$, $\Delta\eta_{jj}$, m_{jj}, central jet veto) provide for a strong rejection of most of the background processes while leaving the signal almost unaffected. This confirms the advantage of the vector-boson fusion channel for the Higgs boson search with respect to the gluon fusion production. The electroweak $Z \to \tau\tau$ events are not as much suppressed by the cuts on the tagging jets as the other backgrounds since they are partly produced by the same mechanism as the signal (see Figure 5.2a) and the Z bosons decay into the same final state. Fortunately, the cross-section for this process is much lower than for the QCD $Z \to \tau\tau$ production. No $Z \to \ell\ell$ + jets and the W + jets background events survive after applying all analysis cuts. About 10-100 times more Monte Carlo events would have to be produced in order to reliably estimate these background contributions after the signal selection. However, the cuts on $E_{\mathrm{T}}^{\mathrm{miss}}$ and

[2] Although the same selection cuts have been applied, the results of the baseline analysis slightly differ from the ones presented in [49] since some corrections on the fast detector simulation and the τ-jet energy contribution to the missing transverse energy have not been applied here.

Table 6.3: Evolution of the cross-section times efficiency [fb] with the applied selection cuts for the signal and background processes in the leptonic decay channel. The uncertainties shown are statistical only. In case of the Z + jets and the W + jets background a "—" indicates that no event remained after all preceding cuts (see text).

	$H \to \tau\tau$	$Z \to \tau\tau$ +jets		$t\bar{t}$		$Z \to \ell\ell$ + jets	W + jets
	120 GeV	QCD	EW	AtlFast	FullSim		
Cross-section	22	1 410	214	833 000	461 000	34 300	347 000
Cuts on τ-decay products:							
Trigger	11.6 ± 0.05	657 ± 2	69 ± 1.0	$242\,000 \pm 610$	$183\,000 \pm 540$	$26\,900 \pm 42$	$165\,000 \pm 320$
N($e + \mu$)=2	5.48 ± 0.04	373 ± 2	7.7 ± 0.2	$30\,900 \pm 220$	$23\,200 \pm 190$	$17\,400 \pm 34$	573 ± 18
Lepton charge	5.46 ± 0.04	371 ± 2	7.5 ± 0.2	$30\,200 \pm 220$	$22\,300 \pm 190$	$17\,400 \pm 34$	363 ± 14
$E_{\mathrm{T}}^{\mathrm{miss}} > 40$ GeV	3.17 ± 0.03	147 ± 1.0	4.6 ± 0.1	$23\,500 \pm 190$	$17\,300 \pm 170$	597 ± 7	195 ± 10
$\Delta\phi_{\tau\tau} < 2.5$	2.85 ± 0.03	137 ± 0.5	4.4 ± 0.1	$16\,500 \pm 160$	$11\,900 \pm 140$	507 ± 6	139 ± 9
$\chi_{1,2} = 0.0 - 1.0$	2.57 ± 0.03	113 ± 0.5	3.6 ± 0.1	$2\,620 \pm 64$	$1\,910 \pm 56$	71 ± 2.4	23 ± 3.3
$m_{\mathrm{T}}^{\ell\nu} < 100$ GeV	2.57 ± 0.03	113 ± 0.4	3.5 ± 0.1	$2\,140 \pm 59$	$1\,570 \pm 51$	68 ± 2.3	19 ± 3
Cuts on the tagging jets:							
N(jets)≥ 2	2.10 ± 0.02	100 ± 0.4	3.4 ± 0.1	$2\,030 \pm 57$	$1\,490 \pm 49$	65 ± 2.3	17 ± 2.8
$\eta_{j1} \times \eta_{j2} < 0$	1.74 ± 0.02	57.5 ± 0.3	2.4 ± 0.1	960 ± 4.1	670 ± 25	36 ± 1.7	9 ± 2.1
$\Delta\eta_{j\tau} > 0$	1.58 ± 0.02	29.7 ± 0.2	1.9 ± 0.1	366 ± 2.6	250 ± 15	13 ± 1.0	2.0 ± 1.1
$\Delta\eta_{jj} > 3.5$	1.32 ± 0.02	13.9 ± 0.1	1.40 ± 0.08	124 ± 1.5	83 ± 8.8	3.6 ± 0.5	0.76 ± 0.62
$m_{jj} > 600$ GeV	1.08 ± 0.02	7.4 ± 0.1	1.20 ± 0.07	84 ± 1.2	64 ± 7.6	2.5 ± 0.4	0.16 ± 0.16
$\Delta\phi_{jj} < 2.2$	0.89 ± 0.03	4.67 ± 0.08	0.62 ± 0.05	58 ± 1.0	48 ± 6.4	1.5 ± 0.3	0.16 ± 0.16
b-jet veto	0.79 ± 0.03	4.02 ± 0.07	0.49 ± 0.05	44.5 ± 0.9	27 ± 4.6	1.1 ± 0.3	0.16 ± 0.16
Cuts on the event topology:							
$\Delta\eta_{j\tau}^{\mathrm{min}} > 0.6$	0.74 ± 0.03	2.78 ± 0.06	0.42 ± 0.04	32.9 ± 0.8	22.0 ± 4.1	0.8 ± 0.2	0.16 ± 0.16
$\Delta\eta_{j\tau}^{\mathrm{max}} > 1.5$	0.73 ± 0.03	2.77 ± 0.06	0.41 ± 0.04	31.2 ± 0.8	22.0 ± 4.1	0.8 ± 0.2	0.16 ± 0.16
Central jet veto	0.63 ± 0.02	1.28 ± 0.04	0.30 ± 0.04	2.9 ± 0.2	2.7 ± 1.7	0.40 ± 0.2	—
$p_{\mathrm{T}}^{\mathrm{tot}} < 30$ GeV	0.58 ± 0.02	1.02 ± 0.03	0.23 ± 0.03	1.5 ± 0.2	1.4 ± 1.0	0.07 ± 0.05	—
Mass window cut:							
$m_{\tau\tau}$ =105-135 GeV	0.49 ± 0.02	0.14 ± 0.01	0.01 ± 0.008	0.13 ± 0.02[a]	0.68 ± 0.7	—	—

[a] The mass window cut efficiency factorization has been applied here. Without factorization the cross-section times efficiency is 0.14 ± 0.05.

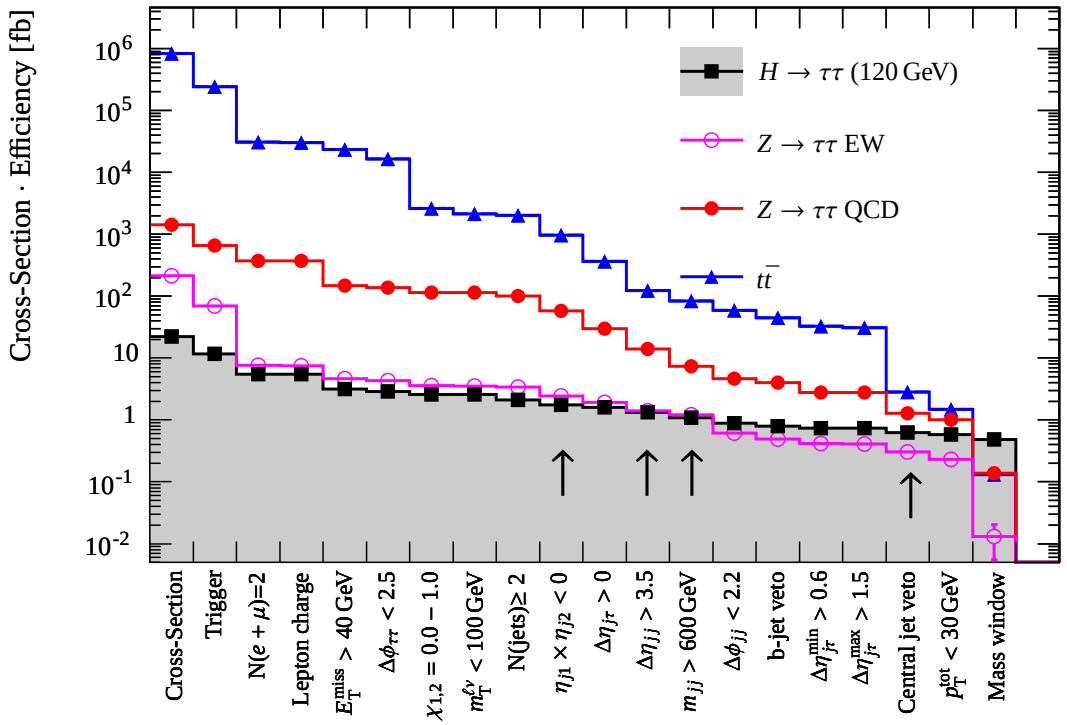


Figure 6.11: Evolution of the cross-sections times efficiency [fb] with the applied selection cuts in the leptonic decay channel for the signal and the most important background processes. The arrows indicate some of the most important cuts related to the vector-boson fusion characteristics (see text).

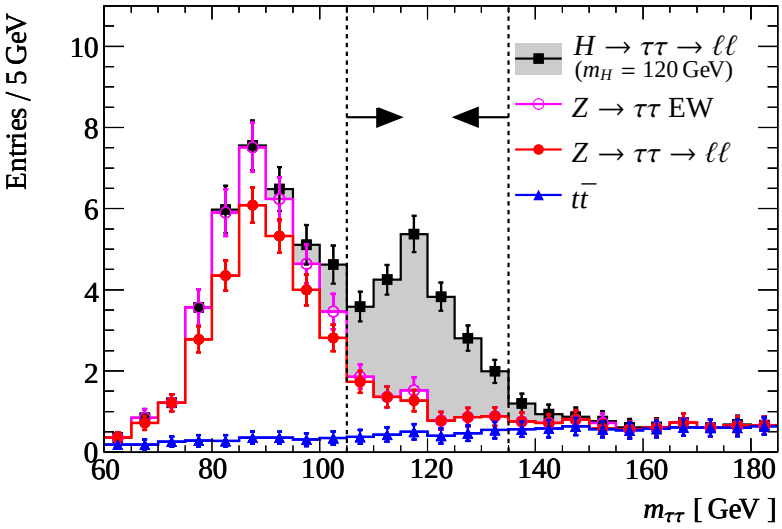


Figure 6.12: Invariant mass distribution $m_{\tau\tau}$ after all selection cuts for the leptonic decay mode and an integrated luminosity of 30 fb^{-1}. The dashed lines indicate the mass window cut.

$\chi_{1,2}$ provide a suppression of the $Z \to \ell\ell$ events, which is about 100 times larger than for $Z \to \tau\tau$. Therefore, the $Z \to \ell\ell$ background is assumed to be negligible. In the CSC analysis [49], the background contribution from W + jets in the leptonic channel has been estimated to be at least a factor of ten smaller than the contribution from $Z \to \tau\tau$ or $t\bar{t}$. Thus, the W + jets background has been neglected in the $\ell\ell$ channel in the remainder of this study.
Figure 6.12 shows the invariant mass distribution $m_{\tau\tau}$ after all cuts except for the mass window cut. The histograms are normalized to an integrated luminosity of 30 fb^{-1}. The error bars indicate the statistical uncertainties of the simulated data.

6.6.2 Results for the Semileptonic Decay Channel

For the semileptonic decay channel, the cuts have been optimized in the same way as for the leptonic channel. As it results in the best signal significance, the iterative cut optimization using all discriminating variables (scenario **D** in the leptonic mode) has been applied.
The resulting cut values are summarized in Table 6.4. For comparison, also the values for the leptonic decay mode are shown in the table. It can be seen, that the cut values are very similar for the two decay channels. The most important differences, the changed cut values on $E_{\mathrm{T}}^{\mathrm{miss}}$ and $m_{\mathrm{T}}^{\ell\nu}$, are due to the different preselection in the semileptonic decay channel, see page 87.
Also the resulting total cross-sections times efficiency, shown in Table 6.5, are quite similar to the ones for the leptonic decay mode. For the signal and the $Z \to \tau\tau$ background, the accepted cross-section is almost the same as in the leptonic mode. The higher initial cross-sections, due to the higher branching ratio of $\tau \to h\nu$ compared to $\tau \to \ell\nu\nu$, is roughly compensated by the lower τ-jet reconstruction and identification efficiency compared to the lepton identification efficiency.

Table 6.4: Cut values obtained with the iterative optimization method for the semileptonic decay channel. For comparison, the cut values for the leptonic decay channel (see Table 6.1, scenario **D**) and the cut values applied in the baseline analysis are shown.

		Optimized cuts		Baseline analysis
Decay channel		$\ell\ell$	ℓh	ℓh
$E_{\mathrm{T}}^{\mathrm{miss}}$	$>$	40 GeV	30 GeV	40 GeV
$\Delta\phi_{\tau\tau}$	$<$	2.5	2.5	2.7
χ_ℓ		0.0 – 1.0	0.0 – 1.0	0.0 – 0.75
χ_h		—	0.0 – 1.0	0.0 – 1.0
$m_{\mathrm{T}}^{\ell\nu}$	$<$	100 GeV	30 GeV	30 GeV
$\Delta\eta_{j\tau}^{\mathrm{min}}$	$>$	0.6	0.8	—
$\Delta\eta_{j\tau}^{\mathrm{max}}$	$>$	1.5	1.5	—
m_{jj}	$>$	600 GeV	600 GeV	700 GeV
$\Delta\eta_{jj}$	$>$	3.5	3.6	4.4
$\Delta\phi_{jj}$	$<$	2.2	2.7	—
$p_{\mathrm{T}}^{\mathrm{tot}}$	$<$	30 GeV	40 GeV	—

Table 6.5: Results for the semileptonic channel: Cross-sections times efficiency [fb] of the signal and the dominant background processes, signal-to-background ratio after applying all cuts except for the mass window cut and expected signal significance $S/\sqrt{S+B}$ within the mass window of $m_{\tau\tau} = m_H \pm 15$ GeV for an integrated luminosity of 30 fb^{-1}. For comparison, the results for the leptonic decay channel (see Table 6.2, scenario **D**) and for the baseline analysis are shown.

	Optimized cuts		Baseline analysis
Decay channel	$\ell\ell$	ℓh	ℓh
Cross-section · efficiency [fb]			
$H \to \tau\tau$ (120 GeV)	0.58 ± 0.02	0.67 ± 0.04	0.67 ± 0.04
$Z \to \tau\tau$ QCD	1.0 ± 0.03	1.0 ± 0.09	1.2 ± 0.10
$Z \to \tau\tau$ EW	0.23 ± 0.03	0.38 ± 0.04	0.46 ± 0.05
$t\bar{t}$	1.5 ± 0.16	0.14 ± 0.05	0.28 ± 0.07
B total	2.7 ± 0.17	1.5 ± 0.11	1.6 ± 0.10
S/B	0.21 ± 0.02	0.44 ± 0.04	0.35 ± 0.03
$S/\sqrt{S+B}$ for a ±15 GeV mass window around $m_H = 120$ GeV and 30 fb^{-1}:	3.0 ± 0.11	3.7 ± 0.15	3.6 ± 0.15

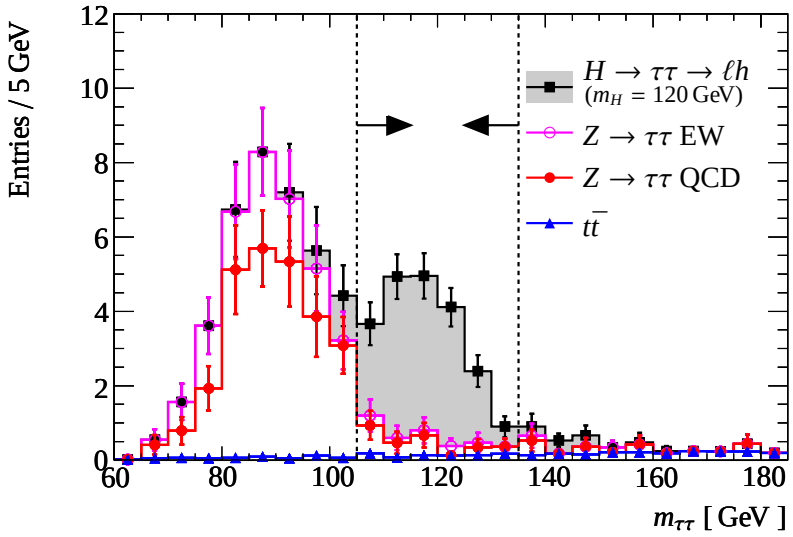


Figure 6.13: Invariant mass distribution $m_{\tau\tau}$ after all selection cuts for the semileptonic decay mode and an integrated luminosity of 30 fb^{-1}. The dashed lines indicate the mass window cut.

A difference is found for the $t\bar{t}$ process which has a lower cross-section times selection efficiency in the semileptonic decay mode. This can be explained by the lower branching ratio of $t\bar{t}$ into semileptonic compared to leptonic final states (see page 43).
Although the branching ratio of $t\bar{t}$ into events that do not contain a τ jet is much higher than into events containing a τ jet, the contribution from $t\bar{t}$ events in the semileptonic channel is rather low. This shows that the rate of misidentified τ jets is small compared to the number of events where a τ jet has been identified correctly. Obviously, the estimation of the $t\bar{t}$ contribution requires a good understanding of the τ jet reconstruction performance. The influence of the different τ jet performance in the fast and full detector simulation are discussed in Section 6.6.3
Figure 6.13 shows the invariant mass distribution $m_{\tau\tau}$ after all analysis cuts except for the cut on the mass window. The histograms are normalized to the number of events expected for an integrated luminosity of 30 fb^{-1}. The error bars indicate the statistical uncertainties of the simulated data.
Figure 6.14 shows the evolution of the cross-section times efficiency with the applied selection cuts for the signal and the dominant background processes. The detailed numerical information can be found in Table 6.6 and A.2 (Appendix).
Similar to the leptonic channel, no $Z \to \ell\ell$ + jets and W + jets events survive after applying all analysis cuts. In the semileptonic channel, $Z \to \ell\ell$ is even more suppressed than in the leptonic channel since it contains no real τ jets and, thus, has also been neglected as a background for the ℓh channel.
The invariant $m_{\tau\tau}$ distribution as well as the cross-section of the W + jets process after applying all analysis cuts in the ℓh channel is expected to be similar to the $t\bar{t}$ background [49]. In the cut-based and the multivariate analysis (Chapter 7) the contribution from W + jets has been neglected. An additional systematic uncertainty of 100 % due to this background in the ℓh channel has been

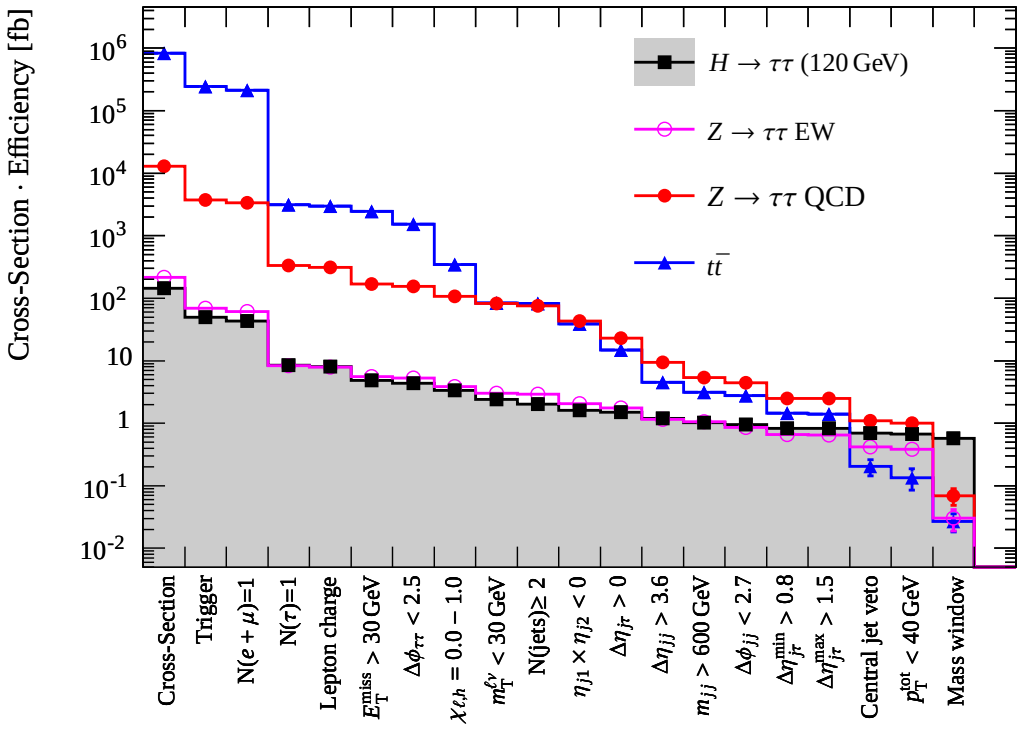


Figure 6.14: Evolution of the cross-sections times efficiency [fb] with the applied selection cuts in the semileptonic decay channel for the signal and the most important background processes.

Table 6.6: Evolution of the cross-sections times efficiency [fb] with the applied selection cuts for the signal and background processes in the semileptonic decay channel. The uncertainties shown are statistical only. In case of the $Z \to \ell\ell$ + jets and the W + jets background a "—" indicates that no event remained after all preceding cuts (see text). The negative cross-section for the $t\bar{t}$ process from full simulation at the end of the analysis is due to a single event with negative weight passing all the cuts.

	$H \to \tau\tau$	$Z \to \tau\tau$ + jets		$t\bar{t}$	$t\bar{t}$	$Z \to \ell\ell$ + jets	W + jets
	120 GeV	QCD	EW	AtlFast	FullSim		
Cross-section	145	12 900	214	833 000	461 000	34 300	347 000
Cuts on τ-decay products:							
Trigger	49.5 ± 0.2	3 740 ± 13	69 ± 0.55	243 000 ± 610	183 000 ± 540	26 900 ± 42	165 000 ± 320
N($e+\mu$)=1	43.4 ± 0.2	3 360 ± 12	61.4 ± 0.52	211 000 ± 570	16 000 ± 510	9 440 ± 25	164 000 ± 320
N(τ)=1	8.45 ± 0.1	337 ± 3.1	8.28 ± 0.19	3 160 ± 69	6 340 ± 100	1 000 ± 8	2 750 ± 41
Lepton charge	8.02 ± 0.1	311 ± 3.1	7.96 ± 0.19	2 960 ± 66	5 290 ± 92	877 ± 7.5	1 910 ± 34
$E_{\mathrm{T}}^{\mathrm{miss}} > 30$ GeV	4.87 ± 0.1	168 ± 1.4	5.59 ± 0.16	2 470 ± 61	4 290 ± 83	90.8 ± 2.7	1 310 ± 28
$\Delta\phi_{\tau\tau} < 2.5$	4.34 ± 0.1	155 ± 1.2	5.27 ± 0.15	1 530 ± 48	2 770 ± 68	76.5 ± 2.5	990 ± 24
$\chi_{\ell,h} = 0.0 - 1.0$	3.35 ± 0.06	107 ± 0.96	3.85 ± 0.13	344 ± 23	542 ± 30	6.65 ± 0.7	96 ± 7
$m_{\mathrm{T}}^{\ell\nu} < 30$ GeV	2.40 ± 0.05	83 ± 0.84	3.03 ± 0.11	84 ± 12	177 ± 16	3.35 ± 0.51	26 ± 3.5
Cuts on the tagging jets:							
N(jets)≥ 2	2.00 ± 0.05	74.9 ± 0.74	2.92 ± 0.11	83 ± 12	169 ± 16	3.13 ± 0.5	22 ± 3.1
$\eta_{j1} \times \eta_{j2} < 0$	1.62 ± 0.04	43 ± 0.56	2.07 ± 0.10	39 ± 0.83	90.9 ± 8.5	1.8 ± 0.4	12 ± 2.3
$\Delta\eta_{j\tau} > 0$	1.49 ± 0.04	22.9 ± 0.41	1.76 ± 0.09	14 ± 0.52	30.5 ± 5.1	0.47 ± 0.2	5.7 ± 1.7
$\Delta\eta_{jj} > 3.6$	1.20 ± 0.04	9.35 ± 0.27	1.15 ± 0.07	4.5 ± 0.29	3.39 ± 2.3	0.22 ± 0.15	2.9 ± 1.2
$m_{jj} > 600$ GeV	1.02 ± 0.03	5.41 ± 0.20	1.06 ± 0.07	3.1 ± 0.24	2.04 ± 2.0	0.13 ± 0.13	2.9 ± 1.2
$\Delta\phi_{jj} < 2.7$	0.95 ± 0.04	4.42 ± 0.18	0.85 ± 0.06	2.8 ± 0.23	1.36 ± 1.9	—	2.5 ± 1.1
Cuts on the event topology:							
$\Delta\eta_{j\tau}^{\mathrm{min}} > 0.8$	0.83 ± 0.04	2.48 ± 0.13	0.66 ± 0.05	1.5 ± 0.17	0 ± 1.7	—	2.3 ± 1.1
$\Delta\eta_{j\tau}^{\mathrm{max}} > 1.5$	0.83 ± 0.04	2.48 ± 0.13	0.65 ± 0.05	1.4 ± 0.17	0 ± 1.7	—	2.3 ± 1.1
Central jet veto	0.70 ± 0.04	1.1 ± 0.09	0.42 ± 0.04	0.20 ± 0.06	-0.68 ± 0.68	—	1.1 ± 0.8
$p_{\mathrm{T}}^{\mathrm{tot}} < 40$ GeV	0.67 ± 0.04	1 ± 0.09	0.38 ± 0.04	0.14 ± 0.05	-0.68 ± 0.68	—	0.6 ± 0.6
Mass window cut:							
$m_{\tau\tau}$ =105-135 GeV	0.57 ± 0.03	0.07 ± 0.02	0.03 ± 0.01	0.03 ± 0.01[a]	-0.68 ± 0.68	—	—

[a] The mass window cut efficiency factorization has been applied here. Without factorization the cross-section times efficiency is 0.014 ± 0.014.

introduced when calculating the signal significances with background estimation based on Monte Carlo in Section 8.2.4, assuming that the shape of the $m_{\tau\tau}$ distribution is the same for the W + jets and $t\bar{t}$ contribution.

6.6.3 Comparison of Detailed and Fast Simulation for the $t\bar{t}$ Background

Due to the high cross-section for $t\bar{t}$ production, it is not possible to simulate a sufficient number of $t\bar{t}$ events using the full detector simulation. Therefore, the data sample simulated with ATLFAST has been used for the described analysis.
In Section 5.2 the detector performance determined from the full and the ATLFAST simulation is described. Here, the impact of using ATLFAST on the prediction of the $t\bar{t}$ background in the presented analysis is discussed. The most relevant differences between the full and the fast simulation are:

- Electron and Muon Trigger
  Since no trigger simulation is performed in the fast simulation, a reconstructed lepton with minimum p_{T} is required to represent the trigger selection in the analysis of events from the fast simulation. An event is considered to be accepted by the trigger if an electron with $p_{\mathrm{T}}^{e} \geq 25\,\mathrm{GeV}$ and/or a muon with $p_{\mathrm{T}}^{\mu} \geq 20\,\mathrm{GeV}$ is present. No τ-jet trigger was used in the analysis (see Section 5.2.1).

  The last two columns of Table 6.7 show the efficiency of the trigger requirements *after* requiring two reconstructed τ-decay products with opposite charge. In the leptonic decay channel, about 10 % of the $t\bar{t}$ events are rejected by the trigger in the detailed simulation, while in the fast simulation only 3 % of the events are not selected by the trigger for all lepton combinations ($ee, e\mu, \mu\mu$). In the semileptonic case, the trigger efficiency after requiring one lepton and one τ jet is 69 % for the full simulation and 84 % for the fast detector simulation.
  Thus, the missing trigger simulation in ATLFAST leads to an overestimation of the $t\bar{t}$ background by ~ 8 % in the leptonic decay channel and by ~ 22 % in the semileptonic mode.

- Lepton reconstruction and identification efficiency
  Figures 5.4 and 5.5 show the reconstruction efficiencies for electrons and muons, both for the full and for the fast simulation. The overall electron efficiency of 59 % in the full simulation significantly differs from 76 % in the fast simulation. However, the muon reconstruction efficiencies agree very well: 81 % in the full and 80 % in the fast simulation respectively.

  In the first two columns of Table 6.7, the cross-sections times efficiency after requiring two reconstructed τ-lepton candidates with opposite charge are shown for the two simulations. The third column shows the ratio of the two. In the fourth column, the expected ratio is given, which is calculated by using the single-electron, single-muon and single-τ-jet reconstruction efficiencies and the rate of misidentified τ jets. The differences between the two simulations are consistent with the differences in the electron and muon reconstruction efficiencies.

Table 6.7: Differences in the $t\bar{t}$ event selection between full and fast simulation for different decay modes. The first two columns show the cross-section times efficiency obtained after requiring two reconstructed τ-decay candidates with opposite charge. The third column shows the ratio between those cross-sections. In the fourth column, the corresponding expected ratios are given, which are calculated using the known electron, muon, τ-jet reconstruction efficiencies as well as the ratio of misidentified τ jets. The last two columns show the trigger efficiency for the two detector simulations after requiring two oppositely charged τ-decay products.

	Cross-section [fb]		Ratio	Expected	Trigger efficiency [%]	
$\tau\tau$-channel	Full sim.	Fast sim.	Full/Fast	Full/Fast	Full sim.	Fast sim.
ee-channel	$4\,610 \pm 90$	$7\,600 \pm 110$	0.61 ± 0.02	0.59	0.92 ± 0.03	0.97 ± 0.02
$e\mu$-channel	$12\,200 \pm 140$	$15\,600 \pm 160$	0.78 ± 0.01	0.78	0.91 ± 0.02	0.97 ± 0.01
$\mu\mu$-channel	$7\,850 \pm 110$	$7\,940 \pm 110$	0.99 ± 0.02	1.03	0.89 ± 0.02	0.98 ± 0.02
all $\ell\ell$	$24\,700 \pm 200$	$31\,100 \pm 220$	0.79 ± 0.01	0.80	0.90 ± 0.01	0.97 ± 0.01
$e\tau$-channel	$3\,300 \pm 70$	$1\,730 \pm 50$	1.9 ± 0.1	2.0	0.70 ± 0.02	0.82 ± 0.04
$\mu\tau$-channel	$4\,350 \pm 90$	$1\,790 \pm 50$	2.4 ± 0.1	2.7	0.68 ± 0.02	0.86 ± 0.04
all ℓh	$7\,650 \pm 110$	$3\,530 \pm 70$	2.2 ± 0.1	2.3	0.69 ± 0.02	0.84 ± 0.03

- Hadronic τ-decay reconstruction and identification efficiency
  The difference in the simulated τ-jet reconstruction performance is shown in Figure 5.7. In the full detector simulation the overall τ-jet reconstruction efficiency is 34 % and the misidentification rate of τ jets is 56 %. In the fast simulation, the τ-jet efficiency of 36 % is similar to the full simulation, but the τ-jet misidentification rate of 10 % is much smaller than in the full simulation. The difference in the τ-jet misidentification rate is even more pronounced at a later stage in the analysis. After requiring one τ jet and one electron or muon with opposite charge, the fraction of events with a misidentified τ jet is 66 % in the detailed simulation while it is only 13 % in the fast simulation.

  Table 6.7 shows the cross-sections times efficiency for the semileptonic decay modes. Also for the semileptonic decay modes, the difference in the cross-sections in the full and the fast simulation is consistent with the different electron, muon and τ-jet reconstruction efficiencies and τ-jet misidentification rates.

- b-Jet veto efficiency
  Also the b-identification efficiency differs significantly between the two detector simulations. As shown in Figure 5.8, the b-jet identification efficiency is higher in the full simulation and, consequently, the b-jet veto rejects $t\bar{t}$ events more efficiently in the full simulation.

  But, since the cuts on the jet kinematics (for example $\Delta\eta_{jj} > 3.5$) shift the jet η-distribution to higher $|\eta|$ values and b-jet identification is only possible in the central detector region, the b-jet veto rejects less events when it is applied *after* the jet kinematical cuts and the difference between fast and full simulation is significantly reduced.

Effects on the Leptonic Analysis

The differences between full and fast simulation affecting the leptonic analysis are well understood and lead to twice the number of remaining $t\bar{t}$ background events after applying all selection cuts in the fast detector simulation compared to the full simulation. Table 6.8 shows the efficiencies of all cuts in the leptonic analysis for the full and the fast detector simulation. The only significant differences are the ones described above: trigger, lepton identification, b-jet veto. All the remaining cuts show similar efficiencies for the full and the fast detector simulation.
In Section 8.1, the Higgs discovery potential with the $t\bar{t}$ background from ATLFAST scaled by a factor of 0.5 to account for the mentioned differences is discussed.

Effects on the Semileptonic Analysis

In the semileptonic case, the underestimated τ-jet misidentification rate in the fast simulation leads to a significantly lower number of $t\bar{t}$ background events. After requiring trigger and two reconstructed τ-lepton decay products with opposite charge, the detailed simulation shows a 1.8 times higher cross-section times efficiency than the ATLFAST simulation which is expected from the different trigger, lepton and τ-jet efficiencies.
Table 6.9 shows a comparison of the efficiencies of all analysis cuts for the fast and the full simulation. The $E_{\mathrm{T}}^{\mathrm{miss}}$ cut and several jet-related cuts show a difference in efficiency, but most cut efficiencies are compatible. The distributions of discriminating variables of the cuts that show the largest difference between the fast and the full detector simulation are compared in Figure 6.15. While $\Delta\eta_{jj}$, m_{jj} and $\Delta\eta_{j\tau}^{\mathrm{min}}$ are very similar for both detector simulations, the $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution is shifted to higher values for the ATLFAST simulation leading to a higher efficiency of the $E_{\mathrm{T}}^{\mathrm{miss}} > 30\,\mathrm{GeV}$ cut. The difference can be explained by a non-optimal calibration of the τ-jet energy in the full detector simulation [49, 70].
In Section 8.1 the influence on the Higgs discovery potential if the $t\bar{t}$ background from ATLFAST is scaled up by a factor of 1.8 to account for the mentioned differences is discussed. To be conservative, in Section 8.2 an uncertainty of 100 % is applied on the $t\bar{t}$ contribution for the studies of the discovery potential taking into account systematic uncertainties.

Table 6.8: Cut efficiencies on $t\bar{t}$ events for the fast and the full detector simulation in the leptonic decay channel. The values represent the cut efficiencies without applying the other cuts before.

Cut	Efficiency (%)	
	Fast sim.	Full sim.
Trigger + N(e,μ)=2	3.8 ± 0.1	2.7 ± 0.1
Opposite charge	97 ± 0.3	96 ± 1.1
$E_{\mathrm{T}}^{\mathrm{miss}} > 40\,\mathrm{GeV}$	78 ± 0.1	78 ± 0.8
$\Delta\phi_{\tau\tau} < 2.5$	71 ± 0.1	70 ± 0.7
$\chi_{1,2} = 0.0 - 1.0$	19 ± 0.1	19 ± 0.3
$m_{\mathrm{T}}^{\ell\nu} < 100\,\mathrm{GeV}$	62 ± 0.1	61 ± 0.7
N(jets)$\geq$ 2	88 ± 0.3	88 ± 1.1
$\eta_{j1} \times \eta_{j2} < 0$	44 ± 0.1	45 ± 0.5
$\Delta\eta_{j\tau} > 0$	24 ± 0.1	24 ± 0.4
$m_{jj} > 600\,\mathrm{GeV}, \Delta\eta_{jj} > 3.5$	4.1 ± 0.1	4.2 ± 0.2
$\Delta\phi_{jj} < 2.2$	51 ± 0.1	53 ± 0.6
$\Delta\eta_{j\tau}^{\mathrm{min}} > 0.6$	7.3 ± 0.1	7.6 ± 0.2
$\Delta\eta_{j\tau}^{\mathrm{max}} > 1.5$	27 ± 0.1	28 ± 0.5
Central jet veto	37 ± 0.1	35 ± 0.5
$p_{\mathrm{T}}^{\mathrm{tot}} < 30\,\mathrm{GeV}$	38 ± 0.1	40 ± 0.5
$m_{\tau\tau} = m_H \pm 15\,\mathrm{GeV}$	5.3 ± 0.1	5.4 ± 0.2
b-jet veto	36 ± 0.1	17 ± 0.3
After $m_{jj} > 600\,\mathrm{GeV}, \Delta\eta_{jj} > 3.5$, $\eta_{j1} \times \eta_{j2} < 0, \Delta\eta_{j\tau} > 0$ and Central jet veto cut:		
b-jet veto	59 ± 1.2	38 ± 8.0

Table 6.9: Cut efficiencies on $t\bar{t}$ events for the fast and the full detector simulation in the semileptonic decay channel. The values represent the cut efficiencies without applying the other cuts before.

Cut	Efficiency (%)	
	Fast sim.	Full sim.
Trigger + N(e, μ)=1	25 ± 0.1	19 ± 0.1
N(τ)=1	1.5 ± 0.1	4.0 ± 0.1
Opposite charge	94 ± 2.9	84 ± 2.0
$E_{\mathrm{T}}^{\mathrm{miss}} > 30\,\mathrm{GeV}$	86 ± 0.3	81 ± 0.3
$\Delta\phi_{\tau\tau} < 2.5$	64 ± 0.3	66 ± 1.5
$\chi_{\ell,h} = 0.0 - 1.0$	25 ± 0.1	22 ± 0.7
$m_{\mathrm{T}}^{\ell\nu} < 30\,\mathrm{GeV}$	15 ± 0.1	16 ± 0.6
N(jets)$\geq$ 2	88 ± 2.3	90 ± 2.3
$\eta_{j1} \times \eta_{j2} < 0$	44 ± 0.2	42 ± 1.1
$\Delta\eta_{j\tau} > 0$	25 ± 0.2	22 ± 0.8
$\Delta\eta_{jj} > 3.6, m_{jj} > 600\,\mathrm{GeV}$	4.1 ± 0.1	3.2 ± 0.3
$\Delta\phi_{jj} < 2.7$	76 ± 0.3	76 ± 1.6
$\Delta\eta_{j\tau}^{\mathrm{min}} > 0.8$	4.4 ± 0.1	3.8 ± 0.3
$\Delta\eta_{j\tau}^{\mathrm{max}} > 1.5$	27 ± 0.2	23 ± 0.8
Central jet veto	35 ± 0.2	26 ± 0.1
$p_{\mathrm{T}}^{\mathrm{tot}} < 40\,\mathrm{GeV}$	50 ± 0.2	50 ± 1.2
$m_{\tau\tau} = m_H \pm 15\,\mathrm{GeV}$	6.4 ± 0.1	6.5 ± 0.4

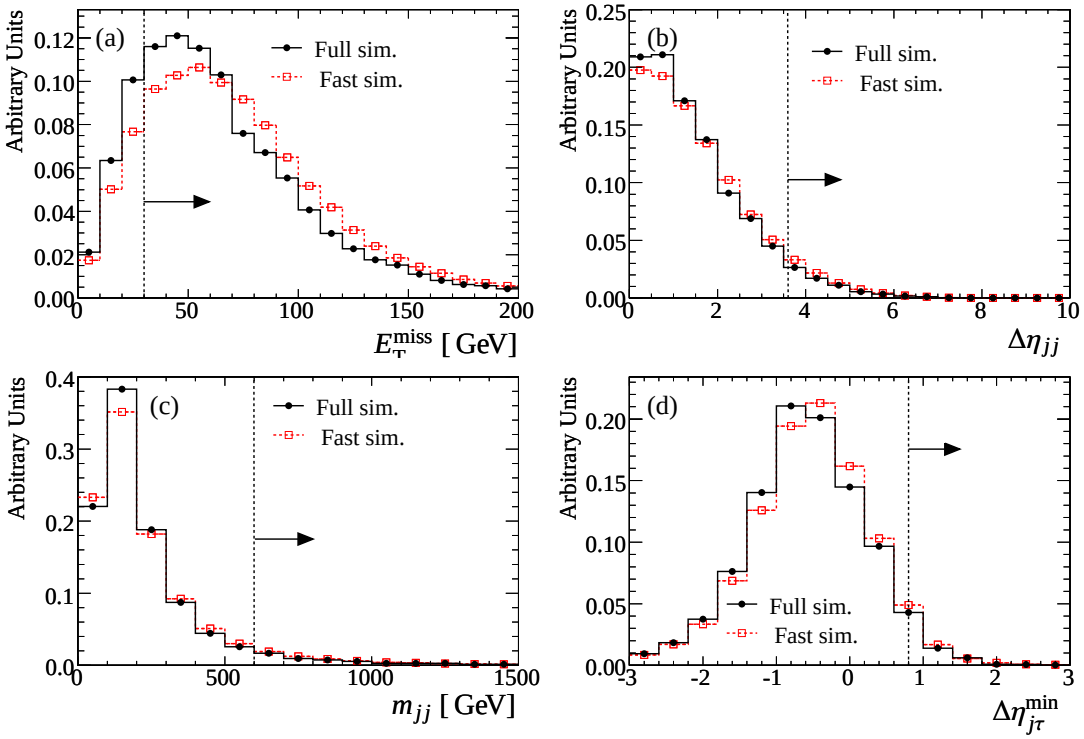


Figure 6.15: Distributions of selected discriminating variables for the full and the fast detector simulation, shown for $t\bar{t}$ in the semileptonic channel. (a) shows the E_T^{miss}, (b) the $\Delta\eta_{jj}$, (c) the m_{jj} and (d) the $\Delta\eta_{j\tau}^{min}$ distribution.

Chapter 7

Multivariate Analysis

In this chapter, the application of Multivariate Analysis (MVA) methods to the search for the Higgs boson is studied. Compared to the cut-based analysis presented in the previous chapter, multivariate analysis methods may provide a better separation between the signal and background by taking into account the information of all discriminating variables including their correlations. The advantages of MVA methods are introduced in Section 7.1.
For this study, the dedicated software package TMVA [71] has been used. The package contains a large collection of multivariate methods, from which the Fisher discriminant method, the projective likelihood method, Boosted Decision Trees (BDT) and Artificial Neural Networks (ANN) have been selected to evaluate their performance. The principle of each of these methods is described very briefly in Section 7.2. Further details can be found in many textbooks or in the TMVA User guide [71].
Before a given MVA method can be used to classify an event as signal- or background-like, it has to be trained with events identified as signal or background. For this purpose Monte Carlo data has been used.
In Section 7.3, the selection of events used for the training and in Section 7.4 the optimal selection of discriminating variables is discussed. Section 7.5 compares the performance of the different methods. The optimal results for the $H \to \tau\tau$ decay search in the leptonic and semileptonic channels are shown in Section 7.6. Finally, in Section 7.7 the stability of the results of the different multivariate methods is discussed.

7.1 Motivation

The cut-based event selection is characterized by a set of cuts applied sequentially on the discriminating variables. In case of correlations between these variables, their discriminating power is not fully exploited. For example, consider the two arbitrary discriminating variables x_1 and x_2. Signal (S) and background (B) events are distributed in the two dimensional x_1-x_2 space as illustrated in Figure 7.1a. The discrimination between signal and background in the cut-based analysis is shown in Figure 7.1b. Obviously, the selection using independent rectangular cuts does not provide optimal separation.
In contrast, MVA methods take into account correlations between the discriminating variables

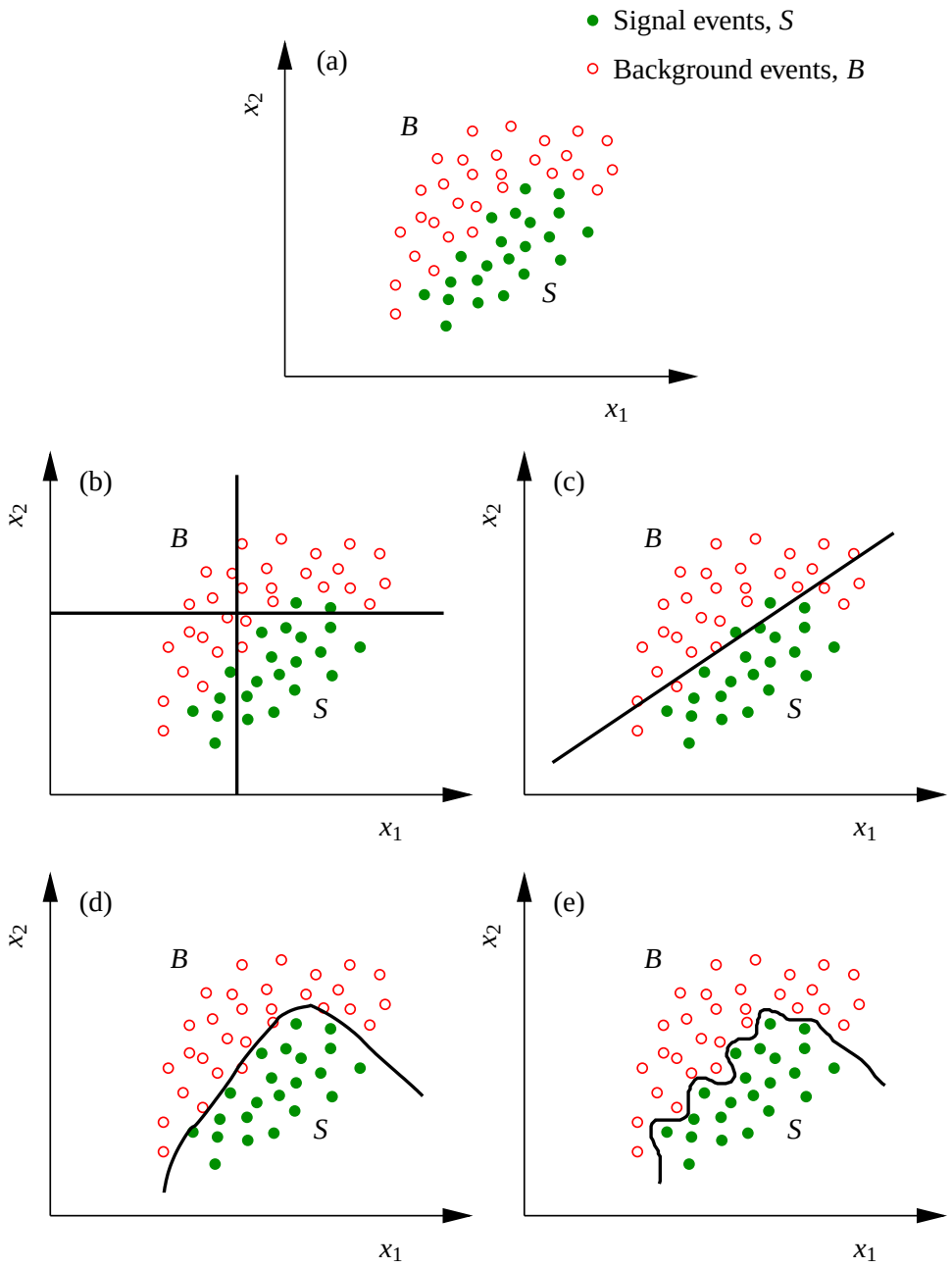


Figure 7.1: Examples for different signal and background classification strategies in two dimensions. The distribution of the events in the plane of the two discriminating variables x_1 and x_2 is shown in (a). The discrimination can be made by (b) two rectangular cuts performed independently in each variable, (c) one linear diagonal cut or (d) a nonlinear boundary between the signal and background distributions. In (e) an extreme case of a very specific nonlinear boundary is shown.

improving the separation. One way to account for correlations is to introduce a linear relationship between the cuts on the different variables resulting in linear boundaries between between signal and background regions in each two-dimensional projection of the discriminating variable space as indicated in Figure 7.1c. The Fisher discriminant method is such a linear MVA method.
However, in many cases the correlations between the variables are such that more complicated boundaries, like the one illustrated in Figure 7.1d, give the best separation. Examples for the corresponding nonlinear MVA methods are Boosted Decision Trees or Artificial Neural Networks.
Multivariate methods map the N-dimensional space of N discriminating input variables onto a single output variable taking into account the correlations in the N-dimensional space of variables. By applying a cut on this single variable, signal events can be efficiently separated from background events, considerably simplifying the determination of the optimal working point.
During the training of a MVA method, care has to be taken to avoid overtraining, which means that the MVA method recognizes individual events rather than the general features of signal and background processes as illustrated in Figure 7.1e. In this case the decision boundary does not follow the shape of the optimal boundary shown in 7.1d. An overtrained MVA will show a better separation performance for the events used in the training than for independent data samples. Usually, MVA methods with a higher number of degrees of freedom (for example BDTs) show a higher susceptibility to overtraining compared to methods with fewer degrees of freedom (for example the Fisher method). In order to obtain a reliable evaluation, it is therefore crucial to test the MVA performance on a statistically independent data sample.

7.2 Overview of Multivariate Analysis Methods

7.2.1 Projective Likelihood Method

Training of the projective likelihood method means to create for every discriminating variable k a probability density function (pdf) for the signal ($p_{S,k}(x_k)$) as well as for the background ($p_{B,k}(x_k)$). For a given event i, the likelihood for the signal S (background B) is given by:

$$L_{S(B)}(i) = \prod_k p_{S(B),k}(x_k(i)). \tag{7.1}$$

In order to classify an event i, the normalized likelihood ratio

$$y(i) = \frac{L_S(i)}{L_S(i) + L_B(i)} \tag{7.2}$$

is calculated which is the output value of the method. Signal-like events will result in high output values and background-like events in low output values.
Usually the difficulty of the likelihood method is to obtain good pdfs. Here, the pdfs have been obtained by fitting the input variable distributions of the training events by interpolating polynomial functions (spline functions) of second order. Before fitting the spline functions, the input variable distributions are smoothed to mitigate overtraining in case of limited statistics.
With this implementation of the likelihood method, correlations among the input variables are *not* taken into account. To overcome this, multidimensional pdfs $p_{S(B)}(x_1(i), \ldots, x_k(i))$ could be used. But since they have the disadvantage that one needs a very huge number of training events in particular for many discriminating variables, these methods have not been used in this study.

7.2.2 Fisher Discriminant Method

The Fisher Linear Discriminant analysis method [72], calculates the line in the multidimensional space of the k input variables that provides the best separation between signal and background. If signal and background events are projected to a line perpendicular to the separation line, the distance between their mean values is maximized and the variance of the signal as well as the background distribution is minimized.
The output variable of the Fisher method is the projected coordinate and it is determined by a linear combination of the input variables,

$$y(i) = F_0 + \sum_k F_k x_k(i), \tag{7.3}$$

where the Fisher coefficients F_k are given by:

$$F_k = \frac{\sqrt{N_S N_B}}{N_S + N_B} \sum_{l=1}^{n_{var}} W_{kl}^{-1} (\bar{x}_{S,l} - \bar{x}_{B,l}) \tag{7.4}$$

with W_{kl} being the sum of the covariant matrices $C_{S,kl}$ and $C_{B,kl}$ for signal and background respectively. $\bar{x}_{S(B),l}$ denotes the mean of the signal (background) training event distributions. $N_{S(B)}$ is the number of signal (background) training events.
By construction, the simple and robust Fisher method provides the optimum classification in case of linearly correlated and Gaussian distributed input variables.

7.2.3 Artificial Neural Networks

The basic idea of an Artificial Neural Network (ANN) is motivated by the working principle of biological neural networks. As in biology, an ANN consists of a number of interconnected "neurons". The connections transfer the output of one neuron with a certain weight to the input of another neuron. Each neuron sums up the input values and calculates an output value according to a certain response function. By feeding the values of discriminating variables to input neurons, the state of the whole network is defined corresponding to the response of a output neuron. If the ANN is trained to distinguish between different input patterns, the output of the ANN can be used to discriminate between them.
ANNs are widely used for pattern recognition and a large collection of ANN implementations exists. The type of network used in this analysis is a so-called feed-forward multilayer perceptron where the neurons are organized in layers and connections exist only between the outputs of one layer and the inputs of the next layer. Figure 7.2 shows the layout of such an ANN. The input values are fed to the neurons in the input layer ("Layer 0") and the output is formed by the single neuron in the last layer ("Layer 3"). The neuron layers 1 and 2 are not accessible from the outside and are therefore called hidden layers. The response function is usually a nonlinear function to enable the ANN to cope with nonlinear correlations. Here, the hyperbolic tangent function is used to describe the response of the neurons. The transfer function of the input neurons is 1. Each neuron is also connected to a "bias node" with constant output value 1, which allows for an additive constant in the network function.

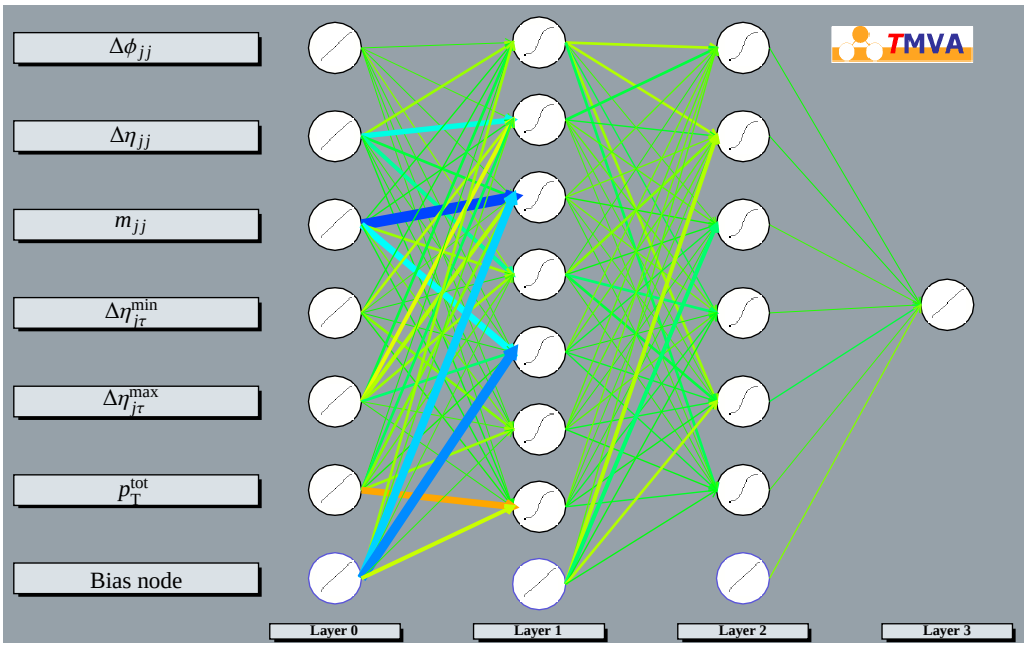


Figure 7.2: Layout of an ANN as used in this analysis. The neuron-neuron connection weights are illustrated by the thickness of the lines.

The output $y(k)$ for a given event k depends on the values of the input variables x_{ik}, the layout of the network, the weights of the neuron-neuron connections w_{ij} determined by the training of the network and the response function. The output value can be calculated by (for simplicity only one hidden layer is assumed here):

$$y(k) = \sum_{j=1}^{n_h} y_j^{(2)} w_{j1}^{(2)} = \sum_{j=1}^{n_h} \tanh\left(\sum_{i=1}^{n_{var}} x_{ik} w_{ij}^{(1)}\right) \cdot w_{j1}^{(2)} \tag{7.5}$$

where n_h (n_{var}) are the number of neurons in the hidden (input) layer, $w_{ij}^{(1)}$ is the weight of the connection between the input neuron i and the hidden neuron j and $w_{j1}^{(2)}$ are the weights between the hidden-layer neurons and the output neuron.
During the training of the ANN, an optimal set of weights $\mathbf{w}$ is determined. Starting from random weights, the weights are changed according to:

$$\mathbf{w}' = \mathbf{w} - \eta \nabla_{\mathbf{w}} E(\mathbf{x}, \mathbf{w}) \tag{7.6}$$

where $E(\mathbf{x}, \mathbf{w})$ is the so-called error function, defined by

$$E(\mathbf{x}, \mathbf{w}) = \sum_{k=1}^{N} E(\mathbf{x_k}, \mathbf{w}) = \sum_{k=1}^{N} \frac{1}{2} \left[y(k) - \hat{y}(k)\right]^2 . \tag{7.7}$$

N is the number of training events and $\hat{y}(k)$ the desired output value (1 for signal, 0 for background) of the ANN. Thus, $E(\mathbf{x}, \mathbf{w})$ is a measure of the misclassification rate for the data $\mathbf{x}$ as a function

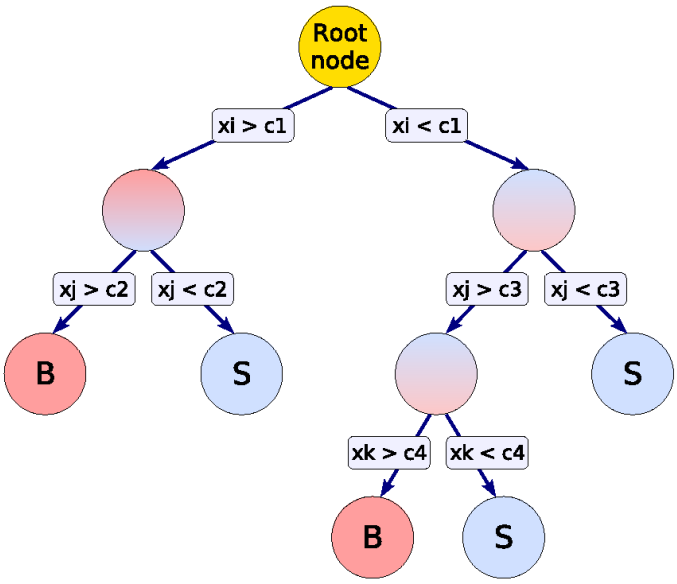


Figure 7.3: Sketch of a decision tree. Each left-right decision is based on one of the discriminating variables x_i, and classifies an event into one of the subsequent nodes [71].

of the set of weights $\mathbf{w}$ and thus $\nabla_{\mathbf{w}} E(\mathbf{x}, \mathbf{w})$ gives the direction in which $E(\mathbf{x}, \mathbf{w})$ decreases most rapidly. The step size is controlled by the factor η called the learning rate.
Usually several hundred training iterations over all training events are necessary to obtain a stable set of weights.

7.2.4 Boosted Decision Trees

Figure 7.3 shows a sketch of a decision tree which consists of repeated yes–no decisions corresponding to cuts on the discriminating variables. Each variable can be used *several times* and thus the multidimensional variable space is segmented into a large number of hypercubes. According to the number of events in a hypercube, it is classified as signal- or background-like. In this way, correlations between the variables are taken into account. Note the difference to a cut-based analysis where every variable is used only *once* to decide whether an event belongs to signal or background. Thus a cut-based analysis selects only one hypercube as signal region.
To make decision trees more robust against statistical fluctuations in the training sample, the outputs of a large collection (forest) of decision trees trained with different weights for every event in the training sample are combined to so-called Boosted Decision Trees (BDT).
During the training, the decision trees "grow". Starting from the root node, an algorithm determines for all training events (signal and background) the variable and the corresponding cut value that provides the best signal-to-background separation. This defines the first splitting in the tree. The events are divided accordingly and the procedure is repeated for both subsamples. The algorithm stops either if only a specified minimal number of events is left in the subsamples or

if a certain purity of the subsample is achieved. The resulting nodes are classified as signal or background according to the majority of the events in the node.

Boosting Before building the next tree, the events of the training sample are reweighted according to the AdaBoost algorithm [73]: The weight of every event which was misclassified by the present tree is increased by a common boost factor

$$\alpha = \frac{1 - \text{err}}{\text{err}} \tag{7.8}$$

where err is the misclassification rate of the tree. With the modified training samples, the procedure of growing and boosting a tree is repeated until the desired number of trees (usually a few hundred) is reached.
After building all trees, insignificant nodes are removed from every tree (pruning) in order to reduce overtraining effects. The Cost Complexity pruning method [74] is used here, which removes branches which have a low gain in classification performance compared to the number of additional nodes involved.
If an event k is presented to the forest of Boosted Decision Trees, the output $h_n(\mathbf{x}_k)$ of each tree n in the forest is determined with $h_n(\mathbf{x}_k) = -1$ if the corresponding node is of background type and $h_n(\mathbf{x}_i) = +1$ if the node is of signal type. The overall response of the forest is determined as the sum of the outputs of the individual trees weighted with the logarithm of their boost factors α_n:

$$y(k) = \sum_n \ln(\alpha_n) \cdot h_n(\mathbf{x}_k). \tag{7.9}$$

The application of BDTs is relatively straightforward. BDTs can cope with huge numbers of input variables while being fairly insensitive to the inclusion of poorly discriminating variables since these variables will be neglected when growing the trees. However, in case of strong linear correlations between the input variables, the performance of BDTs may profit from a decorrelation of the input variables.

7.2.5 Decorrelation of Input Variables

Certain multivariate analysis methods, for example the Projective Likelihood method, do not take into account correlations among the discriminating input variables and thus loose performance. The optimal performance could be recovered by transforming the input variables in such a way that the correlations vanish.
The TMVA package offers the possibility to decorrelate the input variables. This is done by calculating the square-root of the covariance matrix, inverting it and multiplying the input data vectors with this matrix [71]. In this way all *linear* correlations of Gaussian-distributed variables are removed. The decorrelation is done independently for signal and background since they may have different correlations. If there are nonlinear correlations, linear decorrelation provides only little improvement or can even lead to worse performance.
For two of the multivariate methods used, the improvement by applying linear decorrelation has been studied: for the Projective Likelihood which intrinsically cannot treat correlations and for the BDT which is known to potentially profit from decorrelated input variables.

7.3 Training of Multivariate Methods

For the training of multivariate methods, Monte Carlo data samples for signal and background are necessary, which have to be independent of the data used for the evaluation of the analysis performance. The number of events needed for an optimal training depends on the MVA method and on the complexity of the discrimination between signal and background. Usually between several hundred and several ten thousand events are needed to obtain a good result.
In order to obtain a sufficient number of background events for the training, only very loose preselection cuts have been applied, much looser than those used for the cut optimization in Section 6. Table 7.1 summarizes the preselection cut criteria used for the training and the testing of the MVA methods. For the training events, decay products of two τ-lepton candidates with opposite charge and two forward jets in opposite detector hemispheres were required. No trigger cuts, no $E_{\mathrm{T}}^{\mathrm{miss}}$ cut, no central- and b-jet veto and only very loose cuts on the momentum fractions $\chi_{1,2}$ have been applied.
Due to the loose preselection cuts, the multidimensional distributions of discriminating variables in the training can slightly differ from the event shapes on which the trained MVA methods are eventually applied. This leads to a loss of performance, which is also expected in the analysis of real data. The dependence of the performance of the MVA methods on the preselection cuts and on the number of training events has been studied in Section 7.7.4.
An additional preselection cut is needed for the $t\bar{t}$ background, which has been produced with the MC@NLO generator. During the generation of events, one has to avoid double counting of events when matching parton showers to matrix elements. In MC@NLO this is done by introducing negative "counter-events" to keep the cross-section right. Thus, about 13 % of the $t\bar{t}$ events have a negative weight. Since negatively weighted events cannot be used in all MVA methods, $t\bar{t}$ events with negative weight have been excluded from the training. The distributions of the input variables do not change significantly if negatively weighted events are excluded.
The training data statistics is further increased by using additional ATLFAST data samples for the $Z \to \tau\tau$ background. An overview of the data samples used for the training is shown in Table 7.3.

7.4 Selection of Input Variables

In the following, different sets of input variables have been tested in order to select the one with the best separation of signal and background events in the leptonic decay channel. It is assumed that the same variable selection will also provide the best signal-to-background separation for the semileptonic decay channel. In principle, one can include all potential discriminating variables as an input to the MVA methods. However, care has to be taken to avoid the following problems:

- **Correlation with the reconstructed Higgs boson mass**
  In order to evaluate the Higgs boson discovery potential, it is crucial that the event selection criteria do not distort the reconstructed $\tau\tau$ mass distribution for signal and background. For input variables strongly correlated with $m_{\tau\tau}$, the training can result in a mass dependent event selection and thus change or shift the $m_{\tau\tau}$ distribution. As an example for a strongly correlated variable, Figure 7.4 shows the correlation between the input variable $\Delta\phi_{\tau\tau}$ and the invariant mass $m_{\tau\tau}$ of the τ-lepton pair for signal and $t\bar{t}$ events. The resulting correlation of

Table 7.1: Event preselection criteria used for the training and performance tests of the multivariate analysis methods. Full circles indicate that the cut was applied while the cuts marked with open circles were not applied.

	Training	Testing	
		$\ell\ell$-Channel	ℓh-Channel
Trigger	○	●	●
N $(e + \mu + \tau) = 2$	●	●	●
Opposite lepton charge	●	●	●
N (jets) ≥ 2	●	●	●
$\Delta\eta_{j\tau}^{\text{min}} > 0$	●	●	●
$(\eta_1^{jet} \cdot \eta_2^{jet}) < 0$	●	●	●
$E_{\text{T}}^{\text{miss}}$	○	$> 40\,\text{GeV}$	$> 30\,\text{GeV}$
$\lvert\Delta\phi_{\tau\tau}\rvert < 2.9$	●	●	●
$\chi_{1,2}$	-2.0 ... 3.0	0 ... 1.0	0 ... 1.0
$m_{\text{T}}^{\ell\nu}$	○	○	$< 30\,\text{GeV}$
$\Delta\eta_{jj}$	○	○	○
$\Delta\phi_{jj}$	○	○	○
m_{jj}	○	○	○
b-jet veto	○	●	○
Central jet veto	○	●	●
$p_{\text{T}}^{\text{tot}}$	○	○	○
$\Delta\eta_{j\tau}^{\text{min}} > 0$	●	●	●
$\Delta\eta_{j\tau}^{\text{max}} > 0$	●	●	●

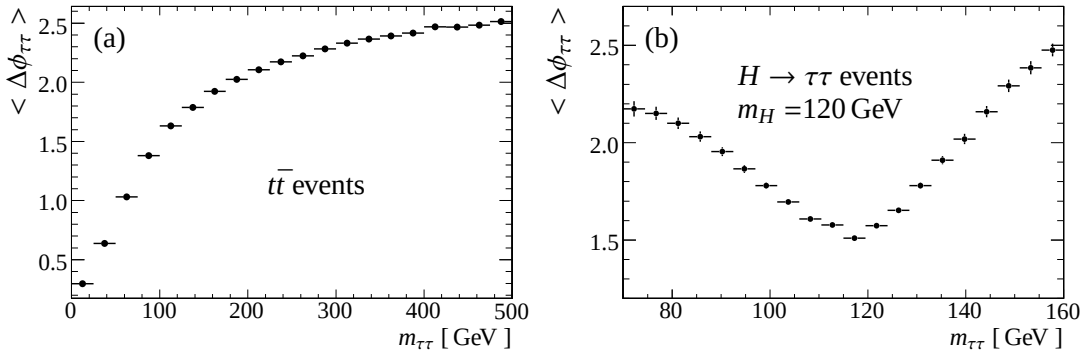


Figure 7.4: Dependence of the mean of the $\Delta\phi_{\tau\tau}$ distribution on the reconstructed invariant mass $m_{\tau\tau}$, shown in (a) for the $t\bar{t}$ background and in (b) for the signal.

the signal selection efficiency with $m_{\tau\tau}$ is shown in Figure 7.5 for an ANN with $\Delta\phi_{\tau\tau}$ as an input variable. Due to the strong mass correlation, the ANN provides a lower suppression (higher selection efficiency) of $t\bar{t}$ events near the signal region (see Figure 7.5a). Also for the signal selection efficiency (see Figure 7.5b) a mass sensitivity is observed, though not as pronounced as for the $t\bar{t}$ events. The aim is to choose a mass independent selection. The correlation of the output of each MVA method with the reconstructed $m_{\tau\tau}$ has been tested for each set of input variables. In case of a significant correlation with $m_{\tau\tau}$, the corresponding set of input variables has not been used. Also nonlinear correlations have been taken into account by looking at the resulting dependencies and accepting only those where no significant correlation is observed.

- **Weak input variables**
  Input variables providing only weak separation between signal and background are called weak variables. Depending on the multivariate method, the inclusion of weak variables can lead to a loss of performance.

- **Strongly correlated input variables**
  Using strongly correlated input variables can also lead to a loss of performance of certain MVA methods. Instead of rejecting strongly correlated variables, decorrelation of the input variables can be used. For MVA methods known to be sensitive to the correlation of input variables (Projective Likelihood, Boosted Decision Trees), the benefit of applying linear decorrelation (see Page 121) has been studied.

The input variables have been selected such that the above problems are avoided. Several sets of input variables have been tested, as summarized in Table 7.2. The selection of the optimal set was performed in three steps:

1. In order to avoid sensitivity to $m_{\tau\tau}$, only input variables without mass correlation have been used. For this purpose, the mass correlation of all potential input variables has been determined for the signal as well as for the two most important background processes, $Z \to \tau\tau$

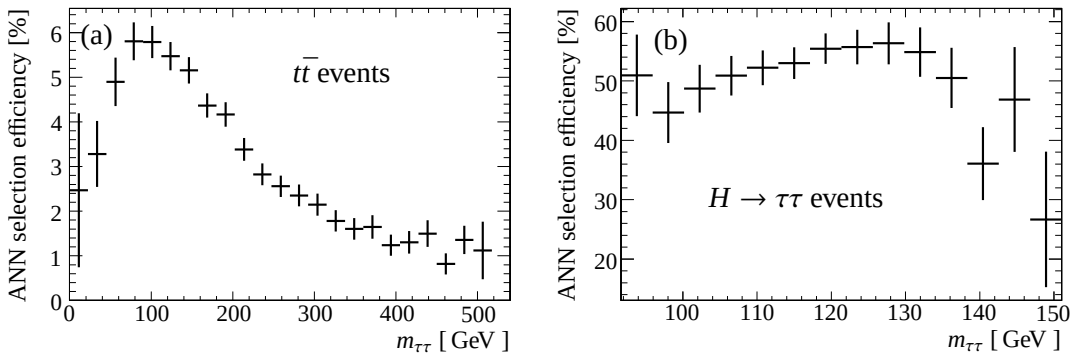


Figure 7.5: Dependence of the selection efficiency of an ANN with $\Delta\phi_{\tau\tau}$ as an input variable on the τ-pair invariant mass $m_{\tau\tau}$, shown (a) for the $t\bar{t}$ background and (b) for the signal.

and $t\bar{t}$. The correlation coefficients[1] are shown in the second, third and fourth column of Table 7.2. In the first set of input variables (set **I**), all variables with more than 10 % correlation with $m_{\tau\tau}$ have been excluded. As expected, the output obtained for this set of input variables shows no dependence on $m_{\tau\tau}$.

2. The exclusion of all mass-correlated input variables potentially reduces the signal-to-background separation power. When trying to include them, one still has to make sure that the multivariate analysis has no mass sensitivity. In the variable sets **II**, **III**, **IV** and **V** one mass dependent variable has been added at a time to the uncorrelated variables of the set **I**. It has turned out that the variables $\Delta\phi_{\tau\tau}$ (set **III**) and $m_{\mathrm{T}}^{\ell\nu}$ (set **V**) lead to a mass sensitive analysis. Therefore, these variables have not been used further.

   On the other hand, by including $E_{\mathrm{T}}^{\mathrm{miss}}$ (set **II**) or $\chi_{1,2}$ (set **IV**) as additional input variable, the multivariate output remains uncorrelated to $m_{\tau\tau}$. Thus $E_{\mathrm{T}}^{\mathrm{miss}}$ and $\chi_{1,2}$ might be used as an input for the multivariate methods. Even if both of them are used at the same time (set **VI**), the output distribution shows no dependence on the invariant mass $m_{\tau\tau}$.

3. In the following step, the possible loss of performance in some multivariate methods due to weak input variables or input variables that are correlated to other input variables, is studied. This is done by excluding certain variables and comparing the performance to the case when these variables are included. The input variable set **VII** excludes $\Delta\phi_{jj}$ as a relatively weak variable, set **VIII** excludes $\Delta\eta_{jj}$ since this variable is strongly correlated to $\Delta\eta_{j\tau}^{\mathrm{min}}$ and set **IX** excludes both $\Delta\phi_{jj}$ and $\Delta\eta_{jj}$.

[1] The correlation coefficient $\rho_{x,y}$ between two variables x and y is given by the covariance $cov(x,y) = < (x - < x >)(y - < y >) >$ of the two variables divided by their standard deviations, $\rho_{x,y} = \frac{cov(x,y)}{\sigma_x \cdot \sigma_y}$. $\rho_{x,y}$ only covers *linear* correlations and ranges from 1 for fully correlated (direct proportional) variables to -1 for fully anticorrelated (indirect proportional) variables.

Table 7.2: Overview of the sets of discriminating variables used for studying multivariate analyses in the leptonic decay channel, where full (open) circles depict variables which have (not) been used as input for the MVA method respectively. The correlation coefficients between the input variables and the invariant mass $m_{\tau\tau}$ are shown for both signal and background samples. The last row indicates whether the MVA output variable shows a correlation with $m_{\tau\tau}$ for any process.

	Correlation coefficient with $m_{\tau\tau}$ [%]			Sets of variables used as input to MVA								
	$t\bar{t}$	$Z \to \tau\tau$	$H \to \tau\tau$	**I**	**II**	**III**	**IV**	**V**	**VI**	**VII**	**VIII**	**IX**
$E_{\mathrm{T}}^{\mathrm{miss}}$	28 ± 2	9 ± 2	10 ± 1	○	●	○	○	○	●	○	○	○
$\Delta\phi_{\tau\tau}$	56 ± 2	28 ± 2	26 ± 1	○	○	●	○	○	○	○	○	○
$\Delta\phi_{jj}$	-3 ± 2	3 ± 2	5 ± 1	●	●	●	●	●	●	○	●	○
$\Delta\eta_{jj}$	7 ± 2	6 ± 2	1 ± 1	●	●	●	●	●	●	●	○	○
m_{jj}	9 ± 2	0 ± 2	2 ± 1	●	●	●	●	●	●	●	●	●
$\chi_{\ell}\chi_{h}$	-30 ± 2	-10 ± 2	-20 ± 1	○	○	○	●	○	●	○	○	○
$\Delta\eta_{j\tau}^{\mathrm{min}}$	-4 ± 2	1 ± 2	1 ± 1	●	●	●	●	●	●	●	●	●
$\Delta\eta_{j\tau}^{\mathrm{max}}$	-5 ± 2	3 ± 2	1 ± 1	●	●	●	●	●	●	●	●	●
$m_{\mathrm{T}}^{\ell\nu}$	48 ± 2	7 ± 2	10 ± 1	○	○	○	○	●	○	○	○	○
$p_{\mathrm{T}}^{\mathrm{tot}}$	3 ± 2	1 ± 2	0 ± 1	●	●	●	●	●	●	●	●	●
Correlation of output variable with $m_{\tau\tau}$				No	No	Yes	No	Yes	No	No	No	No

All sets of variables which provide a mass independent output variable, have been tested for their performance, which is described in the following.

7.5 Performance of the Multivariate Analysis Methods

7.5.1 Event Selection for Performance Tests

In order to avoid a bias in the performance evaluation due to overtraining effects, the available Monte Carlo data has been divided into two parts. Table 7.3 shows the number of signal and background events in the training and the testing samples. The same preselection as for the optimization of the cut-based analysis (see Section 6) has been applied to evaluate the performance of multivariate methods.

7.5.2 Performance Criterion

In order to compare the separation power between signal and background for different multivariate analysis methods, the background selection efficiency is measured for a given signal efficiency of 50 % as obtained in the cut-based analysis in Section 6.6. This allows for a direct comparison of the performances of multivariate and cut-based analysis methods.
Figure 7.6 shows the distributions of the output variables of the ANN for signal and several background processes for one selection of input variables. In Figure 7.7 the background efficiency as a function of the signal efficiency is shown for a range of cut values on the ANN output variable.

Table 7.3: Data samples used for the training of multivariate methods and for the analysis. The numbers of events before and after the preselection cuts are listed.

	Simulation	Training events		Analysis events	
		before preselection	after	before preselection	after
Leptonic channel:					
$H \to \tau\tau \to \ell\ell$ (120 GeV)	Full	60 000	8 272	26 750	1 660
$Z \to \tau\tau$	Fast	$9.7 \cdot 10^6$	9 687	—	—
$Z \to \tau\tau \to \ell\ell$ VBF	Full	—	—	$990 \cdot 10^3$	15 260
$t\bar{t}$	Fast	$10 \cdot 10^6$	13 460	$80 \cdot 10^6$	16 650
Semileptonic channel:					
$H \to \tau\tau \to \ell h$ (120 GeV)	Full	60 000	2 928	72 250	627
$Z \to \tau\tau$	Fast	$9.7 \cdot 10^6$	13 412	—	—
$Z \to \tau\tau$ VBF Filter	Full	—	—	$1.28 \cdot 10^6$	1 554
$t\bar{t}$	Fast	$10 \cdot 10^6$	1 896	$80 \cdot 10^6$	437

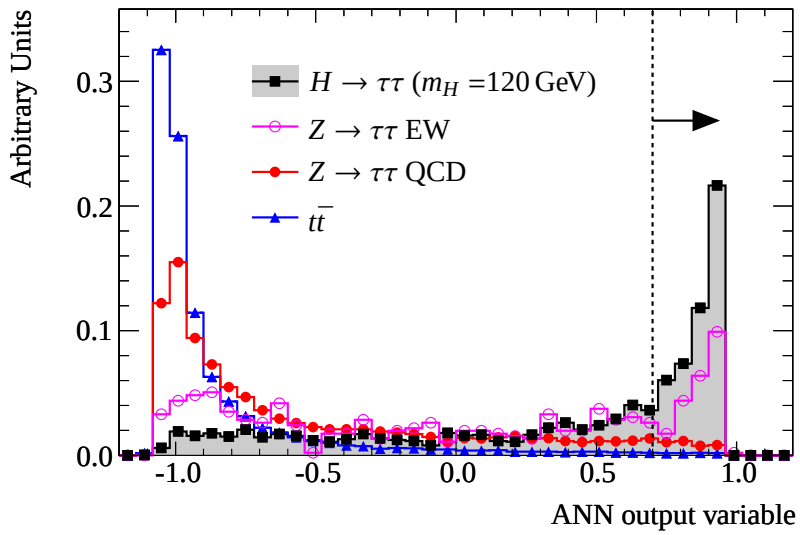


Figure 7.6: Distribution of the output variable from the ANN with input variable set **I** for the signal and the dominant background processes in the leptonic decay channel. The dashed line indicates the cut value resulting in the highest signal significance (see Section 7.6).

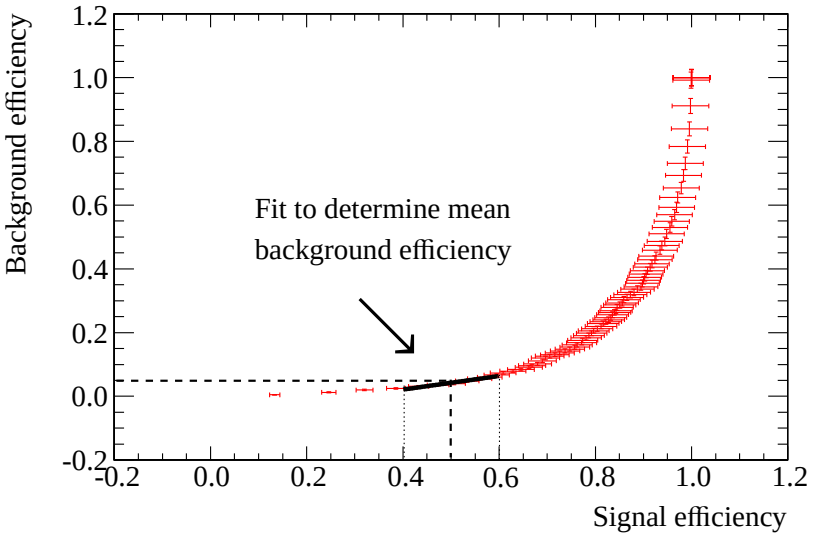


Figure 7.7: Background versus signal selection efficiency in the leptonic decay channel for a scan of cut values on the ANN output variable as shown in Figure 7.6. The mean background efficiency corresponding to 50 % signal efficiency is determined from a straight line fit in the signal efficiency interval from 40 % to 60 % as indicated.

The mean background efficiency corresponding to 50 % signal efficiency is determined by a straight line fit in the signal efficiency interval from 40 % to 60 %. This procedure has been performed for each set of input variables and for each MVA method.

7.5.3 Comparison of Multivariate Analysis Methods

Table 7.4 shows the results of the performance tests for the leptonic decay channel. For comparison, the corresponding background efficiency obtained by the cut-based analysis is 6.8 %.

The projective likelihood method was tested without and with a linear decorrelation of the input variables. For the variable sets **I**, **II**, **IV**, **VI** and **VII** the decorrelation results in a clear improvement of the performance (lower background efficiency). This is expected since the projective likelihood does not take the correlations between the variables into account. Without decorrelation the projective likelihood method shows a better performance when highly correlated variables are not used (**VIII** and **IX**). The input variable sets **II** and **VI**, in which $E_{\mathrm{T}}^{\mathrm{miss}}$ is used as additional correlated variable, are extreme cases for the projective likelihood method. While the method shows a dramatic loss in performance without decorrelation for these variable sets, it has the best performance when decorrelation is applied.

Since it is the intrinsic feature of the Fisher method to take into account linear correlations between the input variables, no decorrelation has been applied in this case. Compared to the other methods, the projective likelihood and the Fisher method show a worse performance for all sets of input variables. This indicates the existence of nonlinear correlations, which cannot be resolved by the

Table 7.4: Background efficiencies at a fixed signal efficiency of 50 %, for the multivariate analysis methods with different sets of input variables (see Table 7.2) with and without decorrelation of the input variables. The statistical error is about ±0.15 % for all values.

Method	Likelihood		Fisher	BDT		ANN
Decorrelation applied	No	Yes	No	No	Yes	No
Variable set	Background efficiency [%]					
I	10.5	7.1	6.6	5.3	4.9	4.3
II	34.5	6.1	6.3	5.0	4.7	5.0
IV	9.5	7.1	6.8	4.9	5.0	4.8
VI	35.3	6.1	6.8	5.0	4.7	5.0
VII	8.6	7.3	6.8	5.3	5.3	5.1
VIII	6.6	7.6	6.7	5.0	5.6	4.4
IX	6.3	7.5	6.9	5.4	5.5	4.8

projective likelihood or the Fisher method.
Comparing the performance for different input variable sets, it turns out that BDTs show a good performance even if many input variables have been used (for example for input variable set **VI**), independently of whether decorrelation has been applied or not. This shows the intrinsic property of BDTs to be very robust against weak variables.
The impact of input variable decorrelation can be seen by comparing the performance of the input variable sets **I** and **VIII**. Set **I** includes $\Delta\eta_{jj}$, which is strongly correlated with $\Delta\eta_{j\tau}^{\min}$, while set **VIII** does not. If $\Delta\eta_{jj}$ is excluded, the BDT with decorrelated input variables shows a loss of performance (4.9 %→5.6 % background efficiency), while the performance of the BDT without input variable decorrelation increases (5.3 %→5.0 %). Thus, for BDTs, strongly correlated variables have to be used with care and decorrelation should be applied.
The ANN method shows the best performance for most input variable sets. In contrast to the BDTs, the performance improves if only a selection of the strongest variables is used as input. If all variables are used (**VI**) the background efficiency is 5.0 %, while for selected input variables (**I**) the background efficiency is only 4.3 %.

*Since it shows the best performance, the Artificial Neural Network with input variable set **I** is used in the remainder of this study.*

7.5.4 The ANN Output Distribution

The output variable distribution of the selected ANN is shown in Figure 7.6. For signal events, the distribution peaks at large values close to 1. However, the distribution also has a significant tail down to the lowest values close to -1 which means that certain signal events are not easy to classify because they do not resemble the typical signature of the signal events. The $t\bar{t}$ background has a pronounced peak at -1 with only a small tail towards larger output values. These events can be strongly suppressed. The QCD $Z \to \tau\tau$ background is also well separated from the signal although

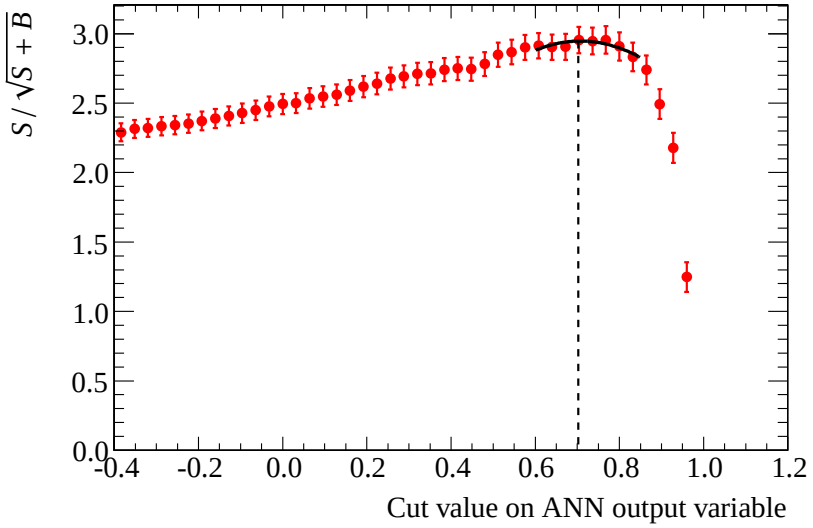


Figure 7.8: The ratio $S/\sqrt{S+B}$ as a function of the cut on the ANN output variable. The optimal cut value of 0.7 is indicated by the dashed line.

less than the $t\bar{t}$ background. For the electroweak $Z \to \tau\tau$ background, roughly half of the events are similar to the signal, while the other half resembles the QCD $Z \to \tau\tau$ background. This can be explained by the different physics processes contributing to the electroweak $Z \to \tau\tau$ events. Events originating from the VBF process (see Figure 5.2a) resemble the signal event structure, while the Z boson radiation processes (see Figure 5.2b) is similar to the QCD $Z \to \tau\tau$ process and can thus be rejected more easily.

7.6 Multivariate Analysis Results

To evaluate the improvement of the signal significance using the chosen ANN method, the output variable is treated as discriminating variable. The corresponding cut value is optimized using the iterative cut optimization procedure described in Chapter 6. In addition to the ANN output, all variables not used as input to the ANN have been included in the cut optimization procedure allowing for further discrimination between signal and background by applying a cut on those variables.

7.6.1 The Leptonic Decay Channel

For the leptonic decay channel, the dependence of $S/\sqrt{S+B}$ on the neural network cut is shown in Figure 7.8. The maximum of $S/\sqrt{S+B}$ is obtained for a cut on the output value of 0.7.
An additional cut on $m_{\mathrm{T}}^{\ell\nu} < 100\,\mathrm{GeV}$ provides a further suppression of the $t\bar{t}$ contribution while having no impact on the signal and $Z \to \tau\tau$ efficiency. Table 7.5 shows the evolution of the cross-

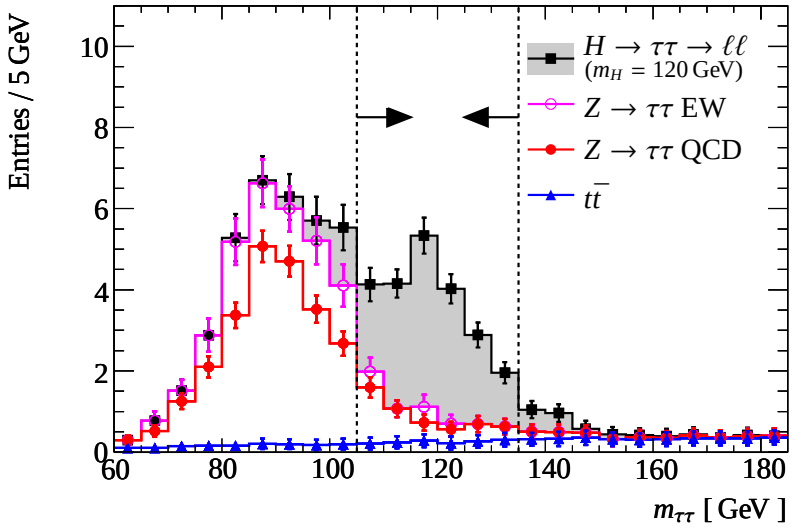


Figure 7.9: Invariant $\tau\tau$ mass distribution for the leptonic decay channel after the ANN analysis cuts. The distributions are normalized to the number of events expected for an integrated luminosity of 30 fb^{-1}.

section during the analysis, including the preselection cuts (up to the central jet veto), the cut on the ANN output variable, the additional cut on $m_{\mathrm{T}}^{\ell\nu}$ and the $m_{\tau\tau}$ mass window cut.
The invariant $m_{\tau\tau}$ spectrum after all cuts is shown in Figure 7.9 for a 120 GeV Higgs boson and an integrated luminosity of 30 fb^{-1}.

7.6.2 The Semileptonic Decay Channel

The same ANN with the same set of input variables which showed the best performance for the leptonic decay channel was used also for the semileptonic decay channel. Only the training has been repeated with semileptonic signal and background events. The optimal cut on the output of the artificial neural network is 0.3 and the signal-to-background separation is not improved by applying an additional cut on one of the variables that have not been used in the neural network.
The evolution of the cross-section is shown in Table 7.6 and the resulting $m_{\tau\tau}$ spectrum for a Higgs boson mass of 120 GeV and an integrated luminosity of 30 fb^{-1} in Figure 7.10.
In Section 8.1.1, the performance of the ANN analysis is compared to the cut-based analysis for several simulated Higgs masses.

Table 7.5: Evolution of the cross-section [fb] with the applied cuts for signal and background in the leptonic channel. The uncertainties shown are statistical only.

	$H \to \tau\tau$	$Z \to \tau\tau$ + jets		$t\bar{t}$	$t\bar{t}$
	120 GeV	QCD	EW	AtlFast	FullSim
Cross-section [fb]	22	1 410	214	833 000	461 000
Trigger	11.6 ± 0.05	657 ± 2	69 ± 1.0	$242\,000 \pm 610$	$183\,000 \pm 540$
N($e + \mu$)=2	5.5 ± 0.04	373 ± 2	7.7 ± 0.2	$30\,900 \pm 220$	$23\,200 \pm 190$
Lepton charge	5.5 ± 0.04	371 ± 2	7.5 ± 0.2	$30\,200 \pm 220$	$22\,300 \pm 190$
$E_{\mathrm{T}}^{\mathrm{miss}} > 40$ GeV	3.2 ± 0.03	147 ± 1.0	4.6 ± 0.1	$23\,500 \pm 190$	$17\,300 \pm 170$
$\Delta\phi_{\tau\tau} < 2.9$	3.1 ± 0.03	143 ± 0.44	4.5 ± 0.14	$20\,800 \pm 180$	$15\,100 \pm 160$
$\chi_{1,2} = 0.0 - 1.0$	2.7 ± 0.03	117 ± 0.39	3.7 ± 0.13	$4\,280 \pm 82$	$3\,100 \pm 71$
N(jets)≥ 2	2.2 ± 0.02	103 ± 0.36	3.6 ± 0.12	$4\,000 \pm 79$	$2\,900 \pm 68$
$\eta_{j1} \times \eta_{j2} < 0$	1.8 ± 0.02	59.1 ± 0.27	2.5 ± 0.10	$1\,820 \pm 5.6$	$1\,300 \pm 35$
$\Delta\eta_{j\tau} > 0$	1.7 ± 0.02	30.5 ± 0.19	2.0 ± 0.09	698 ± 3.5	500 ± 22
b-jet veto	1.4 ± 0.03	24.6 ± 0.17	1.6 ± 0.08	321 ± 2.5	150 ± 12
Central jet veto	1.2 ± 0.03	11.6 ± 0.11	1.0 ± 0.07	61 ± 1.0	15 ± 4
ANN > 0.7	0.67 ± 0.02	0.90 ± 0.03	0.36 ± 0.04	1.3 ± 0.16	1.4 ± 1.0
$m_{\mathrm{T}}^{\ell\nu} < 100$ GeV	0.67 ± 0.02	0.90 ± 0.03	0.35 ± 0.04	1.0 ± 0.15	1.4 ± 1.0
$m_{\tau\tau}$ =105-135 GeV	0.54 ± 0.02	0.13 ± 0.01	0.03 ± 0.01	0.07 ± 0.02[a]	0 ± 0

[a] The mass window cut factorization has been applied here; without the factorization the cross-section is 0.03 ± 0.05.

Table 7.6: Evolution of the cross-section [fb] with the applied cuts for signal and background in the semileptonic channel. The uncertainties shown are statistical only. The negative cross-section for the $t\bar{t}$ process from full simulation at the end of the analysis is due to a single event with negative weight passing all the cuts.

	$H \to \tau\tau$	$Z \to \tau\tau$ + jets		$t\bar{t}$	$t\bar{t}$
	120 GeV	QCD	EW	AtlFast	FullSim
Cross-section [fb]	145	12 900	214	833 000	461 000
Trigger	49.5 ± 0.2	3 740 ± 13	69 ± 0.55	243 000 ± 610	183 000 ± 540
N($e+\mu$)=1	43.4 ± 0.2	3 360 ± 12	61.4 ± 0.52	211 000 ± 570	16 000 ± 510
N(τ)=1	8.45 ± 0.1	337 ± 3.1	8.28 ± 0.19	3 160 ± 69	6 340 ± 100
Lepton charge	8.02 ± 0.1	311 ± 3.1	7.96 ± 0.19	2 960 ± 66	5 290 ± 92
$E_{\mathrm{T}}^{\mathrm{miss}} > 30$ GeV	4.87 ± 0.1	168 ± 1.4	5.59 ± 0.16	2 470 ± 61	4 290 ± 83
$\Delta\phi_{\tau\tau} < 2.9$	4.69 ± 0.07	163 ± 1.3	5.46 ± 0.15	2 110 ± 57	3 690 ± 77
$\chi_{\ell,h} = 0.0 - 1.0$	3.53 ± 0.06	111 ± 1.0	3.99 ± 0.13	59 ± 30	889 ± 38
$m_{\mathrm{T}}^{\ell\nu} < 30$ GeV	2.53 ± 0.05	86.1 ± 0.9	3.12 ± 0.12	110 ± 14	227 ± 19
N(jets)≥ 2	2.08 ± 0.04	76.5 ± 0.8	2.98 ± 0.11	100 ± 13	212 ± 18
$\eta_{j1} \times \eta_{j2} < 0$	1.68 ± 0.04	44.0 ± 0.6	2.11 ± 0.10	49 ± 1.0	108 ± 10
$\Delta\eta_{j\tau} > 0$	1.55 ± 0.04	23.4 ± 0.4	1.79 ± 0.09	19 ± 0.6	37 ± 6
Central jet veto	1.26 ± 0.05	10.2 ± 0.3	1.05 ± 0.07	4.6 ± 0.3	5.4 ± 2.1
ANN > 0.3	0.71 ± 0.04	1.01 ± 0.08	0.41 ± 0.04	0.15 ± 0.05	-0.68 ± 0.68
$m_{\tau\tau}$ =105-135 GeV	0.60 ± 0.04	0.08 ± 0.02	0.04 ± 0.01	0.030 ± 0.009[a]	-0.68 ± 0.68

[a]The mass window cut factorization has been applied here; without the factorization the cross-section is 0.014 ± 0.014.

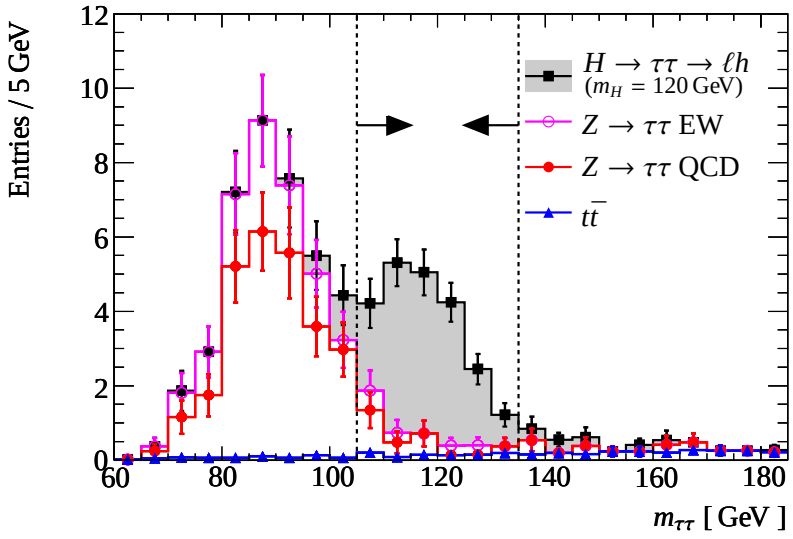


Figure 7.10: Invariant $\tau\tau$ mass distribution for the semileptonic decay channel after the ANN analysis cuts. The distributions are normalized to the number of events expected for an integrated luminosity of 30 fb^{-1}.

7.7 Systematic Tests

In this section several subjects specific to the performance and the stability of MVA methods are studied: the separate treatment of the two different background types, the influence of a different event preselection for the training, the number of events used for the training, the ANN architecture, the number of training cycles and the learning rate. Finally, the influence of the MVA method on the invariant $\tau\tau$ mass distribution is studied.

7.7.1 Separate Treatment of Backgrounds

In the described analysis, a mixture of the two main background processes $Z \to \tau\tau$ and $t\bar{t}$ weighted according to their cross-sections was used for the training of the MVA methods. Another possibility is to train one MVA method for every background process separately since in general the input variables are differently distributed for different background processes. Examples are the transverse mass $m_{\mathrm{T}}^{\ell\nu}$ (Figure 5.20) or the visible momentum fractions $\chi_{1,2}$ (Figure 5.17). In these cases, separate treatment of the backgrounds could provide better separation. Another advantage of treating the backgrounds separately is the possibility to use different sets of input variables for the rejection of different backgrounds.

A disadvantage is the need for several MVA instances. To suppress all background contributions, cuts are applied on the output variables of each MVA instance. This leads to a loss of performance since the output variables are usually correlated.

An ANN has been used to suppress the $Z \to \tau\tau$ and $t\bar{t}$ background. For discriminating between

Table 7.7: Comparison of combined and separate background treatment in the ANN analysis. The cross-sections of the signal and dominant background processes and the signal-to-background ratio after applying all analysis cuts except the mass window cut are shown.

	Cross-section [fb]			
	$H \to \tau\tau$	$Z \to \tau\tau$	$t\bar{t}$	S/B
Combined backgrounds	0.67	1.3	1.0	0.30
Separated backgrounds	0.67	1.3	1.6	0.23

signal and $t\bar{t}$ events, the same input variable set (**I**) as used in the combined background procedure shows the best performance. The best discrimination between signal and $Z \to \tau\tau$ is achieved with the input variable set **IX**. The cuts on both output variables have been optimized by the iterative cut optimization, resulting in a cut on the $Z \to \tau\tau$ discriminating ANN of ≥ 0.85 and a cut on the $t\bar{t}$ discriminating ANN of ≥ 0.8. Table 7.7 shows the result of treating the backgrounds separately. While the signal and the $Z \to \tau\tau$ cross-sections after the selection cuts are similar for the separate and the combined treatment, the $t\bar{t}$ cross-section is higher if this background is treated separately. Thus, the separate background treatment performs worse which can be explained by the fact that the distributions of the input variables used are very similar for both background types.

7.7.2 Number of Training Events

A sufficient number of events is crucial to reliably train a MVA method. In order to verify that the number of training events is sufficient, the stability of the MVA performance has been studied as a function of the number of training events.

The number of background events was varied from 100 up to 23 000 events in 90 steps. Up to 8 000 background events, the number of signal events was chosen to be the same as the number of background events. For higher background statistics, the number of signal events was kept at the maximum available number of 8 000.

Figure 7.11 shows the background efficiency for a signal efficiency of 50 % as a function of the number of background events used for the training, for the ANN and for the BDT analysis. The larger the number of training events, the more stable is the background efficiency. Above $\sim$ 8 000 training events, the performance of both methods is stable.

7.7.3 Sensitivity of the Training to Statistical Fluctuations

In order to test their sensitivity to statistical fluctuations in the training, 60 ANNs and BDTs have been trained with the same number of events, but slightly different subsets of the event sample. For the background, 20 000 out of 23 000, and for the signal, 7 000 out of 8 000 generated events have been randomly selected 60 times. The background efficiencies obtained are shown in Figures 7.12. The distributions have a Gaussian shape, both with a width of 0.15 % which is comparable to the statistical uncertainty in the determination of the background efficiency. Thus, the MVA methods are not very sensitive to statistical fluctuations in the training sample.

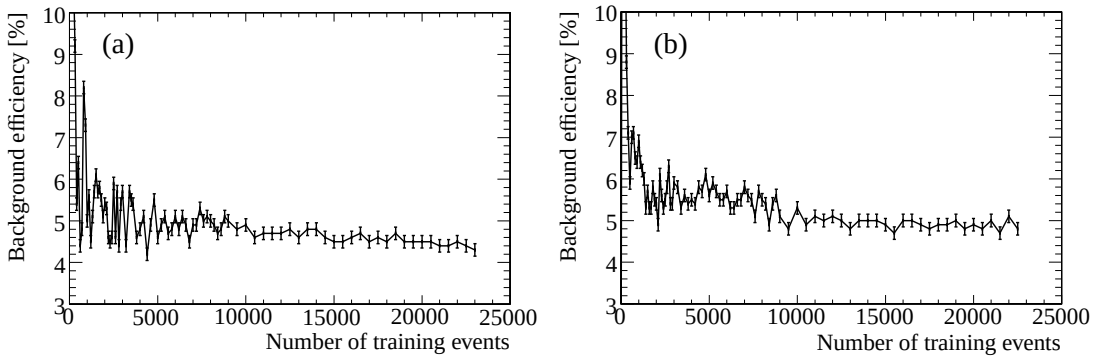


Figure 7.11: Background ($Z \to \tau\tau$ and $t\bar{t}$) efficiency for a signal efficiency of 50 % as a function of the number of background events used for the training (a) of the ANN and (b) of the BDT analysis in the leptonic channel.

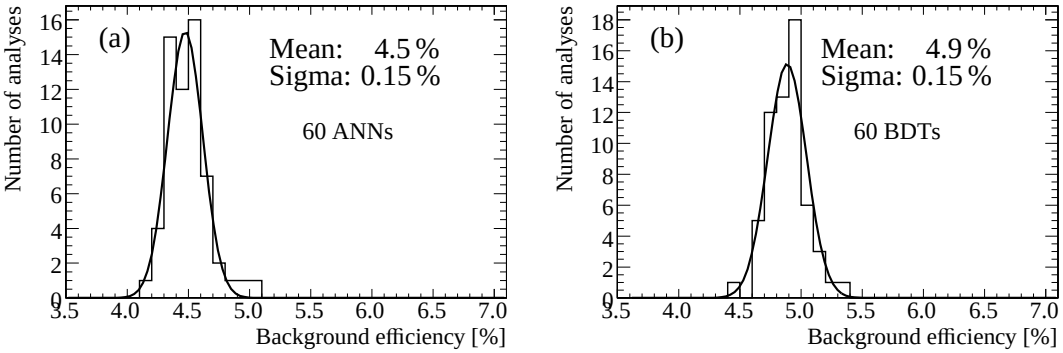


Figure 7.12: Background efficiencies obtained from (a) 60 ANN analyses and (b) 60 BDTs trained with slightly different event subsamples.

Table 7.8: Number of events obtained for the training of multivariate methods after loose and tight preselection cuts.

	Preselection	
	Loose	Tight
$H \to \tau\tau$	8 272	3 863
$Z \to \tau\tau$	6 987	3 188
$t\bar{t}$	13 460	667

7.7.4 Influence of the Preselection

In order to increase the number of training events, very loose preselection cuts have been applied up to now. The *b*-jet and central jet veto as well as the cuts on the collinear approximation and $E_{\mathrm{T}}^{\mathrm{miss}}$ have not been applied.
A potential drawback of this procedure is, that the multidimensional shape of the signal and background distributions change compared to tighter preselection cuts. To study the influence of the preselection cuts on the MVA performance, the ANN has also been trained with tighter preselection cuts where *b*-jet and central jet vetoes have been applied, the visible momentum fractions χ_1 and χ_2 have been constrained to $0 \leq \chi_{1,2} \leq 1$, and a cut on the missing transverse energy of $E_{\mathrm{T}}^{\mathrm{miss}} \geq 40\,\mathrm{GeV}$ has been applied.
Table 7.8 shows the numbers of events available for the training after applying the loose and the tight preselection criteria. For the signal and the $Z \to \tau\tau$ background, the number of available events with tight cuts is roughly two times smaller than with loose cuts while for the $t\bar{t}$ process the number of events is reduced by a factor of 20.
The background efficiency of the ANN trained after the tight preselection is $4.8\,\% \pm 0.15\,\%$ whereas the efficiency obtained with the loose preselection cuts was $4.3\,\% \pm 0.15\,\%$. This shows the advantage of a higher number of events for the training, even if the event shapes are somewhat altered compared to the actual signal selection.

7.7.5 Influence of the MVA Configuration

Several ANN configurations have been studied to test their influence on the analysis performance:

- Number of training cycles:
  The number of training cycles was varied from 1 to 5 000. The performance of the analysis increases with the number of cycles and reaches a plateau for about 100 cycles. The default setting of 200 cycles used in the analysis is therefore a reasonable value.

- Number of hidden layers and neurons:
  Several ANN architectures have been tested. If only a few neurons are added in an additional hidden layer or to one of the existing layers, the performance of the analysis is similar to the default one. If $\gtrsim$ 10 neurons are added, the analysis performance decreases. Since many more weights of neuron–neuron connections have to be tuned in such a case, the

performance decrease may be due to an insufficient number of training cycles and/or training events.

On the other hand, removing neurons leads to a decreased performance as well. With too few neurons in the network, there are not enough degrees of freedom to describe the signal and background shapes in the multidimensional space.

The default setting of two hidden layers with $N + 1$ and N neurons, where N is the number of input variables, is best suited for this analysis.

- Learning rate:
  The learning rate η was increased and decreased by factors of 5. While the ANN trained with the lower learning rate shows a selection efficiency comparable to the default setup, the higher learning rate results in a degraded performance. Again, the default setting of 0.02 for the learning rate is optimal for this analysis.

7.7.6 Influence on the $m_{\tau\tau}$ Distribution

Since the signal significance depends on the $m_{\tau\tau}$ spectrum after applying all analysis cuts, it is important that the signal selection criteria do not influence the $m_{\tau\tau}$ distributions of signal and background. Figure 7.13 shows the shape for the signal as well as for the $Z \rightarrow \tau\tau$ and the $t\bar{t}$ background separately for signal-like events (high ANN output values) and background-like events (low output values).
No significant differences in the mass shapes are observed as expected, since the set of input variables has been selected in order to minimize the bias of the $m_{\tau\tau}$ distribution (see Section 7.4).

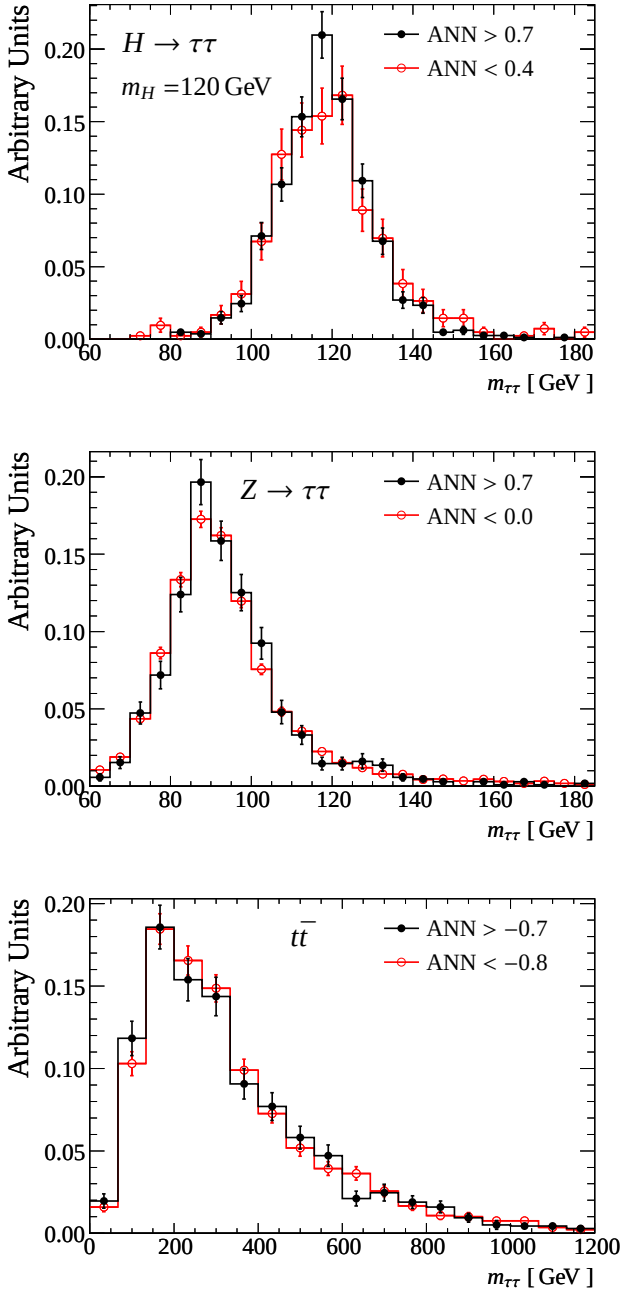


Figure 7.13: $m_{\tau\tau}$ distributions after applying the cut on the ANN output in the leptonic channel for the signal and the two most important background processes. The cuts on the neural network output variable ANN have been chosen separately for each process in order to have enough signal-like (full circles) and background-like events (open circles).

Chapter 8

Higgs Discovery Potential

This chapter discusses the discovery potential for the Higgs boson in the $H \to \tau\tau$ decay channel. After a short introduction to the calculation of the signal significance in Section 8.1, the results for the cut-based and the multivariate analysis are shown in Section 8.1.1. Finally, the impact of systematic uncertainties is evaluated in Section 8.2.

8.1 Signal Significance Determination

In order to investigate the existence of a new particle, a statistical hypothesis test is performed. H_0 be the background-only hypothesis with no Higgs boson and H_1 the hypothesis which assumes the presence of the Standard Model Higgs boson. Claiming discovery of the Higgs boson means that the H_0 hypothesis has to be rejected at a high level of confidence. A $5\,\sigma$ significance is required to claim discovery. The corresponding small probability of a wrongly claimed discovery is known as Type I error probability or p-value.
The probability p for rejecting the background-only hypothesis H_0 even though it is true is the probability of observing $n \geq N$ events in the case that only background is present:

$$p = \int_N^\infty \rho_B(n) dn \tag{8.1}$$

where N is the number of selected events in the experiment and $\rho_B(n)$ is the probability density function (pdf) for the background-only hypothesis given by the Poisson distribution

$$\rho_B(n) = \frac{\mathrm{e}^{-B} \cdot B^n}{n!} \tag{8.2}$$

with B being the number of background events.
The signal significance Z is defined as the number of standard deviations at which a Gaussian distribution centered around zero ($\mu = 0$) and standard deviation $\sigma = 1$ gives an one-sided tail area equal to the p-value. Thus, the significance Z is related to the p-value by

$$p = \int_Z^\infty \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{x^2}{2}} dx = 1 - \Phi(Z) \tag{8.3}$$

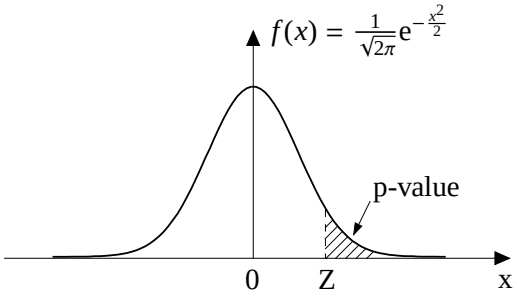


Figure 8.1: Correspondence between the signal significance Z and the p-value.

where Φ is the cumulative distribution of the normalized Gaussian function. This relation is illustrated in Figure 8.1. A signal significance of $Z = 5$ corresponds to a p-value of $2.87 \cdot 10^{-7}$.
In the limit of large background event numbers, the pdf $\rho_B(n)$ can be approximated by a Gaussian distribution with mean $\mu = B$ and a standard deviation of $\sigma = \sqrt{B}$,

$$\rho'_B(n) = \frac{1}{\sqrt{2\pi B}} e^{-\frac{(n-B)^2}{2B}}, \tag{8.4}$$

leading to the relation

$$Z = \frac{S}{\sqrt{B}} \tag{8.5}$$

with S being the number of signal events.
In a similar way, one can determine the probability to reject the signal-plus-background hypothesis H_1 even though it is true. In the approximation of Gaussian probability density the signal significance is given by

$$Z = \frac{S}{\sqrt{S+B}}. \tag{8.6}$$

In the following sections, two methods are used to determine the expected signal significance:

1. **Number counting in the mass window**
   In this approach, the number of signal plus background events is counted in the mass window $m_H - 15\,\text{GeV} \leq m_{\tau\tau} \leq m_H + 15\,\text{GeV}$ around the generated Higgs mass m_H. The signal significance is given by equation 8.6. It allows for a straightforward comparison of the discovery potential obtained with different analysis methods (baseline, optimized cut-based and ANN analysis), in particular when the influence of systematic uncertainties is considered.

   The ratio $S/\sqrt{S+B}$ has also been used as criterion for the optimization of the cut values in Section 6.

2. **Likelihood ratio hypothesis test**
   Instead of using only the single-bin information about signal and background events in the signal region of the $m_{\tau\tau}$ spectrum as described above, one can make use of the additional information given by the shapes of the invariant signal and background mass distributions.

The number of events observed is considered separately for each bin i of the mass distribution by parameterizing the signal and background distributions as follows:

$$s_i = s_{tot} \cdot \int_{\text{bin } i} F_s(m_{\tau\tau}; \boldsymbol{\theta}_s) dm_{\tau\tau}, \tag{8.7}$$

$$b_i = b_{tot} \cdot \int_{\text{bin } i} F_b(m_{\tau\tau}; \boldsymbol{\theta}_b) dm_{\tau\tau}. \tag{8.8}$$

s_{tot} (b_{tot}) be the total number of signal (background) events selected and $F_{s,b}$ be the corresponding probability density functions with sets of shape parameters $\boldsymbol{\theta}_{s,b}$.

The number of observed events n_j in a given bin j can then be compared to the prediction $(\mu s_j + b_j)$, where μ is the signal strength parameter with $0 \leq \mu \leq 1$. Since the number of events in each bin can be small, the Poisson probability density is used to define the likelihood function as

$$L(\mu, \boldsymbol{\theta}) = \prod_{j=1}^{N_{bin}} \frac{(\mu s_j + b_j)^{n_j} e^{-(\mu s_j + b_j)}}{n_j!}. \tag{8.9}$$

A binned maximum likelihood $L(\hat{\mu}, \hat{\boldsymbol{\theta}})$ fit to the measured mass distribution is performed with μ as a free parameter and the shape parameters $\boldsymbol{\theta} = (\boldsymbol{\theta}_s, \boldsymbol{\theta}_b)$. The result of the fit is denoted as $\hat{\mu}$ and $\hat{\boldsymbol{\theta}}$.

The background shape can additionally be constrained by subsidiary measurements of the so-called control data samples which are free of signal contribution. These measurements enter the likelihood function as additional background-only bins.

The background-only hypothesis H_0 corresponds to $\mu = 0$ and can be tested by evaluating the profile likelihood ratio

$$\lambda(\mu = 0) = \frac{L(\mu = 0, \hat{\hat{\boldsymbol{\theta}}})}{L(\hat{\mu}, \hat{\boldsymbol{\theta}})} \tag{8.10}$$

where $\hat{\hat{\boldsymbol{\theta}}}$ is the set of shape parameters which maximizes the likelihood for a fixed signal strength of $\mu = 0$. According to [75, 49] the signal significance can be approximated by

$$Z = \sqrt{-2 \ln \lambda(0)}. \tag{8.11}$$

Since λ always fulfills $0 \leq \lambda(0) \leq 1$, a small value of $\lambda(0)$ (corresponding to a low probability of the H_0 hypothesis for the observed data) will lead to a high signal significance Z.

The advantage of the likelihood ratio hypothesis test is that the systematic uncertainties on the shapes of the signal and background distributions can be directly taken into account by introducing additional fit parameters. Furthermore, different datasets k can be combined by multiplying the individual likelihood ratios $\lambda_k(0)$. This has been done for the combination of the leptonic and the semileptonic channel. Further details of the profile likelihood method can be found in [49].

8.1.1 Results without Systematic Uncertainties

The performance of the cut-based and the ANN analysis are compared for the leptonic and the semileptonic decay mode. Both methods have been optimized for a Higgs mass of $m_H = 120\,\text{GeV}$ and then used to evaluate the discovery potential also for Higgs boson masses of $m_H = 105$, 110, 115, 120, 125, 130 and 135 GeV.
Table 8.1 gives an overview of the results for the leptonic and the semileptonic channel. In order to allow for a comparison, the results obtained with the baseline analysis [49] are also given. The expected signal and background cross-sections after the signal selection as well as the signal significances $S/\sqrt{S+B}$ are shown for an integrated luminosity of $30\,\text{fb}^{-1}$ in the $1.5\,\sigma$ window around the generated Higgs mass ($m_H - 15\,\text{GeV} \leq m_{\tau\tau} \leq m_H + 15\,\text{GeV}$).
Compared to the baseline analysis, the number of signal events selected by the optimized cut-based analysis is very similar while the number of background events (in particular in the $t\bar{t}$ background) is significantly reduced. This leads to higher signal significances for all Higgs mass points in the leptonic as well as in the semileptonic channel. The background contribution in the ANN analysis is similar to the one obtained from the optimized cut-based analysis while there is a gain in the number of signal events in the ANN analysis. This increase of the signal efficiency is more pronounced in the leptonic than in the semileptonic mode. Due to the lower $t\bar{t}$ background contribution, the semileptonic mode provides a higher signal significance than the leptonic mode for all three analysis methods.
Although the cut-based and the ANN analysis have been optimized for a Higgs boson mass of $m_H = 120\,\text{GeV}$, the performance is similar for all other masses. The $S/\sqrt{S+B}$ signal significances are shown in Figure 8.2 as a function of the Higgs boson mass. The improved performance of the optimized analysis methods is independently observed of the Higgs boson mass. This is particularly important for the ANN analysis, where the training for a 120 GeV signal could have led to an ANN that shows a good performance only for the trained Higgs mass. The highest signal significance is obtained for a Higgs boson mass of about 120 GeV. The lower signal significances observed at lower Higgs masses are due to the increase of the $Z \to \tau\tau$ background contribution as one approaches the peak of the Z-resonance. The decrease of the $H \to \tau\tau$ branching ratio leads to a decreasing signal significance for masses above $m_H \gtrsim 120\,\text{GeV}$.

8.2 Systematic Uncertainties

The results presented so far do not take systematic uncertainties into account. In the following, theoretical uncertainties and systematic uncertainties originating from imperfect knowledge of the detector response and their impact on the expected discovery potential are discussed. For the $t\bar{t}$ background, the use of the fast instead of the full detector simulation is an additional source of systematic uncertainty.

8.2.1 Theoretical Uncertainties

Uncertainties on the predicted cross-sections of the signal and background processes are due to the renormalization and factorization scale dependence as well as uncertainties on the parton distribution function. For the signal, the total cross-section uncertainty has been estimated to be less

Table 8.1: Overview of the expected signal and background cross-sections [fb] after all analysis cuts except for the mass window cut and the signal significances $S/\sqrt{S+B}$ for an integrated luminosity of $30\,\text{fb}^{-1}$ calculated in the 1.5σ mass window ($m_H - 15\,\text{GeV} \leq m_{\tau\tau} \leq m_H + 15\,\text{GeV}$) around the generated Higgs boson mass. The results are shown for the optimized cut-based analysis and the ANN analysis as a function of the Higgs mass separately for the leptonic ($\ell\ell$) and the semileptonic (ℓh) decay channel. The results obtained by the baseline analysis are given for comparison. The uncertainties are due to the limited Monte Carlo statistics.

	Baseline analysis		Optimized cut-based analysis		ANN analysis	
	cross-section	$S/\sqrt{S+B}$	cross-section	$S/\sqrt{S+B}$	cross-section	$S/\sqrt{S+B}$
$\ell\ell$ channel	[fb]		[fb]		[fb]	
$Z \to \tau\tau$ QCD	1.3 ± 0.04		1.02 ± 0.03		0.9 ± 0.03	
$Z \to \tau\tau$ EW	0.43 ± 0.04		0.23 ± 0.03		0.35 ± 0.04	
$t\bar{t}$	4.7 ± 0.28		1.5 ± 0.16		1.0 ± 0.15	
H (105 GeV)	0.61 ± 0.03	2.3 ± 0.09	0.65 ± 0.03	2.7 ± 0.10	0.74 ± 0.03	2.9 ± 0.10
H (110 GeV)	0.59 ± 0.02	2.4 ± 0.10	0.62 ± 0.02	2.9 ± 0.10	0.73 ± 0.03	3.2 ± 0.11
H (115 GeV)	0.62 ± 0.02	2.6 ± 0.10	0.63 ± 0.02	3.1 ± 0.11	0.73 ± 0.03	3.4 ± 0.11
H (120 GeV)	0.54 ± 0.02	2.5 ± 0.10	0.58 ± 0.02	3.0 ± 0.11	0.67 ± 0.02	3.4 ± 0.11
H (125 GeV)	0.49 ± 0.02	2.3 ± 0.10	0.53 ± 0.02	2.9 ± 0.11	0.60 ± 0.02	3.2 ± 0.11
H (130 GeV)	0.44 ± 0.02	2.1 ± 0.10	0.44 ± 0.02	2.5 ± 0.11	0.53 ± 0.02	2.9 ± 0.11
H (135 GeV)	0.37 ± 0.01	1.9 ± 0.09	0.38 ± 0.01	2.3 ± 0.10	0.44 ± 0.01	2.6 ± 0.11
ℓh channel						
$Z \to \tau\tau$ QCD	1.2 ± 0.10		1.0 ± 0.09		1.0 ± 0.08	
$Z \to \tau\tau$ EW	0.46 ± 0.05		0.38 ± 0.04		0.41 ± 0.04	
$t\bar{t}$	0.28 ± 0.07		0.14 ± 0.05		0.15 ± 0.05	
H (105 GeV)	0.74 ± 0.04	2.9 ± 0.17	0.75 ± 0.05	3.2 ± 0.17	0.81 ± 0.05	3.3 ± 0.17
H (110 GeV)	0.72 ± 0.04	3.2 ± 0.16	0.73 ± 0.04	3.5 ± 0.16	0.81 ± 0.04	3.7 ± 0.16
H (115 GeV)	0.78 ± 0.07	3.6 ± 0.24	0.75 ± 0.07	3.8 ± 0.24	0.80 ± 0.07	3.9 ± 0.24
H (120 GeV)	0.67 ± 0.04	3.6 ± 0.15	0.67 ± 0.04	3.7 ± 0.15	0.71 ± 0.04	3.8 ± 0.15
H (125 GeV)	0.58 ± 0.03	3.4 ± 0.13	0.59 ± 0.03	3.4 ± 0.13	0.68 ± 0.03	3.7 ± 0.13
H (130 GeV)	0.53 ± 0.05	3.1 ± 0.20	0.55 ± 0.05	3.2 ± 0.20	0.64 ± 0.05	3.4 ± 0.20
H (135 GeV)	0.45 ± 0.02	2.9 ± 0.12	0.44 ± 0.02	2.9 ± 0.12	0.50 ± 0.02	3.1 ± 0.11

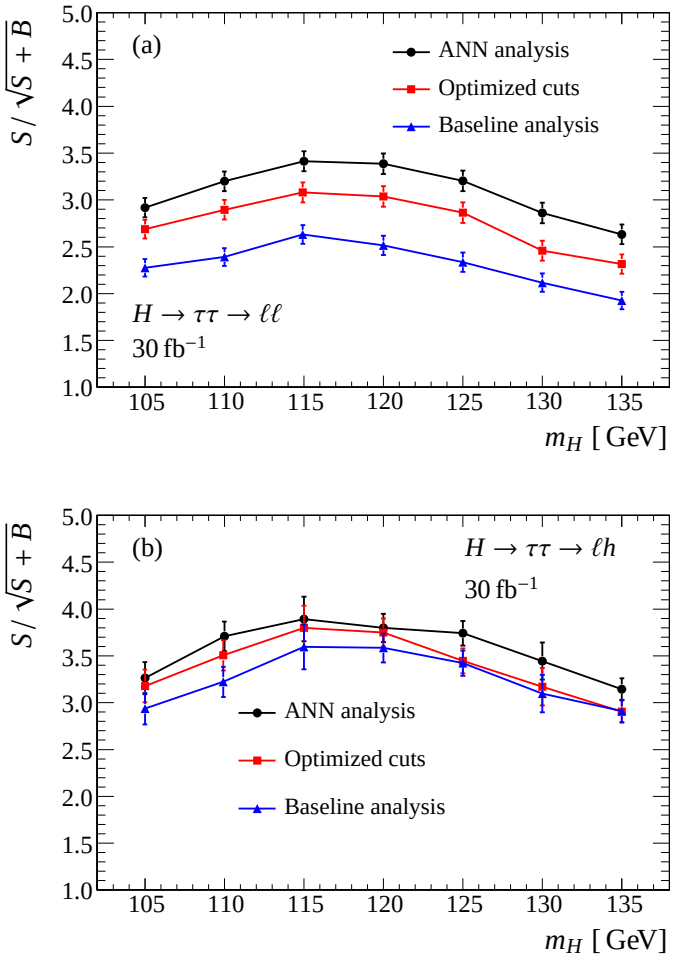


Figure 8.2: Expected signal significance $S/\sqrt{S+B}$ of the optimized cut-based and the ANN analysis for (a) the leptonic and (b) the semileptonic channel and an integrated luminosity of 30 fb^{-1} as a function of the Higgs mass. For comparison, the results obtained by the baseline analysis are shown. The error bars indicate the uncertainty due to the limited Monte Carlo statistics.

than 10 % [49]. Also for the Z production [76] and for the $t\bar{t}$ process [77] the theoretical uncertainties are expected to be $\leq 10\%$. The impact of the systematic uncertainties on the Higgs discovery potential is discussed in Section 8.2.4.

8.2.2 Uncertainty due to the Fast Simulation for the $t\bar{t}$ Background

As described in Section 5.2, the electron, muon, τ-jet and b-jet reconstruction and identification efficiencies in the fast detector simulation ATLFAST differ from the predictions of the full simulation. The impact of the differences on the selected $t\bar{t}$ cross-section has been estimated in Section 6.6.3. In case of the leptonic channel, the $t\bar{t}$ contribution is overestimated in the fast simulation by a factor of 2, while for the semileptonic channel it is underestimated by a factor of 1.8.
Table 8.2 shows the resulting signal significances $S/\sqrt{S+B}$ after rescaling of the $t\bar{t}$ background contribution, neglecting all other uncertainties. Compared to the results without rescaling of the $t\bar{t}$ background (see Table 8.1), one observes slightly higher signal significances for the leptonic channel, while for the semileptonic channel the significances slightly decrease. However, the observed differences of the signal significance of less than 5 % are rather small.
Since the scaling is only a rough estimate with limited reliability, it has *not* been applied for the final results. The described difference is taken into account as an additional systematic uncertainty when studying the impact on the signal significance in Section 8.2.4: Instead of reducing the $t\bar{t}$ contribution by 50 %, an additional uncertainty of 50 % is assigned to the $t\bar{t}$ background in the leptonic channel and instead of scaling with a factor 1.8 an uncertainty of 100 % is added to the other uncertainties on the $t\bar{t}$ background in the semileptonic channel.

8.2.3 Detector-Related Experimental Uncertainties

Uncertainties in the determination of the energy calibration, resolution and efficiency of the detector components can lead to uncertainties in the prediction of the analysis results from Monte Carlo data. The detector-related systematic effects studied for both the signal and the background simulations are summarized in Table 8.3. The magnitude of systematic uncertainties corresponds to a knowledge of the detector response achievable with an integrated luminosity of $10\,\mathrm{fb}^{-1}$ as agreed in the ATLAS collaboration for the CSC analyses [49].
The reconstruction efficiency, the energy resolution and the energy scale have been varied for electrons, muons, τ-jets and jets, by the amounts summarized in Table 8.3. The uncertainty of the missing transverse energy ($E_{\mathrm{T}}^{\mathrm{miss}}$) measurement is taken into account by varying the electron, muon, τ-jet and jet energies and recalculating the $E_{\mathrm{T}}^{\mathrm{miss}}$ vector. Since additional information about the $E_{\mathrm{T}}^{\mathrm{miss}}$ measurement can be obtained from real data, only a 5 % variation in the jet energy is used for the calculation of the missing transverse energy.
The uncertainty on the b-jet identification efficiency is $\pm 5\%$ and an uncertainty of $\pm 10\%$ is assigned to the rejection of light quark jets. However, the influence of systematic effects related to the b-jet identification cannot be studied in detail since it requires a full simulation of the $t\bar{t}$ background with very high statistics. Nevertheless, it can be assumed that the uncertainty in the b-jet identification efficiency leads to an uncertainty of $\leq 5\,\%$ in the $t\bar{t}$ contribution. b-jets are neither involved in the signal nor in the $Z \rightarrow \tau\tau$ production and thus, these two processes are only sensitive to the light jet rejection uncertainty. Since the b-jet veto requirement rejects only $10 - 20\,\%$

Table 8.2: Expected signal significances $S/\sqrt{S+B}$ for the baseline analysis, the optimized cut-based and the ANN analysis after taking into account the differences between the fast and the full detector simulation as described in Section 6.6.3: For the leptonic (semileptonic) channel, the $t\bar{t}$ cross-section has been scaled by a factor of 0.5 (1.8).

	Baseline analysis	Optimized cut based analysis	ANN analysis
$\ell\ell$ channel, $t\bar{t}$ cross-section $\times 0.5$			
H (105 GeV)	2.4 ± 0.09	2.8 ± 0.10	3.0 ± 0.10
H (110 GeV)	2.5 ± 0.09	3.0 ± 0.10	3.3 ± 0.10
H (115 GeV)	2.8 ± 0.10	3.2 ± 0.10	3.5 ± 0.10
H (120 GeV)	2.7 ± 0.10	3.2 ± 0.10	3.5 ± 0.10
H (125 GeV)	2.5 ± 0.10	3.0 ± 0.10	3.3 ± 0.10
H (130 GeV)	2.3 ± 0.09	2.6 ± 0.09	3.0 ± 0.09
H (135 GeV)	2.2 ± 0.09	2.5 ± 0.09	2.8 ± 0.08
ℓh channel, $t\bar{t}$ cross-section $\times 1.8$			
H (105 GeV)	2.9 ± 0.17	3.2 ± 0.17	3.2 ± 0.17
H (110 GeV)	3.2 ± 0.16	3.5 ± 0.16	3.7 ± 0.16
H (115 GeV)	3.6 ± 0.24	3.8 ± 0.24	3.9 ± 0.24
H (120 GeV)	3.5 ± 0.16	3.7 ± 0.16	3.7 ± 0.16
H (125 GeV)	3.4 ± 0.14	3.4 ± 0.14	3.7 ± 0.14
H (130 GeV)	3.0 ± 0.21	3.1 ± 0.21	3.4 ± 0.21
H (135 GeV)	2.8 ± 0.13	2.8 ± 0.13	3.1 ± 0.13

Table 8.3: Systematic uncertainties on the knowledge of the detector response corresponding to an integrated luminosity of 10 fb^{-1}.

Observable	Relative uncertainty
Electron energy scale	$\pm 0.5\,\%$
Electron energy resolution	$\sigma(E_\text{T}) \oplus 7.3 \cdot 10^{-3} E_\text{T}$
Electron reconstruction efficiency	$-0.2\,\%$
Muon energy scale	$\pm 1.0\,\%$
Muon momentum resolution	$\sigma(p_\text{T}) \oplus 0.011 p_\text{T} \oplus 1.7 \cdot 10^{-4} p_\text{T}^2$
Muon reconstruction efficiency	$-1\,\%$
τ-jet energy scale	$\pm 5.0\,\%$
τ-jet momentum resolution	$\sigma(E) \oplus 0.45\sqrt{E}$
τ-jet reconstruction efficiency	$-5\,\%$
Jet energy scale	$\pm 7.0\,\%$ ($\lvert\eta\rvert < 3.2$)
	$\pm 15.0\,\%$ ($\lvert\eta\rvert > 3.2$)
Jet energy resolution	$\sigma(E) \oplus 0.45\sqrt{E}$ ($\lvert\eta\rvert < 3.2$)
	$\sigma(E) \oplus 0.67\sqrt{E}$ ($\lvert\eta\rvert > 3.2$)

of signal and $Z \to \tau\tau$ events, the light jet rejection uncertainty of ±10% is assumed to change the signal and $Z \to \tau\tau$ contribution only by a few percent. These uncertainties are rather small compared to other detector-related effects and have thus been neglected in the following.

The impact of the systematic variations has been studied by applying the full set of analysis cuts. The ANN analysis was performed on data samples with systematic variations introduced while the ANN was still trained with unmodified events.

Figures 8.3, 8.4 and 8.5 show the distributions of important discriminating variables with and without the systematic shift of the jet energy scale towards lower values. This systematic effect has the biggest impact on the event selection as shown below.

p_T^tot is the variable which is most sensitive to the jet energy scale uncertainty. The shift of p_T^tot (calculated by equation 5.12) to higher values can be explained by the fact that the uncertainty on the jet energy scale affects the missing transverse energy less than the visible jet energy as explained above. A large fraction of signal events has low p_T^tot values. Thus, a change of the jet energy scale influences the efficiency of the p_T^tot <30 GeV selection cut of the signal more than the selection efficiency of the background.

The effect of the negative shift of the jet energy scale on the distribution of the ANN output variable is shown in Figure 8.6. One can see that the output values for the signal as well as for the two main background contributions are shifted towards lower values resulting in a lower selection efficiency for all processes.

Table 8.4 shows the impact of the shift in the jet energy scale on the selection efficiencies in the leptonic channel. As expected, the effect is largest for the cuts on E_T^miss, on the number of jets

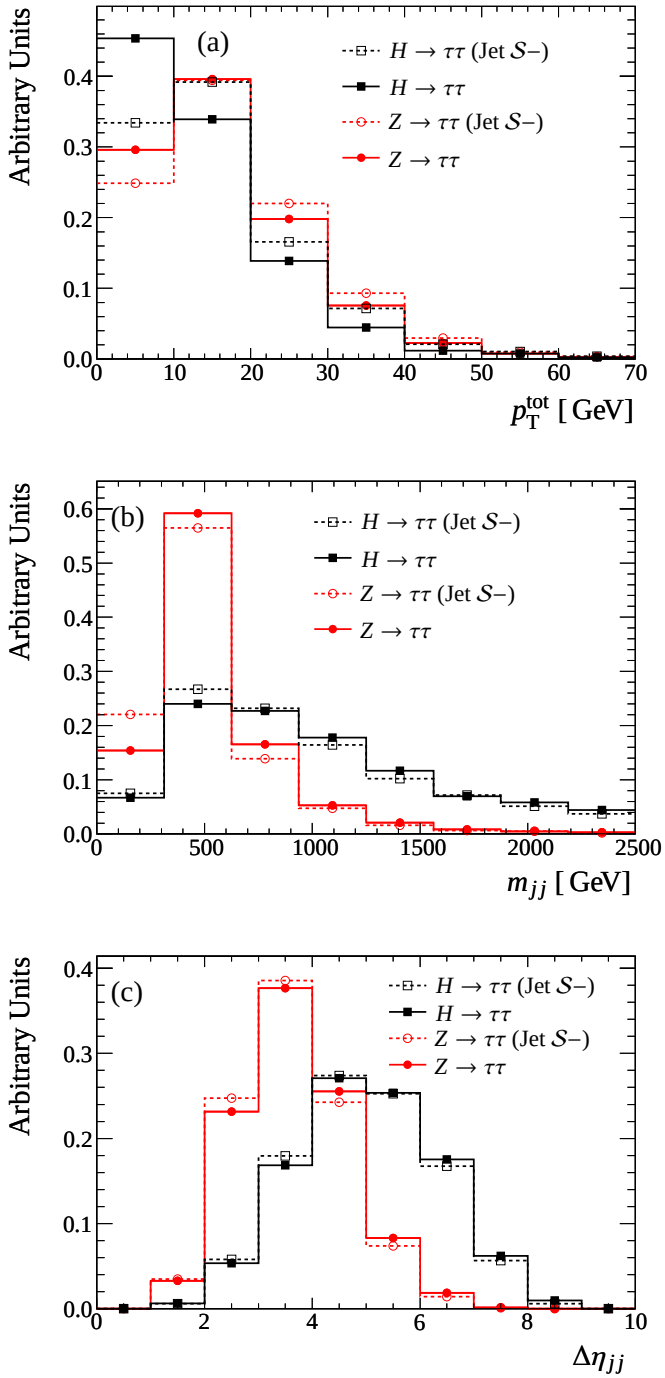


Figure 8.3: Distributions of (a) p_T^{tot}, (b) m_{jj} and (c) $\Delta\eta_{jj}$ for the signal (squares) and the $Z \to \tau\tau$ background (circles) without (full lines) and with a systematic shift of the jet energy scale towards lower values (Jet $\mathcal{S}-$, dashed lines).

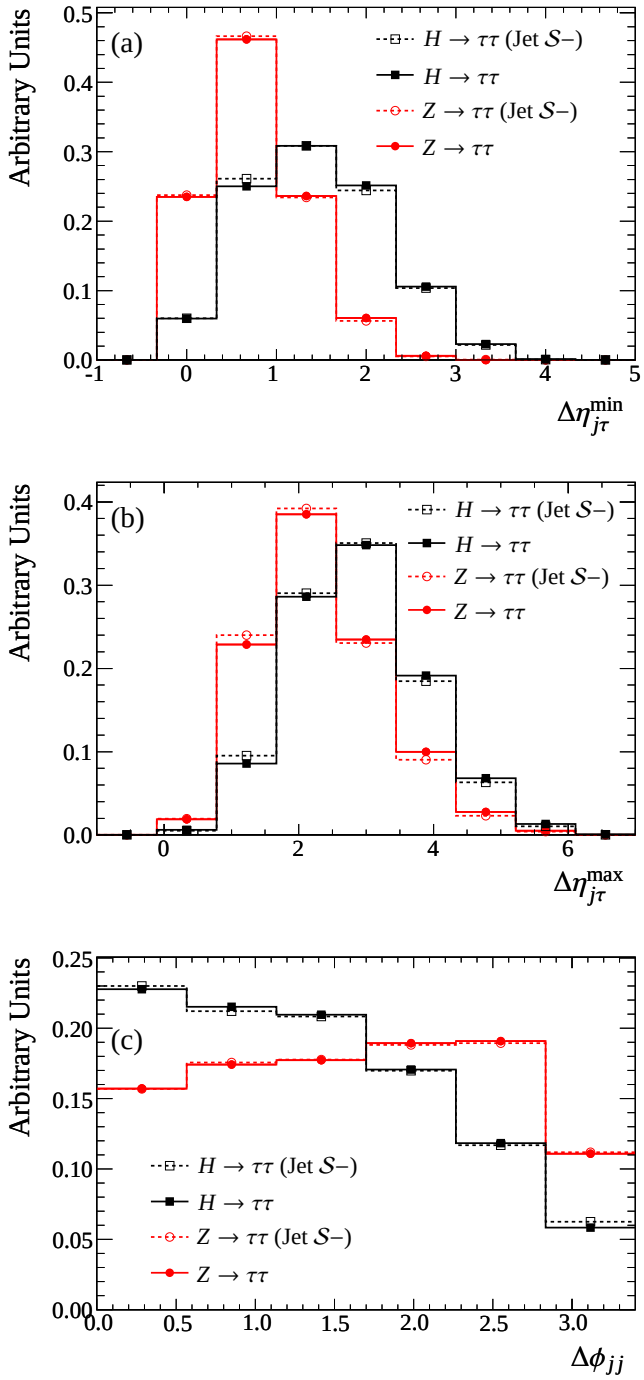


Figure 8.4: Distributions of (a) $\Delta\eta_{j\tau}^{\text{min}}$, (b) $\Delta\eta_{j\tau}^{\text{max}}$ and (c) $\Delta\phi_{jj}$ for the signal (squares) and the $Z \to \tau\tau$ background (circles) without (full lines) and with a systematic shift of the jet energy scale towards lower values (Jet $\mathcal{S}-$, dashed lines).

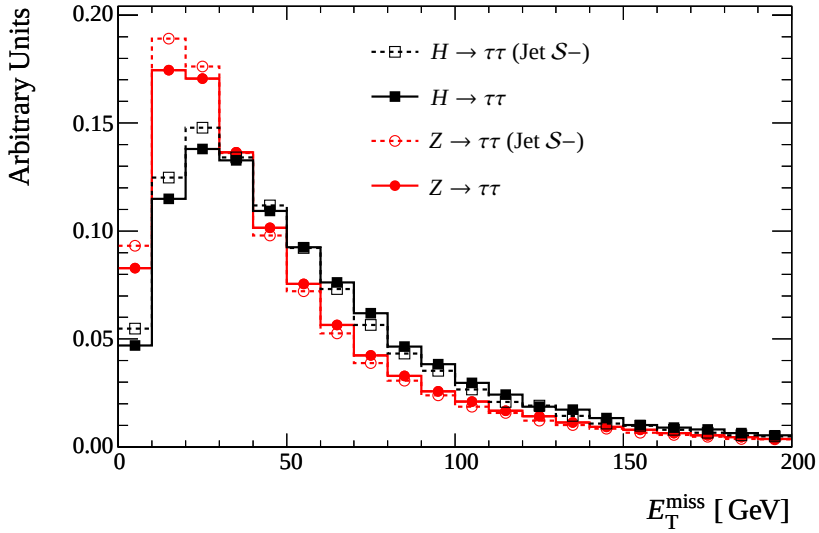


Figure 8.5: Distribution of the missing transverse energy E_T^{miss} for the signal (squares) and the $Z \to \tau\tau$ background (circles) without (full lines) and with a systematic shift of the jet energy scale towards lower values (Jet $\mathcal{S}-$, dashed lines).

(with $p_T > 20\,\text{GeV}$), on the jet pair mass m_{jj} and on p_T^{tot}.
Tables 8.5 and 8.6 show the impact of the systematic effects on the signal and background cross-sections in the leptonic and semileptonic channel respectively after applying all analysis cuts except for the mass window cut. The systematic uncertainties related to the electron and muon reconstruction performance are negligible. The largest uncertainty is caused by the jet energy scale. The uncertainties related to the τ-jet reconstruction have a significant impact on the final cross-sections in the semileptonic channel, too, although the changes are smaller than for the jet-related uncertainties.
The optimized cut-based and the ANN analysis show similar sensitivity to the systematic uncertainties indicating the robustness of the ANN analysis.

8.2.4 Impact of Systematic Uncertainties on the Expected Discovery Potential

Systematic uncertainties of the detector performance change the number of signal and background events expected to be observed on real data as discussed above and lead to a change of the predicted signal significance. In addition, a systematic uncertainty has to be added to the significance calculation depending on how the background contribution is determined.
Three different scenarios have been considered for the prediction of the signal significance using different assumptions regarding the uncertainties on the expected signal and background contributions: Without and with systematic uncertainties on the background estimation from Monte Carlo data as well as the case when the background contribution is is determined from real data (so-called data-driven background estimation by means of control data samples).

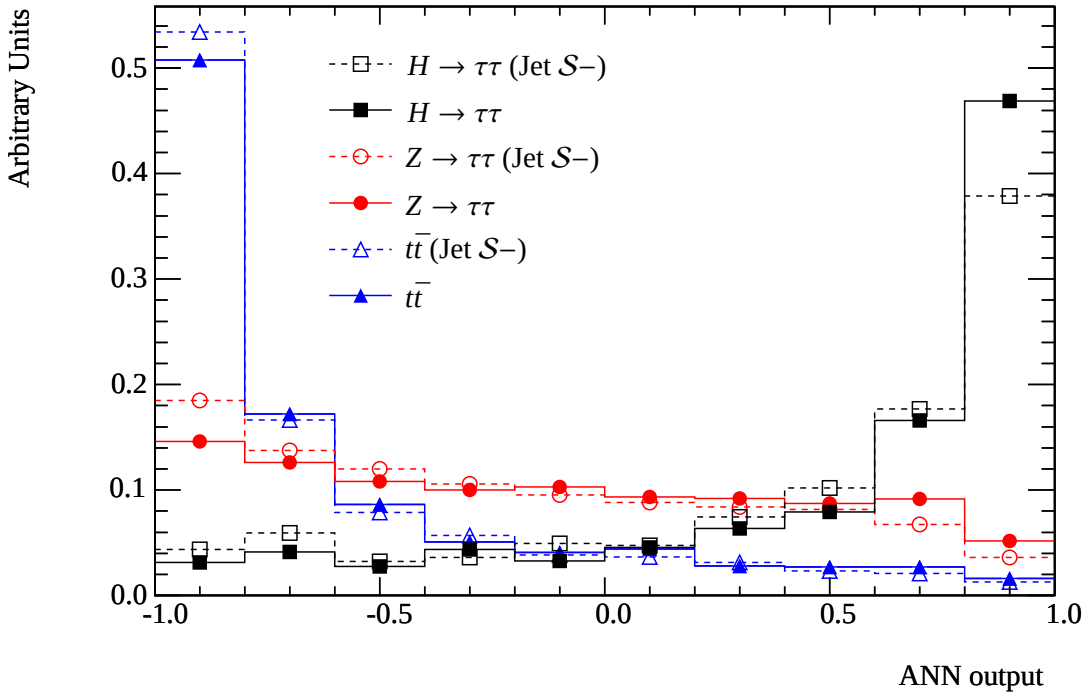


Figure 8.6: Distribution of the ANN output value for the signal and the two most important background processes without (full lines) and with a systematic shift of the jet energy scale towards lower values (dashed lines).

Table 8.4: Comparison of the cut efficiencies [%] with and without a shift of the jet energy scale towards lower values (Jet $\mathcal{S}-$) for the signal and the $Z \to \tau\tau$ and $t\bar{t}$ backgrounds in the leptonic channel. The discriminating variables which are affected the most are indicated by the shaded lines.

	$H \to \tau\tau$		$Z \to \tau\tau$ + jets		$t\bar{t}$	
Jet $\mathcal{S}-$ applied?	No	Yes	No	Yes	No	Yes
Trigger	53	53	49	49	29	29
N($e+\mu$)=2	47	47	58	58	13	13
Lepton charge	100	100	100	100	98	98
$E_{\mathrm{T}}^{\mathrm{miss}} > 40\,\mathrm{GeV}$	58	55	45	42	78	77
$\Delta\phi_{\tau\tau} < 2.5$	90	91	94	94	70	70
$\chi_{1,2} = 0.0 - 1.0$	90	89	85	84	16	15
$m_{\mathrm{T}}^{\ell\nu} < 100\,\mathrm{GeV}$	100	100	100	100	82	84
N(jets)$\geq$ 2	82	79	89	85	95	92
$\eta_{j1} \times \eta_{j2} < 0$	83	84	62	62	47	46
$\Delta\eta_{j\tau} > 0$	91	91	57	56	38	38
$\Delta\eta_{jj} > 3.5$	83	83	50	48	34	34
$m_{jj} > 600\,\mathrm{GeV}$	82	79	44	38	68	65
$\Delta\phi_{jj} < 2.2$	82	82	65	64	70	70
b-jet veto	89	90	87	86	76	78
$\Delta\eta_{j\tau}^{\mathrm{min}} > 0.6$	93	93	70	71	74	75
$\Delta\eta_{j\tau}^{\mathrm{max}} > 1.5$	99	100	100	100	95	94
Central jet veto	85	87	86	89	9	10
$p_{\mathrm{T}}^{\mathrm{tot}} < 30\,\mathrm{GeV}$	93	88	90	85	52	40

Table 8.5: Impact of the detector-related systematic effects (in %) on the selected cross-section for signal and backgrounds in the leptonic channel in the optimized cut-based analysis, in the ANN analysis and, for comparison, in the baseline analysis [49]. $\mathcal{S}+$ and $\mathcal{S}-$ denote the positive and negative shifts of the energy scale, $\mathcal{R}$ the degradation of the energy resolution and $\mathcal{E}$ the degradation of the reconstruction efficiency.

$\ell\ell$-channel		cross-section w/o sys. effects [fb]	Systematic effects										
			Electrons				Muons				Jets		
			$\mathcal{S}+$	$\mathcal{S}-$	$\mathcal{R}$	$\mathcal{E}$	$\mathcal{S}+$	$\mathcal{S}-$	$\mathcal{R}$	$\mathcal{E}$	$\mathcal{S}+$	$\mathcal{S}-$	$\mathcal{R}$
Baseline analysis	$Z \to \tau\tau$	1.69 ± 0.057	+1	±0	±0	±0	±0	±0	±0	−1	+27	−25	−1
	$t\bar{t}$	4.7 ± 0.3	−3	−3	−3	−3	−4	−2	−3	−3	+9	−10	−4
	H (105 GeV)	0.61 ± 0.03	±0	−1	±0	±0	±0	±0	±0	±0	+13	−15	+1
	H (110 GeV)	0.59 ± 0.02	±0	±0	±0	±0	±0	±0	−1	−2	+13	−13	+1
	H (115 GeV)	0.62 ± 0.02	±0	±0	±0	±0	±0	±0	±0	−2	+9	−16	±0
	H (120 GeV)	0.54 ± 0.02	±0	±0	±0	±0	−1	+1	±0	−1	+15	−16	+2
	H (125 GeV)	0.49 ± 0.02	±0	±0	±0	±0	−1	+1	−1	−1	+11	−16	±0
	H (130 GeV)	0.44 ± 0.02	±0	±0	−1	±0	±0	±0	±0	−2	+11	−14	−2
	H (135 GeV)	0.37 ± 0.01	±0	+1	±0	±0	−1	+1	±0	−1	+13	−18	±0
Optimized cuts	$Z \to \tau\tau$	1.25 ± 0.05	+1	±0	±0	±0	±0	±0	±0	−1	+24	−28	−2
	$t\bar{t}$	1.49 ± 0.16	−5	−5	−3	−5	−6	−5	−6	−6	+13	−29	−7
	H (105 GeV)	0.65 ± 0.03	±0	−1	±0	±0	±0	±0	±0	−1	+5	−16	−2
	H (110 GeV)	0.62 ± 0.02	±0	−1	−1	±0	−1	±0	±0	−2	+6	−15	±0
	H (115 GeV)	0.63 ± 0.02	±0	±0	±0	±0	±0	±0	±0	−1	+6	−17	−1
	H (120 GeV)	0.58 ± 0.02	±0	±0	±0	±0	±0	±0	±0	−1	+3	−15	+1
	H (125 GeV)	0.53 ± 0.02	±0	±0	±0	±0	±0	±0	±0	−1	+3	−18	−2
	H (130 GeV)	0.44 ± 0.02	±0	±0	±0	±0	±0	±0	±0	−2	+4	−16	−1
	H (135 GeV)	0.38 ± 0.01	±0	+1	±0	±0	±0	±0	±0	−1	+4	−16	±0
ANN analysis	$Z \to \tau\tau$	1.25 ± 0.05	±0	±0	−1	±0	−1	±0	±0	−1	+19	−32	−2
	$t\bar{t}$	1.0 ± 0.15	±0	±0	±0	−1	±0	−1	−3	−3	+14	−16	−11
	H (105 GeV)	0.74 ± 0.03	±0	−1	−1	−1	±0	±0	±0	−1	+4	−22	−2
	H (110 GeV)	0.73 ± 0.03	±0	±0	±0	±0	−1	±0	−1	−2	+2	−21	−1
	H (115 GeV)	0.73 ± 0.03	±0	±0	−1	±0	±0	±0	±0	−1	±0	−23	−3
	H (120 GeV)	0.67 ± 0.02	±0	±0	±0	±0	±0	±0	±0	−1	+2	−20	−1
	H (125 GeV)	0.60 ± 0.02	±0	±0	±0	±0	±0	±0	−1	−1	+2	−21	−2
	H (130 GeV)	0.53 ± 0.02	±0	±0	±0	±0	±0	±0	±0	−1	+1	−23	−5
	H (135 GeV)	0.44 ± 0.01	±0	+1	±0	±0	±0	±0	±0	−1	+5	−23	±0

Table 8.6: Impact of the detector-related systematic effects (in %) on the selected cross-section for signal and backgrounds in the semileptonic channel in the optimized cut-based analysis, in the ANN analysis and, for comparison, in the baseline analysis [49]. $\mathcal{S}+$ and $\mathcal{S}-$ denote the positive and negative shifts of the energy scale, $\mathcal{R}$ the degradation of the energy resolution and $\mathcal{E}$ the degradation of the reconstruction efficiency.

ℓh-channel		cross-section w/o sys. effects [fb]	Systematic effects														
			Electrons				Muons				τ-Jets				Jets		
			$\mathcal{S}+$	$\mathcal{S}-$	$\mathcal{R}$	$\mathcal{E}$	$\mathcal{S}+$	$\mathcal{S}-$	$\mathcal{R}$	$\mathcal{E}$	$\mathcal{S}+$	$\mathcal{S}-$	$\mathcal{R}$	$\mathcal{E}$	$\mathcal{S}+$	$\mathcal{S}-$	$\mathcal{R}$
Baseline analysis	$Z \to \tau\tau$	1.63 ± 0.11	±0	±0	+1	±0	−1	±0	±0	±0	+2	−7	−2	−5	+11	−16	−1
	$t\bar{t}$	0.28 ± 0.07	±0	+5	+5	±0	+5	±0	+5	±0	+9	±0	+8	−7	−2	+5	−14
	H (105 GeV)	0.74 ± 0.04	−1	±0	±0	−1	±0	+1	±0	±0	+3	−1	±0	−6	+11	−11	−3
	H (110 GeV)	0.73 ± 0.04	±0	+1	+1	±0	±0	+1	±0	−1	−2	−5	−2	−5	+9	−9	−2
	H (115 GeV)	0.78 ± 0.07	−1	−1	±0	±0	−1	±0	−1	−1	−1	−1	−7	−5	+7	−7	−1
	H (120 GeV)	0.67 ± 0.04	−1	+1	−1	±0	±0	±0	−1	±0	−1	−3	+2	−5	+11	−13	−2
	H (125 GeV)	0.58 ± 0.03	+1	+1	±0	±0	±0	±0	±0	±0	−2	+1	−3	−5	+14	−8	+2
	H (130 GeV)	0.53 ± 0.05	±0	±0	±0	±0	±0	+1	+1	±0	+3	−3	−2	−5	+10	−10	−2
	H (135 GeV)	0.45 ± 0.02	±0	±0	+1	±0	−1	±0	−1	±0	+5	±0	±0	−5	+9	−12	±0
Optimized cuts	$Z \to \tau\tau$	1.38 ± 0.10	±0	−1	+1	±0	±0	±0	±0	±0	+1	−2	+2	−6	+12	−20	−1
	$t\bar{t}$	0.14 ± 0.05	±0	±0	±0	±0	±0	±0	±0	±0	±0	±0	±0	±0	−10	−20	+10
	H (105 GeV)	0.75 ± 0.05	−1	±0	±0	−1	−1	±0	±0	±0	+1	±0	±0	−6	+5	−11	−3
	H (110 GeV)	0.73 ± 0.04	±0	+1	+1	±0	±0	+1	±0	−1	−2	−5	−3	−5	+5	−6	−2
	H (115 GeV)	0.75 ± 0.07	±0	−1	±0	±0	−1	±0	−1	−1	−2	±0	−6	−7	+5	−4	+1
	H (120 GeV)	0.67 ± 0.04	−1	±0	±0	±0	±0	+1	−1	±0	−2	−1	+3	−5	+4	−15	−3
	H (125 GeV)	0.59 ± 0.03	±0	±0	±0	±0	±0	±0	±0	±0	−1	±0	−3	−5	+4	−9	+2
	H (130 GeV)	0.55 ± 0.05	±0	−1	±0	±0	±0	+1	±0	−1	−2	−4	−4	−7	+6	−7	−3
	H (135 GeV)	0.44 ± 0.02	±0	±0	±0	±0	−1	±0	±0	±0	+4	±0	±0	−6	+4	−7	−1
ANN analysis	$Z \to \tau\tau$	1.4 ± 0.09	±0	−1	−1	±0	±0	−1	−1	±0	+5	−14	−4	−6	−5	−26	−1
	$t\bar{t}$	0.15 ± 0.05	±0	±0	±0	±0	±0	±0	±0	±0	±0	±0	+9	±0	+27	−36	−18
	H (105 GeV)	0.81 ± 0.05	±0	+1	+1	−1	−1	±0	±0	±0	+1	−2	+2	−6	−2	−15	±0
	H (110 GeV)	0.81 ± 0.04	±0	+1	+1	±0	±0	+1	±0	−1	−5	−8	−3	−5	±0	−16	−2
	H (115 GeV)	0.80 ± 0.07	+1	+1	+1	±0	±0	+1	±0	−1	−5	−4	−9	−7	−8	−8	−3
	H (120 GeV)	0.71 ± 0.04	−1	±0	±0	±0	±0	±0	−1	±0	−3	−3	+5	−5	+3	−17	−1
	H (125 GeV)	0.68 ± 0.03	±0	±0	±0	±0	±0	±0	±0	±0	−4	−3	−5	−5	+2	−16	±0
	H (130 GeV)	0.64 ± 0.05	±0	±0	+1	±0	±0	+1	±0	−1	+3	−11	−3	−6	+2	−17	−5
	H (135 GeV)	0.50 ± 0.02	±0	±0	±0	±0	−1	±0	±0	±0	+2	−4	−1	−5	+2	−18	−3

a) Without systematic uncertainties on the signal and background predictions

In the ideal case that the background is perfectly known, the signal significance including systematic effects can be evaluated using equation 8.6 with the number of signal and background events being coherently varied according to the detector-related systematic effects summarized in Tables 8.5 and 8.6:

$$\begin{aligned} S &\rightarrow S' = S + \delta_{S\,i} S \quad \text{and} \\ B &\rightarrow B' = B + \delta_{Bi} B, \end{aligned} \tag{8.12}$$

where $\delta_{S\,i}$ (δ_{Bi}) is the relative change of the selected signal (background) cross-section due to the i-th systematic effect.

Figure 8.7 shows the resulting signal significances for the optimized cut-based, the ANN and the baseline analysis separately for the leptonic and the semileptonic channel. The line shows the significance without systematic variations. The shaded band represents the maximum deviation arising from detector-related systematic variations where the shift of the jet energy scale dominates.

b) Background estimation from simulation including systematic uncertainties

Assuming no correlation between the different systematic effects, the total systematic background uncertainty is conservatively given by $\delta_{B,max} = \sqrt{\sum_i \delta_{B,i}^2}$, with $\delta_{B,i}$ being the i-th systematic effect, including:

1. the theoretical uncertainties of 10% (see Section 8.2.1)
2. the uncertainty of 50 % (100 %) on the $t\bar{t}$ background in the leptonic (semileptonic) channel due to the fast detector simulation (see Section 8.2.2)
3. an uncertainty on the W + jets background contribution in the semileptonic channel (see Section 6.6.2) which is assumed to have a similar invariant $\tau\tau$ mass distribution as the $t\bar{t}$ background. This uncertainty has thus been included by adding a 100 % uncertainty on the $t\bar{t}$ contribution.
4. the detector-related effects in Tables 8.5 and 8.6.

The values $\delta_{B,max}$ are given in Table 8.6. Taking the uncertainties into account, equation 8.6 changes to:

$$\frac{S}{\sqrt{N}} \rightarrow \frac{N' - B}{\sqrt{N' + (\delta_{max} \cdot B)^2}} = \frac{S' + B' - B}{\sqrt{N' + (\delta_{max} \cdot B)^2}} \tag{8.13}$$

where $N' = S' + B'$ be the total number of observed events *including* systematic detector effects and B be the number of expected background events from Monte Carlo data *without* any systematic effect.

Figure 8.8 shows the predicted signal significances with and without systematic uncertainties on the background estimation from Monte Carlo data. The lines represent the values

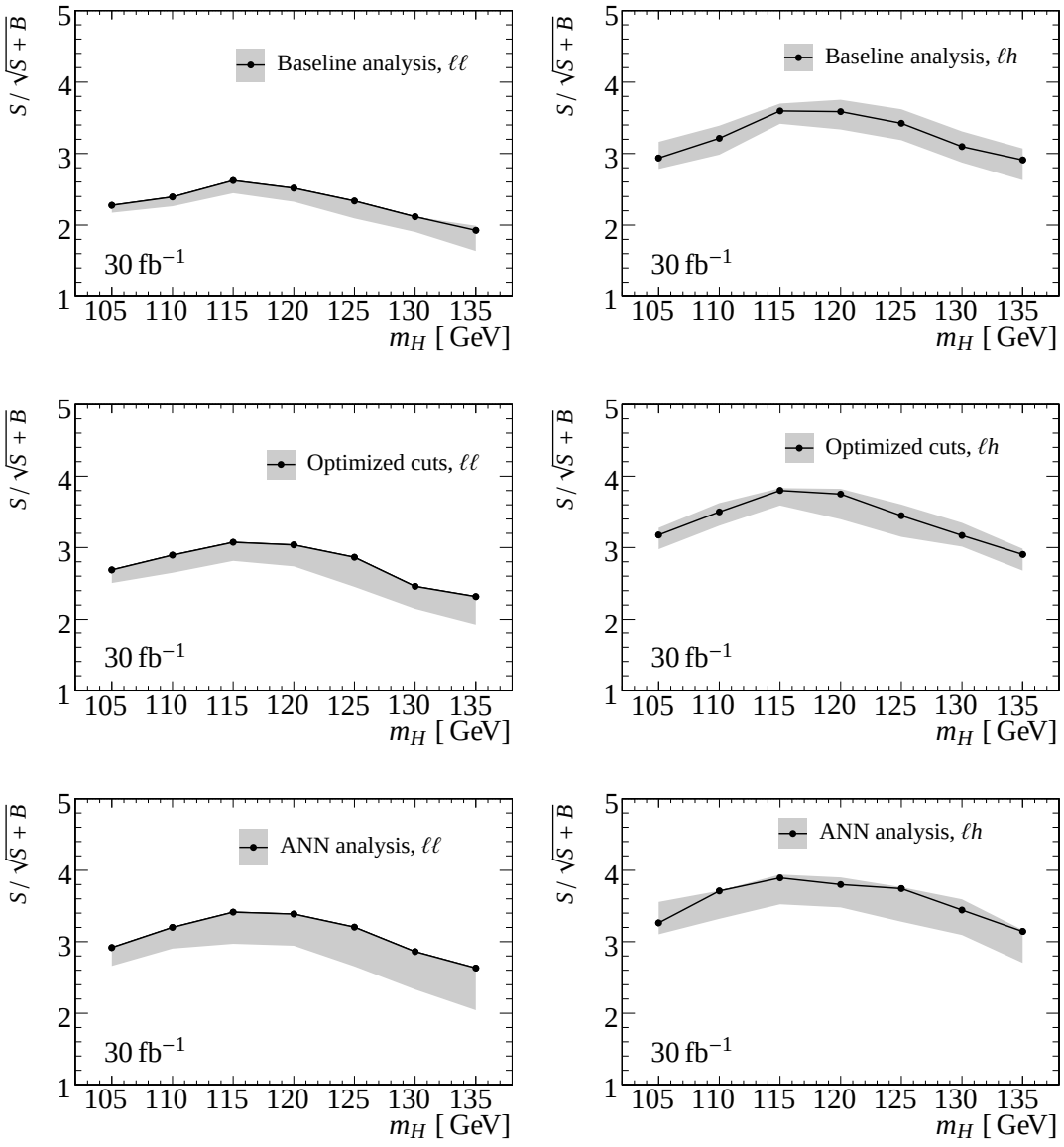


Figure 8.7: Signal significances $S/\sqrt{S+B}$ of the baseline, the optimized cut-based and the ANN analysis as a function of the Higgs boson mass for a perfectly known background contribution and an integrated luminosity of $30\,\mathrm{fb}^{-1}$. The lines represent the case of no systematic detector effects. The shaded bands represent the maximum deviation taking into account detector-related systematic effects. The results are shown separately for the leptonic (left) and the semileptonic channel (right).

Table 8.7: Systematic uncertainty $\delta_{B,max}$ [%] on the background contribution estimated by Monte Carlo simulations of the dominant background processes $Z \to \tau\tau$ and $t\bar{t}$. The W + jets background in the semileptonic channel has been neglected but a systematic uncertainty of 100 % was assigned assuming the same invariant $\tau\tau$ mass distribution as for the $t\bar{t}$ background.

$\delta_{B,max}$ [%]	$\ell\ell$ channel			ℓh channel		
	Baseline analysis	Optimized cuts	ANN analysis	Baseline analysis	Optimized cuts	ANN analysis
$Z \to \tau\tau$	30	30	35	20	20	30
$t\bar{t}$	55	60	60	100	100	110
W + jets		neglected		100	100	100

without variations in the number of signal and background events caused by detector-related systematic effects. A degradation of the signal significance due to the systematic uncertainty $\delta_{B,max}$ is observed.

The corresponding uncertainty band due to detector-effects becomes much broader since the term $B' - B$ in the numerator of equation 8.13 strongly influences the calculated number of signal events. This is most pronounced for low Higgs masses where the vicinity of the Z-resonance leads to a larger number of background events than for higher Higgs masses.

c) Background estimation from control data samples

In order to avoid large systematic uncertainties in the estimation on the background contribution, the background has to be determined using real data rather than Monte Carlo data. This can be done by means of control data samples which are selected free of signal contributions to provide a reliable and unbiased background determination. With the relative uncertainty of such a data-driven background estimation being σ, the estimation of the signal significance becomes

$$\frac{N' - B'}{\sqrt{N' + (\sigma \cdot B')^2}} = \frac{S'}{\sqrt{N' + (\sigma \cdot B')^2}} \tag{8.14}$$

with B' being the number of background events measured with the data-driven method taking into account the variations in the selected cross-section caused by systematic effects.

One can determine the $Z \to \tau\tau$ background contribution from reconstructed $Z \to \mu\mu$ events by replacing the muons with simulated τ decays. In this way, deviations of the background contribution due to systematic effects are directly taken into account. Studies have shown that with this method one can estimate the $Z \to \tau\tau$ background with an uncertainty of $\sigma = 10\,\%$ [78, 49].

A method for determining the $t\bar{t}$ background from real data has not yet been developed. However, it is expected that the $t\bar{t}$ cross-section measurement at the LHC will reduce the present uncertainty on the Monte Carlo simulation. Thus, a rather conservative uncertainty of 50 % is assigned to this background contribution as it has been done in the CSC analysis [49].

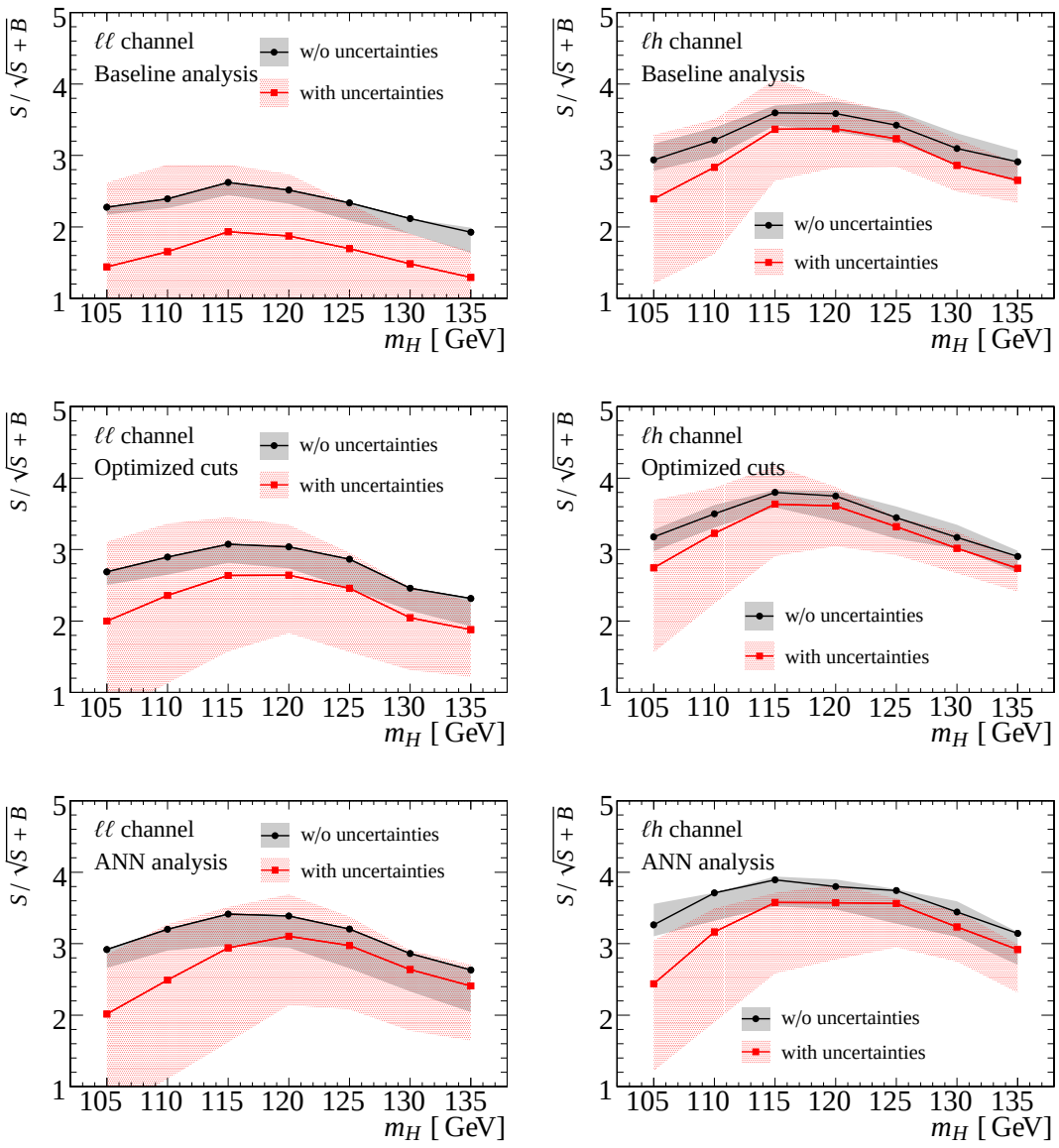


Figure 8.8: Signal significances $S/\sqrt{S+B}$ with (squares) and without (circles, see Figure 8.7) systematic uncertainties on the background estimation from Monte Carlo simulations as a function of the Higgs mass, calculated for an integrated luminosity of 30 fb^{-1}. The shaded bands indicate the variations of the predicted significance due to detector-related systematic effects in both cases.

Figure 8.9 shows the resulting predictions of the signal significance for the leptonic and semileptonic channels and the analysis methods studied. The lines represent the significance predictions without systematic variations of signal and background which are in turn indicated by the shaded bands.

Compared to the background estimation from Monte Carlo simulations (Figure 8.8), the estimation of the background contribution by data-driven methods provides higher signal significances and much narrower uncertainty bands. Thus, methods for a reliable estimation of the background from experimental data are crucial for a Higgs boson discovery in the $H \to \tau\tau$ channel.

8.2.5 Profile Likelihood Calculation

For the case of data-driven background estimation, the signal significances have also been calculated by means of the profile likelihood method (see Section 8.1) in the same manner as it was done in the CSC analysis [49]. A simultaneous fit to the $m_{\tau\tau}$ distributions of the signal candidates as well as to the $Z \to \tau\tau$ and the $t\bar{t}$ background estimations from data is performed with the systematic uncertainties of 10 % and 50 % on the $Z \to \tau\tau$ and $t\bar{t}$ backgrounds being taken into account in this calculation.

The probability density functions $F(m_{\tau\tau};\boldsymbol{\theta})$ for the signal and each background process have been determined from fits to the corresponding $m_{\tau\tau}$ distributions after all analysis cuts. Both the $Z \to \tau\tau$ and the signal distribution are described by the sum of three Gaussians with identical mean and fixed ratios of the standard deviations ($\sigma_2 = 2.45 \cdot \sigma_1$, $\sigma_3 = 5.02 \cdot \sigma_1$) and normalizations ($\mathcal{N}_2 = 0.635 \cdot \mathcal{N}_1$, $\mathcal{N}_3 = 0.292 \cdot \mathcal{N}_1$). In addition, an efficiency weighting function is used to take into account the asymmetry of the $Z \to \tau\tau$ and $H \to \tau\tau$ resonances which is introduced by the cuts on the visible momentum fractions $\chi_{1,2}$ and on the missing transverse energy $E_{\mathrm{T}}^{\mathrm{miss}}$. It is parameterized as $1/2 + 1/2 \cdot \mathrm{erf}\{[m_{\tau\tau} - m_{\mathrm{vis}}]/\sqrt{2}\sigma_{\mathrm{vis}}\}$ where "erf" denotes the Gaussian error function. m_{vis} and σ_{vis} represent the mean and standard deviation of the distribution of the events rejected by the $\chi_{1,2}$ cuts, where m_{vis} is parameterized as $m_{\mathrm{vis}} = 0.576\, m_{Z/H} + 6.2\,\mathrm{GeV}$ and σ_{vis} is fixed to 10 GeV.

Thus, the pdfs are:

$$F_{Z/H}(m_{\tau\tau}; m_{Z/H}, \sigma_{Z/H}) = c_{n1}\left[\frac{1}{2} + \frac{1}{2}\cdot \mathrm{erf}\left(\frac{m_{\tau\tau} - m_{\mathrm{vis}}}{\sqrt{2}\sigma_{\mathrm{vis}}}\right)\right] \times \sum_{i=1}^{3} \mathcal{N}_i\, \mathrm{e}^{(m_{\tau\tau} - m_{Z/H})^2/2\sigma^2_{Z/H,i}} \qquad (8.15)$$

The shape of the $m_{\tau\tau}$ distribution for the $t\bar{t}$ background is parameterized by

$$F_{t\bar{t}}(m_{\tau\tau}; a_1, a_2, a_3) = c_{n2}\left(\frac{1}{m_{\tau\tau} + a_1}\right)^{a_2} m_{\tau\tau}^{a_3} \qquad (8.16)$$

where the parameters c_{n1} and c_{n2} ensure the normalization of the pdfs.

The pdf used for the fit to the signal candidates distribution is the sum of the pdfs of the three contributions (signal, $Z \to \tau\tau$ and $t\bar{t}$)

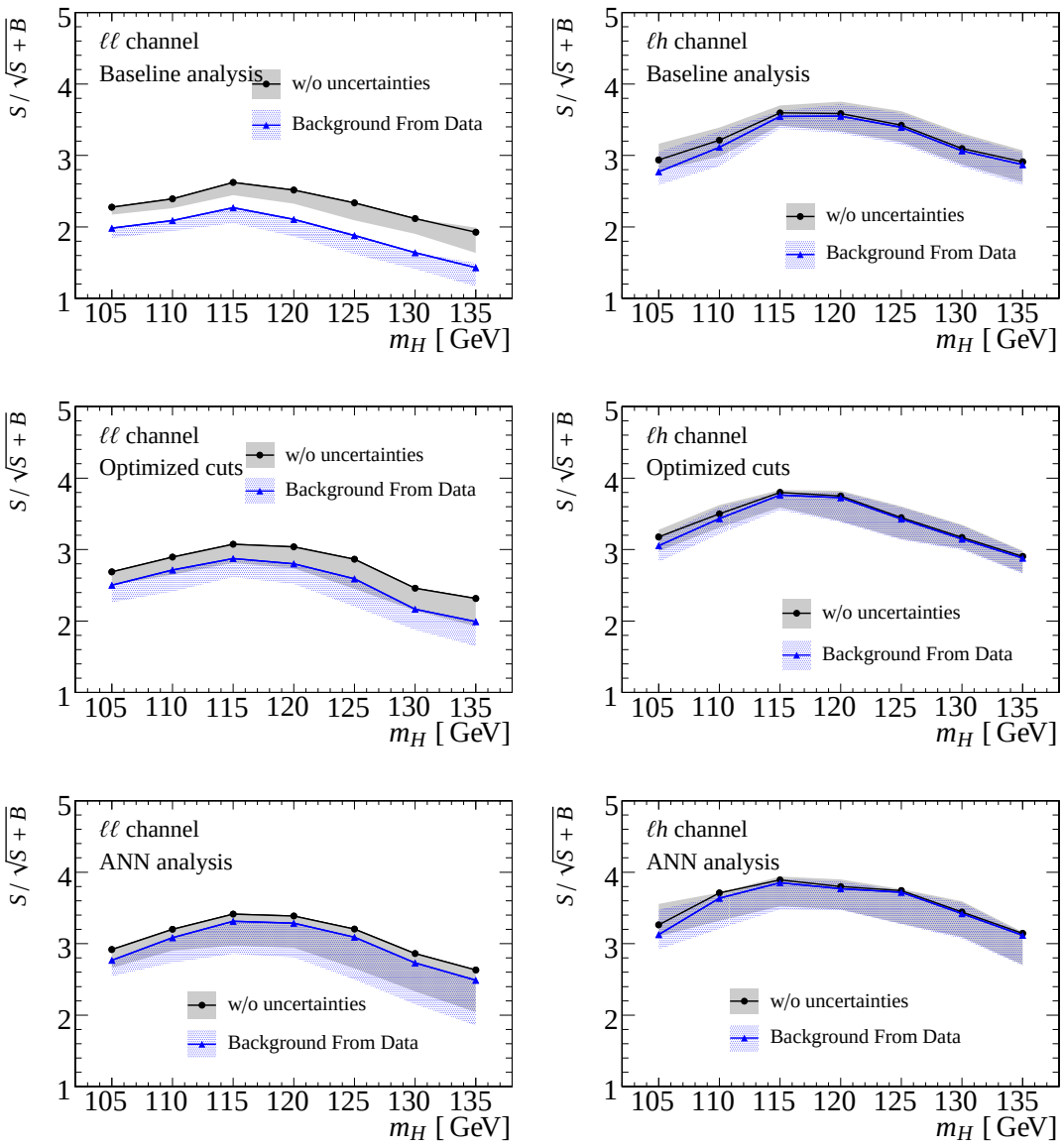


Figure 8.9: Signal significances $S/\sqrt{S+B}$ for background estimation from data (triangles) compared to the case without systematic uncertainties (circles, see Figure 8.7), calculated for an integrated luminosity of 30 fb^{-1}. The bands indicate the systematic variations in the signal significance prediction due to detector-related uncertainties.

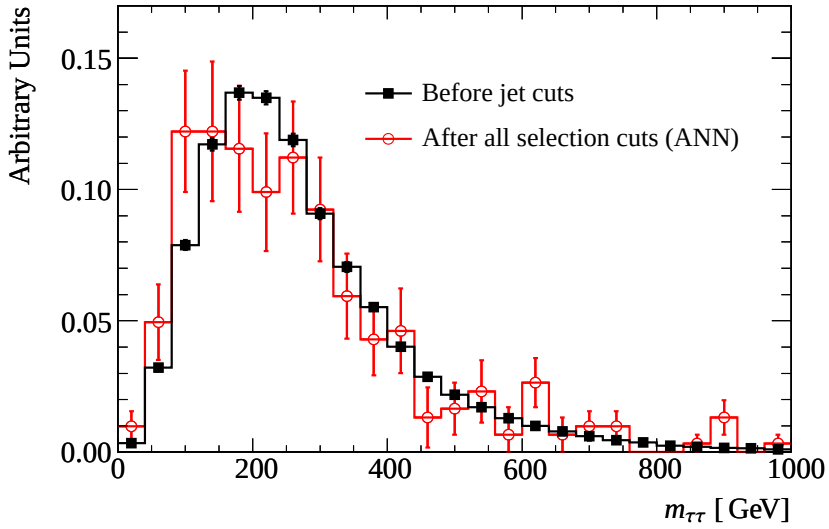


Figure 8.10: $m_{\tau\tau}$ distributions of the $t\bar{t}$ background in the leptonic analysis, shown before applying jet cuts (full squares) and after applying all selection cuts of the ANN analysis (open circles).

$$\begin{aligned} F_{\text{signal}}(m_{\tau\tau};\, f_H, m_H, \sigma_H,\, f_Z, m_Z, \sigma_Z, a_1, a_2, a_3) &= f_H \cdot F_H(m_{\tau\tau}; m_H, \sigma_H) \\ &+ f_Z \cdot F_Z(m_{\tau\tau}; m_Z, \sigma_Z) \\ &+ (1 - f_H + f_Z) \cdot F_{t\bar{t}}(m_{\tau\tau}; a_1, a_2, a_3) \end{aligned} \tag{8.17}$$

where f_H and f_Z are the relative contributions of the signal and the $Z \to \tau\tau$ processes. The total likelihood is calculated by multiplying the likelihoods of the three fits:

$$\begin{aligned} L(m_{\tau\tau}|\, f_H, \nu) &= L_{Z\ \text{control}}(m_{\tau\tau}|m_Z, \sigma_Z) \\ &\times L_{t\bar{t}\ \text{control}}(m_{\tau\tau}|a_1, a_2, a_3) \\ &\times L_{\text{signal}}(m_{\tau\tau}|\, f_H, m_H, \sigma_H,\, f_Z, m_Z, \sigma_Z, a_1, a_2, a_3). \end{aligned} \tag{8.18}$$

To test the H_0 hypothesis, the signal contribution f_H is fixed to 0, which corresponds to $\mu = 0$ as described in Section 8.1. The other parameters $(m_H, \sigma_H,\, f_Z, m_Z, \sigma_Z, a_1, a_2, a_3)$ are left free in the fit and are called *nuisance parameters*, denoted by ν.

For the $Z \to \tau\tau$ and the $t\bar{t}$ control distributions the same Monte Carlo data and the same selection criteria as for the signal selection have been used to simulate the data-driven methods with systematic uncertainties on the normalization of the control distribution of 10 % for $Z \to \tau\tau$ and 50 % for $t\bar{t}$ as discussed above. Due to the low Monte Carlo statistics after applying all cuts, the control distribution for the $t\bar{t}$ background has been taken before applying the cuts on forward and central jets. As indicated in Figure 8.10 the influence of the jet cuts on the $m_{\tau\tau}$ shape is expected to be small and has thus been neglected.

Figures 8.11 and 8.12 show the signal and background pdfs used for the likelihood fits in the leptonic and semileptonic channel respectively. The fits of $F_Z(m_{\tau\tau}; m_Z, \sigma_Z)$ and $F_{t\bar{t}}(m_{\tau\tau}; a_1, a_2, a_3)$ to

the corresponding control samples are shown in Figures 8.11a and 8.12a as well as Figures 8.11d and 8.12d. Figures 8.11b and 8.12b show the fit to the $m_{\tau\tau}$ distribution obtained by the ANN analysis for an integrated luminosity of 30 fb^{-1} in the leptonic decay channel. The signal strength parameter μ is fixed to 0 representing the H_0 hypothesis. Figures 8.11c and 8.12c show the corresponding fit with μ as free parameter. The resulting likelihood values are used to calculate the signal significance as described in Section 8.1. The combination of the leptonic and the semileptonic decay channel is done by multiplying the individual likelihoods.
Figure 8.13 shows the results of the significance calculations for the optimized cut-based, the ANN and the baseline analysis in the leptonic, the semileptonic and the combination of both channels. The lines represent the significances obtained without systematic variations of the detector performance and the shaded bands indicate the maximum deviations due to the systematic detector-related effects.
While both methods take into account the systematic uncertainties from estimating the background from data-driven methods, the significances obtained by the profile likelihood method are higher than the $S/\sqrt{S+B}$ values calculated in the $m_{\tau\tau}$ mass window (see Figure 8.9). This can be explained by the fact that the profile likelihood method uses the information of the full $m_{\tau\tau}$ distribution.
Comparing the three analysis methods, one observes that the significances resulting from the optimized cut-based analysis are significantly higher than from the baseline analysis. Furthermore, the significances resulting from the ANN analysis are higher than those provided by the optimized cut-based analysis. Table 8.8 summarizes the signal significances calculated by the profile likelihood method with uncertainties corresponding to the maximum variations caused by detector-related effects. Uncertainties due to limited Monte Carlo statistics have not been taken into account.
With the optimized cut-based analysis, it will be possible to discover the Standard Model Higgs boson with $\geq 5\,\sigma$ significance after three years of ATLAS operation corresponding to an integrated luminosity of 30 fb^{-1}, for Higgs masses above 113 GeV. The ANN analysis allows for a discovery for Higgs masses even above 110 GeV. It has been demonstrated that the ANN is robust against the expected detector-related systematic uncertainties.

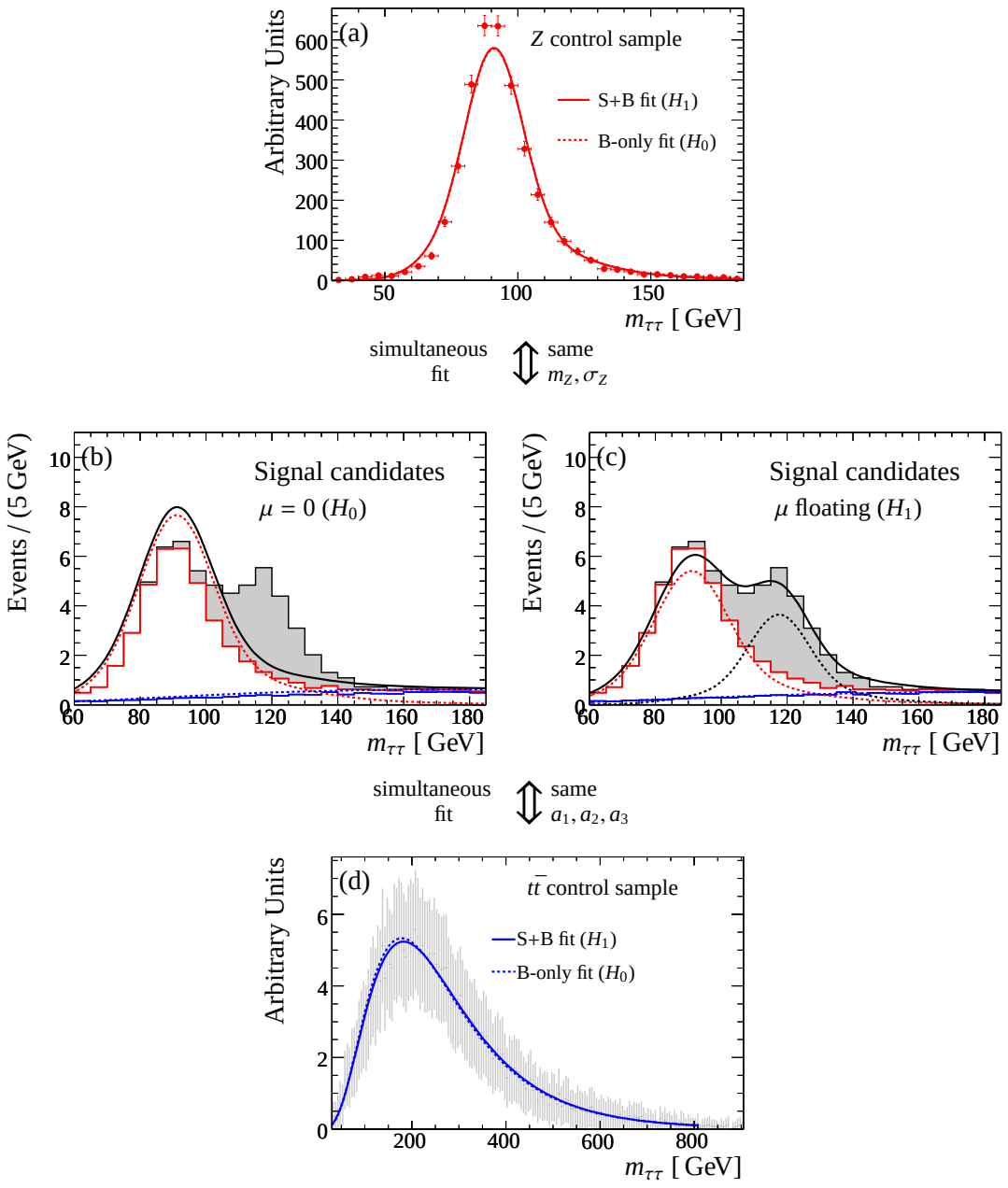


Figure 8.11: Maximum likelihood fit to the signal and background pdfs, shown (a) for the $Z \to \tau\tau$ control sample, (b) and (c) for the $m_{\tau\tau}$ distribution resulting from the ANN analysis and (d) for the $t\bar{t}$ control sample for the leptonic channel. In (b) the signal contribution was set to $\mu = 0$ (H_0 hypothesis) and in (c) μ was left floating (H_1 hypothesis). The systematic uncertainties of estimating the backgrounds from data are taken into account by scaling the error bars for $Z \to \tau\tau$ to 10 % (a) and for $t\bar{t}$ to 50 % (d). The errors for (b) and (c) are as shown in Figure 7.9.

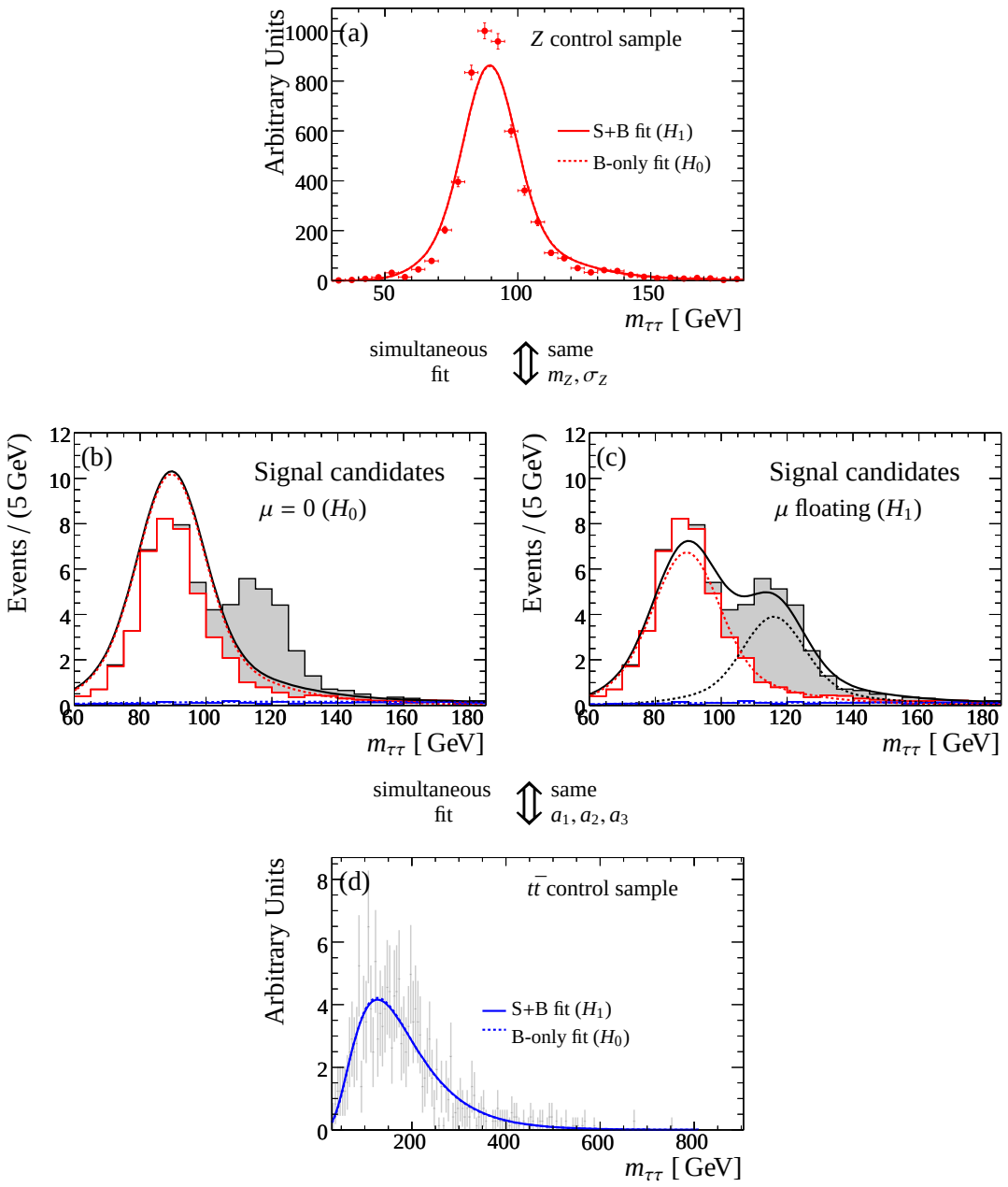


Figure 8.12: Maximum likelihood fit to the signal and background pdfs, shown (a) for the $Z \to \tau\tau$ control sample, (b) and (c) for the $m_{\tau\tau}$ distribution resulting from the ANN analysis and (d) for the $t\bar{t}$ control sample for the semileptonic channel. In (b) the signal contribution was set to $\mu = 0$ (H_0 hypothesis) and in (c) μ was left floating (H_1 hypothesis). The systematic uncertainties of estimating the backgrounds from data are taken into account by scaling the error bars for $Z \to \tau\tau$ to 10 % (a) and for $t\bar{t}$ to 50 % (d). The errors for (b) and (c) are as shown in Figure 7.10.

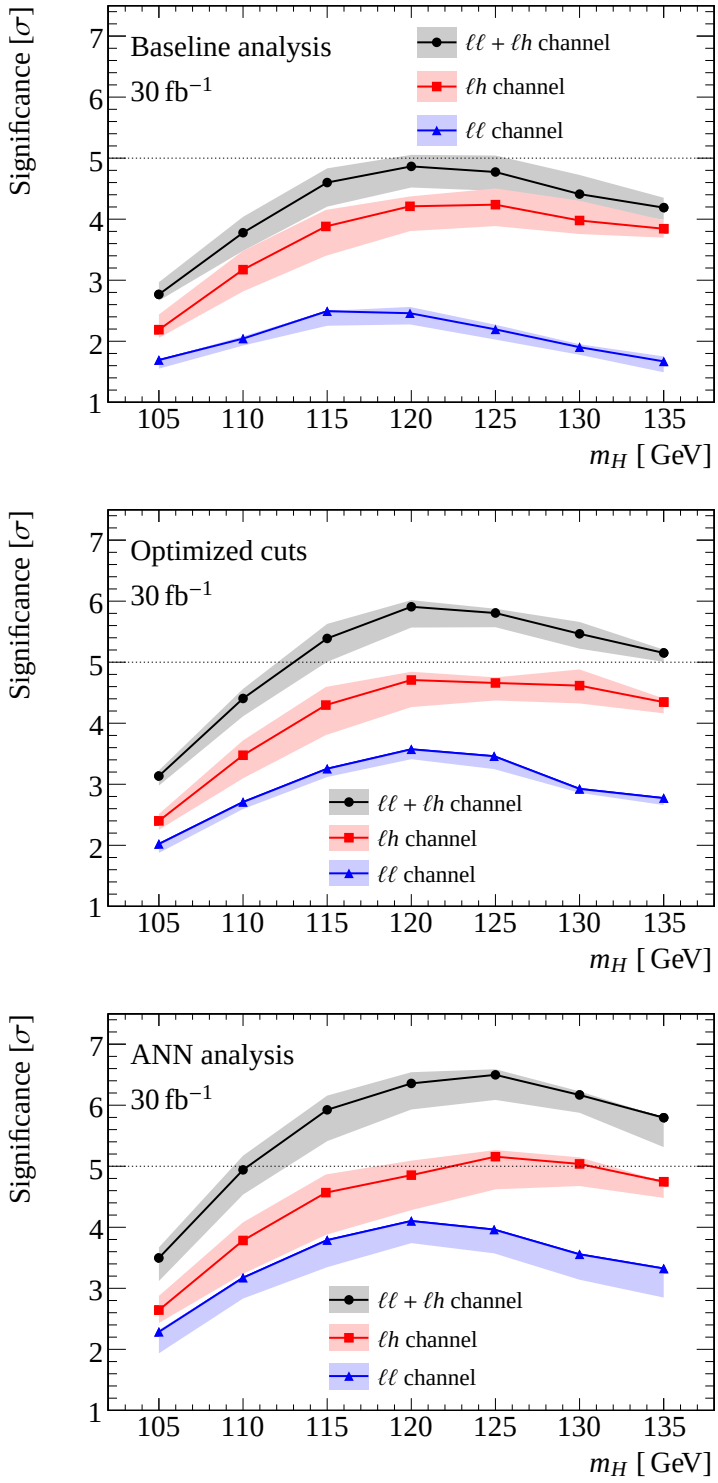


Figure 8.13: Signal significances determined by the profile likelihood method (see text) as a function of the Higgs mass, calculated for an integrated luminosity of 30 fb^{-1}. For each of the three analysis methods studied, the significances are shown for the leptonic and for the semileptonic channel as well as for the combination of the two. The lines represent the signal significance without systematic variations of the detector performance which are taken into account in the shaded bands.

Table 8.8: Expected signal significances for an integrated luminosity of $30\,\text{fb}^{-1}$ determined with the profile likelihood method for the baseline, the optimized cut-based and the ANN analysis for the leptonic and the semileptonic channel as well as the combination of the two. The uncertainties correspond to the systematic variations of the detector performance (see text).

	Baseline analysis	Optimized cuts	ANN analysis
$\ell\ell$ channel			
H (105 GeV)	$1.7^{+0.0}_{-0.1}$	$2.0^{+0.0}_{-0.1}$	$2.3^{+0.0}_{-0.3}$
H (110 GeV)	$2.0^{+0.0}_{-0.1}$	$2.7^{+0.0}_{-0.1}$	$3.2^{+0.0}_{-0.3}$
H (115 GeV)	$2.5^{+0.0}_{-0.2}$	$3.3^{+0.0}_{-0.1}$	$3.8^{+0.0}_{-0.4}$
H (120 GeV)	$2.5^{+0.1}_{-0.2}$	$3.6^{+0.0}_{-0.2}$	$4.1^{+0.0}_{-0.4}$
H (125 GeV)	$2.2^{+0.1}_{-0.2}$	$3.5^{+0.0}_{-0.2}$	$4.0^{+0.0}_{-0.4}$
H (130 GeV)	$1.9^{+0.1}_{-0.1}$	$2.9^{+0.0}_{-0.1}$	$3.6^{+0.0}_{-0.4}$
H (135 GeV)	$1.7^{+0.1}_{-0.2}$	$2.8^{+0.0}_{-0.1}$	$3.3^{+0.0}_{-0.5}$
ℓh channel			
H (105 GeV)	$2.2^{+0.2}_{-0.1}$	$2.4^{+0.1}_{-0.1}$	$2.6^{+0.2}_{-0.2}$
H (110 GeV)	$3.2^{+0.3}_{-0.4}$	$3.5^{+0.2}_{-0.4}$	$3.8^{+0.3}_{-0.5}$
H (115 GeV)	$3.9^{+0.3}_{-0.5}$	$4.3^{+0.3}_{-0.5}$	$4.6^{+0.3}_{-0.7}$
H (120 GeV)	$4.2^{+0.2}_{-0.4}$	$4.7^{+0.1}_{-0.4}$	$4.9^{+0.2}_{-0.6}$
H (125 GeV)	$4.2^{+0.3}_{-0.4}$	$4.7^{+0.1}_{-0.3}$	$5.2^{+0.1}_{-0.5}$
H (130 GeV)	$4.0^{+0.3}_{-0.2}$	$4.6^{+0.3}_{-0.3}$	$5.0^{+0.1}_{-0.4}$
H (135 GeV)	$3.8^{+0.1}_{-0.1}$	$4.3^{+0.1}_{-0.2}$	$4.7^{+0.0}_{-0.3}$
$\ell\ell + \ell h$ channel			
H (105 GeV)	$2.8^{+0.2}_{-0.1}$	$3.1^{+0.1}_{-0.2}$	$3.5^{+0.2}_{-0.4}$
H (110 GeV)	$3.8^{+0.3}_{-0.3}$	$4.4^{+0.2}_{-0.3}$	$4.9^{+0.2}_{-0.4}$
H (115 GeV)	$4.6^{+0.2}_{-0.4}$	$5.4^{+0.2}_{-0.4}$	$5.9^{+0.2}_{-0.5}$
H (120 GeV)	$4.9^{+0.2}_{-0.3}$	$5.9^{+0.1}_{-0.3}$	$6.4^{+0.2}_{-0.4}$
H (125 GeV)	$4.8^{+0.3}_{-0.3}$	$5.8^{+0.1}_{-0.2}$	$6.5^{+0.1}_{-0.4}$
H (130 GeV)	$4.4^{+0.3}_{-0.2}$	$5.5^{+0.2}_{-0.2}$	$6.2^{+0.1}_{-0.3}$
H (135 GeV)	$4.2^{+0.2}_{-0.2}$	$5.2^{+0.1}_{-0.1}$	$5.8^{+0.0}_{-0.5}$

Chapter 9

Summary

The ATLAS experiment at the Large Hadron Collider is designed to explore a wide range of new phenomena in particle physics with proton-proton collisions at energies up to 14 TeV. An integral part of the ATLAS detector is the muon spectrometer for which the Max-Planck-Institut für Physik in Munich in collaboration with the Ludwig-Maximilians University has built 88 precision drift chambers. The installation of the muon chambers in the ATLAS cavern took place in the years 2005 – 2007. Almost 1 200 chambers had to be positioned with an accuracy of 1 mm with respect to the interaction point.

Gas leak and high voltage tests have been carried out before and after installation in order to detect possible damages during installation. Afterwards, the performance of the muon spectrometer has been tested with cosmic ray muons. An average muon detection efficiency of the drift tubes of 97 % has been measured in agreement with the expectation. A part of the barrel muon chambers has been aligned with muon tracks with the desired accuracy in the regions where a sufficient number of cosmic muons could be detected.

The muon spectrometer performance has an important impact on searches for Higgs decays with muons in the final state. The discovery of the Higgs boson is the main objective of the ATLAS experiment. Experimental and theoretical bounds constrain the Standard Model Higgs boson mass to the range $114.4\,\text{GeV} \leq m_H \leq 185\,\text{GeV}$. One of the most promising discovery channel in this mass range is the vector-boson fusion production of the Higgs boson with the subsequent decay into a τ-lepton pair. The decay modes where both τ leptons decay leptonically or where one τ lepton decays leptonically and the other one hadronically have been studied. The Higgs boson mass can be reconstructed with a resolution of about 10 GeV in this channel.

Various discriminating variables can be used to suppress the dominant backgrounds from top quark pair $t\bar{t}$ and Z boson production and decay to τ leptons. The vector-boson fusion process is characterized by two highly energetic jets in the forward regions of the detector providing a particularly important signature for Higgs detection in this production process.

Two signal selection methods have been studied in this work, a cut-based analysis and a multivariate analysis method. In the case of the cut-based analysis, the discovery potential of the search channel is significantly increased compared to previously published results by using additional discriminating variables and optimizing the selection cuts by an iterative procedure. While the signal efficiency remained unchanged the background suppression improved significantly, in

particular for the $t\bar{t}$ background contribution to the leptonic channel which could be reduced by approximately 70 %.
In order to fully take into account correlations between the discriminating variables, the multivariate analysis methods Projective Likelihood, Fisher Linear Discriminant, Boosted Decision Trees and Artificial Neural Networks have been studied. The input variables have been carefully selected in order to obtain best discrimination while avoiding any bias of the reconstructed Higgs boson mass distribution.
The Artificial Neural Network algorithm showed the best performance in terms of achievable signal significance including systematic uncertainties and stability of the results with respect to the choice of discriminating variables. The background suppression of the Artificial Neural Network analysis is similar to the optimized cut-based analysis while the signal efficiency is further increased by approximately 10 % for all tested Higgs boson masses in the studied range from 105 to 135 GeV, although the network has only been trained for a Higgs mass of 120 GeV.
The impact of systematic uncertainties, in particular the simulated detector performance has been discussed. The largest systematic effect on the selection efficiencies for signal and background is the uncertainty in the jet energy calibration. Both the cut-based and the Artificial Neural Network analysis show a similar sensitivity to the systematic uncertainties. Their impact on the expected Higgs discovery potential shows the need for a reliable determination of the background contributions from the data rather than from Monte Carlo simulations.
The final results for the signal significance for the leptonic and semileptonic channel as well as for the combination of the two are derived from a maximum likelihood fit to the τ-pair invariant mass spectra for signal and background using a profile likelihood ratio method and assuming background estimation from data with uncertainties of 10 % for the $Z \to \tau\tau$ and 50 % for the $t\bar{t}$ background.
Combining both channels one can expect to discover the Higgs boson with $\geq 5\,\sigma$ significance in the mass range $115\,\text{GeV} \leq m_H \leq 135\,\text{GeV}$ and a maximum signal significance of $5.9\,\sigma$ for a Higgs mass of 120 GeV with an integrated luminosity of $30\,\text{fb}^{-1}$ corresponding to the first three years of ATLAS operation. With the Artificial Neural Network method, the mass range for $\geq 5\,\sigma$ Higgs discovery with $30\,\text{fb}^{-1}$ is extended down to 110 GeV with a maximum signal significance of $6.5\,\sigma$ for $m_H = 125\,\text{GeV}$. For both analyses, the systematic uncertainty in the prediction of the signal significance is approximately 10 %.

Appendix A

Details for all Higgs Masses

On the following pages, the evolution of the cross-sections with the applied analysis cuts are shown for all studied Higgs mass datasets for the cut-based and the multivariate analysis and for both the leptonic and the semileptonic decay channel.

Table A.1: Evolution of the signal cross-section times efficiency [fb] in the optimized cut-based analysis, shown for all studied Higgs mass points in the leptonic decay channel (extension of Table 6.3). The uncertainties shown are statistical only.

	105 GeV	110 GeV	115 GeV	120 GeV	125 GeV	130 GeV	135 GeV
Cross-section	26	25	24	22	19	16	13
Cuts on τ-decay products:							
Trigger	12.7 ± 0.11	12.8 ± 0.11	12.5 ± 0.10	11.6 ± 0.05	10.3 ± 0.09	8.82 ± 0.07	7.32 ± 0.06
N($e + \mu$)=2	5.75 ± 0.08	5.89 ± 0.07	5.82 ± 0.07	5.48 ± 0.04	4.95 ± 0.06	4.31 ± 0.05	3.62 ± 0.04
Lepton charge	5.73 ± 0.08	5.87 ± 0.07	5.79 ± 0.07	5.46 ± 0.04	4.93 ± 0.06	4.30 ± 0.05	3.61 ± 0.04
$E_{\mathrm{T}}^{\mathrm{miss}} > 40\,\mathrm{GeV}$	3.41 ± 0.06	3.49 ± 0.06	3.45 ± 0.05	3.17 ± 0.03	2.93 ± 0.05	2.55 ± 0.04	2.17 ± 0.03
$\Delta\phi_{\tau\tau} < 2.5$	3.22 ± 0.06	3.22 ± 0.05	3.16 ± 0.05	2.85 ± 0.03	2.61 ± 0.04	2.21 ± 0.04	1.86 ± 0.03
$\chi_{1,2} = 0.0 - 1.0$	2.85 ± 0.05	2.87 ± 0.05	2.84 ± 0.05	2.57 ± 0.03	2.33 ± 0.04	2.01 ± 0.04	1.69 ± 0.03
$m_{\mathrm{T}}^{\ell\nu} < 100\,\mathrm{GeV}$	2.85 ± 0.05	2.87 ± 0.05	2.84 ± 0.05	2.57 ± 0.03	2.32 ± 0.04	1.99 ± 0.04	1.68 ± 0.03
Cuts on the tagging jets:							
N(jets)$\geq$ 2	2.37 ± 0.05	2.35 ± 0.05	2.33 ± 0.05	2.10 ± 0.02	1.91 ± 0.04	1.64 ± 0.03	1.38 ± 0.03
$\eta_{j1} \times \eta_{j2} < 0$	1.94 ± 0.04	1.94 ± 0.04	1.94 ± 0.04	1.74 ± 0.02	1.60 ± 0.04	1.38 ± 0.03	1.14 ± 0.02
$\Delta\eta_{j\tau} > 0$	1.79 ± 0.04	1.80 ± 0.04	1.77 ± 0.04	1.58 ± 0.02	1.44 ± 0.03	1.24 ± 0.03	1.04 ± 0.02
$\Delta\eta_{jj} > 3.5$	1.48 ± 0.04	1.48 ± 0.04	1.48 ± 0.04	1.32 ± 0.02	1.20 ± 0.03	1.06 ± 0.03	0.89 ± 0.02
$m_{jj} > 600\,\mathrm{GeV}$	1.22 ± 0.04	1.18 ± 0.03	1.20 ± 0.03	1.08 ± 0.02	1.00 ± 0.03	0.87 ± 0.02	0.73 ± 0.02
$\Delta\phi_{jj} < 2.2$	0.99 ± 0.03	0.93 ± 0.03	0.95 ± 0.03	0.89 ± 0.03	0.81 ± 0.03	0.68 ± 0.02	0.59 ± 0.02
b-jet veto	0.88 ± 0.03	0.84 ± 0.03	0.84 ± 0.03	0.79 ± 0.03	0.72 ± 0.02	0.61 ± 0.02	0.53 ± 0.02
Cuts on the event topology:							
$\Delta\eta_{j\tau}^{\mathrm{min}} > 0.6$	0.80 ± 0.03	0.78 ± 0.03	0.78 ± 0.03	0.74 ± 0.03	0.67 ± 0.02	0.57 ± 0.02	0.49 ± 0.02
$\Delta\eta_{j\tau}^{\mathrm{max}} > 1.5$	0.80 ± 0.03	0.78 ± 0.03	0.78 ± 0.03	0.73 ± 0.03	0.66 ± 0.02	0.56 ± 0.02	0.49 ± 0.02
Central jet veto	0.70 ± 0.03	0.68 ± 0.03	0.67 ± 0.02	0.63 ± 0.02	0.57 ± 0.02	0.48 ± 0.02	0.42 ± 0.01
$p_{\mathrm{T}}^{\mathrm{tot}} < 30\,\mathrm{GeV}$	0.65 ± 0.03	0.62 ± 0.02	0.63 ± 0.02	0.58 ± 0.02	0.53 ± 0.02	0.44 ± 0.02	0.38 ± 0.01
Mass window cut:							
$m_{\tau\tau} = m_H \pm 15\,\mathrm{GeV}$	0.54 ± 0.02	0.54 ± 0.02	0.53 ± 0.02	0.49 ± 0.02	0.43 ± 0.02	0.34 ± 0.01	0.30 ± 0.01

Table A.2: Evolution of the signal cross-section times efficiency [fb] in the optimized cut-based analysis, shown for all studied Higgs mass points in the semileptonic decay channel (extension of Table 6.6). The uncertainties shown are statistical only.

	105 GeV	110 GeV	115 GeV	120 GeV	125 GeV	130 GeV	135 GeV
Cross-section	180	173	160	145	127	108	87
Cuts on τ-decay products:							
Trigger	56.4 ± 0.39	56.2 ± 0.34	53.4 ± 0.55	49.5 ± 0.23	44.7 ± 0.26	38.8 ± 0.4	31.8 ± 0.18
N($e+\mu$)=1	50 ± 0.36	49.6 ± 0.32	46.7 ± 0.51	43.4 ± 0.22	39 ± 0.24	34.1 ± 0.37	27.6 ± 0.17
N(τ)=1	8.25 ± 0.15	8.67 ± 0.13	8.67 ± 0.22	8.45 ± 0.096	7.92 ± 0.11	7.24 ± 0.17	5.99 ± 0.078
Lepton charge	7.79 ± 0.14	8.16 ± 0.13	8.24 ± 0.22	8.02 ± 0.094	7.5 ± 0.1	6.88 ± 0.17	5.73 ± 0.076
$E_{\mathrm{T}}^{\mathrm{miss}} > 30\,\mathrm{GeV}$	4.74 ± 0.11	4.98 ± 0.1	5.18 ± 0.17	4.87 ± 0.073	4.55 ± 0.082	4.2 ± 0.13	3.52 ± 0.06
$\Delta\phi_{\tau\tau} < 2.5$	4.37 ± 0.11	4.59 ± 0.097	4.62 ± 0.16	4.34 ± 0.069	3.95 ± 0.076	3.54 ± 0.12	2.92 ± 0.055
$\chi_{\ell,h} = 0.0 - 1.0$	3.3 ± 0.094	3.47 ± 0.084	3.5 ± 0.14	3.35 ± 0.061	3.07 ± 0.067	2.74 ± 0.11	2.29 ± 0.048
$m_{\mathrm{T}}^{\ell\nu} < 30\,\mathrm{GeV}$	2.49 ± 0.082	2.59 ± 0.073	2.49 ± 0.12	2.4 ± 0.051	2.14 ± 0.056	1.89 ± 0.088	1.52 ± 0.039
Cuts on the tagging jets:							
N(jets)≥ 2	2.12 ± 0.075	2.14 ± 0.066	2.1 ± 0.11	2 ± 0.047	1.71 ± 0.05	1.52 ± 0.079	1.24 ± 0.036
$\eta_{j1} \times \eta_{j2} < 0$	1.74 ± 0.068	1.81 ± 0.061	1.81 ± 0.1	1.62 ± 0.042	1.46 ± 0.046	1.26 ± 0.072	1.05 ± 0.033
$\Delta\eta_{j\tau} > 0$	1.63 ± 0.066	1.63 ± 0.058	1.68 ± 0.098	1.49 ± 0.04	1.34 ± 0.044	1.12 ± 0.068	0.968 ± 0.031
$\Delta\eta_{jj} > 3.6$	1.3 ± 0.059	1.29 ± 0.051	1.39 ± 0.089	1.2 ± 0.036	1.08 ± 0.04	0.974 ± 0.063	0.78 ± 0.028
$m_{jj} > 600\,\mathrm{GeV}$	1.11 ± 0.054	1.08 ± 0.047	1.18 ± 0.082	1.02 ± 0.033	0.887 ± 0.036	0.799 ± 0.057	0.661 ± 0.026
$\Delta\phi_{jj} < 2.7$	1.04 ± 0.053	0.996 ± 0.045	1.1 ± 0.079	0.949 ± 0.044	0.822 ± 0.035	0.734 ± 0.055	0.61 ± 0.025
Cuts on the event topology:							
$\Delta\eta_{j\tau}^{\mathrm{min}} > 0.8$	0.904 ± 0.049	0.855 ± 0.042	0.923 ± 0.072	0.831 ± 0.041	0.714 ± 0.032	0.66 ± 0.052	0.523 ± 0.023
$\Delta\eta_{j\tau}^{\mathrm{max}} > 1.5$	0.901 ± 0.049	0.853 ± 0.042	0.906 ± 0.072	0.827 ± 0.041	0.714 ± 0.032	0.656 ± 0.052	0.521 ± 0.023
Central jet veto	0.771 ± 0.045	0.751 ± 0.039	0.787 ± 0.067	0.698 ± 0.037	0.611 ± 0.03	0.566 ± 0.048	0.453 ± 0.022
$p_{\mathrm{T}}^{\mathrm{tot}} < 40\,\mathrm{GeV}$	0.755 ± 0.045	0.729 ± 0.039	0.753 ± 0.065	0.67 ± 0.037	0.593 ± 0.03	0.55 ± 0.047	0.438 ± 0.021
Mass window cut:							
$m_{\tau\tau} = m_H \pm 15\,\mathrm{GeV}$	0.648 ± 0.042	0.647 ± 0.036	0.646 ± 0.06	0.572 ± 0.034	0.485 ± 0.027	0.412 ± 0.041	0.346 ± 0.019

Table A.3: Evolution of the signal cross-section times efficiency [fb] in the ANN analysis, shown for all studied Higgs mass points in the leptonic decay channel (extension of Table 7.5). The uncertainties shown are statistical only.

	105 GeV	110 GeV	115 GeV	120 GeV	125 GeV	130 GeV	135 GeV
Cross-section [fb]	26	25	24	22	19	16	13
Trigger	12.7 ± 0.11	12.8 ± 0.11	12.5 ± 0.1	11.6 ± 0.05	10.3 ± 0.09	8.8 ± 0.07	7.3 ± 0.06
N($e + \mu$)=2	5.8 ± 0.08	5.9 ± 0.07	5.8 ± 0.07	5.5 ± 0.04	5.0 ± 0.06	4.3 ± 0.05	3.6 ± 0.04
Lepton charge	5.7 ± 0.08	5.9 ± 0.07	5.8 ± 0.07	5.5 ± 0.04	4.9 ± 0.06	4.3 ± 0.05	3.6 ± 0.04
$E_{\mathrm{T}}^{\mathrm{miss}} > 40\,\mathrm{GeV}$	3.4 ± 0.06	3.5 ± 0.06	3.5 ± 0.05	3.2 ± 0.03	2.9 ± 0.05	2.6 ± 0.04	2.2 ± 0.03
$\Delta\phi_{\tau\tau} < 2.9$	3.3 ± 0.06	3.4 ± 0.06	3.3 ± 0.05	3.1 ± 0.03	2.8 ± 0.05	2.4 ± 0.04	2.1 ± 0.03
$\chi_{1,2} = 0.0 - 1.0$	2.9 ± 0.05	3.0 ± 0.05	3.0 ± 0.05	2.7 ± 0.03	2.5 ± 0.04	2.2 ± 0.04	1.9 ± 0.03
N(jets)$\geq$ 2	2.4 ± 0.05	2.4 ± 0.05	2.4 ± 0.05	2.2 ± 0.02	2.0 ± 0.04	1.8 ± 0.03	1.5 ± 0.03
$\eta_{j1} \times \eta_{j2} < 0$	2.0 ± 0.05	2.0 ± 0.04	2.0 ± 0.04	1.8 ± 0.02	1.7 ± 0.04	1.5 ± 0.03	1.2 ± 0.02
$\Delta\eta_{j\tau} > 0$	1.8 ± 0.04	1.9 ± 0.04	1.8 ± 0.04	1.7 ± 0.02	1.5 ± 0.03	1.3 ± 0.03	1.1 ± 0.02
b-jet veto	1.6 ± 0.04	1.6 ± 0.04	1.6 ± 0.04	1.4 ± 0.03	1.3 ± 0.03	1.2 ± 0.03	0.96 ± 0.02
Central jet veto	1.3 ± 0.04	1.4 ± 0.04	1.3 ± 0.03	1.2 ± 0.03	1.1 ± 0.03	0.99 ± 0.03	0.81 ± 0.02
ANN > 0.7	0.74 ± 0.03	0.74 ± 0.03	0.73 ± 0.03	0.67 ± 0.02	0.60 ± 0.02	0.53 ± 0.02	0.44 ± 0.02
$m_{\mathrm{T}}^{\ell\nu} < 100\,\mathrm{GeV}$	0.74 ± 0.03	0.74 ± 0.03	0.73 ± 0.03	0.67 ± 0.02	0.60 ± 0.02	0.53 ± 0.02	0.44 ± 0.01
$m_{\tau\tau}$ =105-135 GeV	0.60 ± 0.02	0.62 ± 0.02	0.61 ± 0.02	0.54 ± 0.02	0.47 ± 0.02	0.39 ± 0.02	0.33 ± 0.01

Table A.4: Evolution of the signal cross-section times efficiency [fb] in the ANN analysis, shown for all studied Higgs mass points in the semileptonic decay channel (extension of Table 7.6). The uncertainties shown are statistical only.

	105 GeV	110 GeV	115 GeV	120 GeV	125 GeV	130 GeV	135 GeV
Cross-section [fb]	180	173	160	145	127	108	87
Trigger	56.4 ± 0.4	56.2 ± 0.3	53.4 ± 0.6	49.5 ± 0.2	44.7 ± 0.3	38.8 ± 0.4	31.8 ± 0.2
N($e + \mu$)=1	50.0 ± 0.4	49.6 ± 0.3	46.7 ± 0.5	43.4 ± 0.2	39.0 ± 0.2	34.1 ± 0.4	27.6 ± 0.2
N(τ)=1	8.3 ± 0.2	8.7 ± 0.1	8.7 ± 0.2	8.45 ± 0.1	7.9 ± 0.1	7.2 ± 0.2	6.0 ± 0.08
Lepton charge	7.8 ± 0.1	8.2 ± 0.1	8.2 ± 0.2	8.02 ± 0.1	7.5 ± 0.1	6.9 ± 0.2	5.7 ± 0.08
$E_{\mathrm{T}}^{\mathrm{miss}} > 30\,\mathrm{GeV}$	4.7 ± 0.1	5.0 ± 0.1	5.2 ± 0.2	4.87 ± 0.1	4.6 ± 0.08	4.2 ± 0.1	3.5 ± 0.06
$\Delta\phi_{\tau\tau} < 2.9$	4.6 ± 0.1	4.9 ± 0.1	5.0 ± 0.2	4.69 ± 0.07	4.3 ± 0.08	4.0 ± 0.1	3.3 ± 0.06
$\chi_{\ell,h} = 0.0 - 1.0$	3.4 ± 0.1	3.6 ± 0.09	3.6 ± 0.1	3.53 ± 0.06	3.3 ± 0.07	3.0 ± 0.1	2.5 ± 0.05
$m_{\mathrm{T}}^{\ell\nu} < 30\,\mathrm{GeV}$	2.6 ± 0.08	2.7 ± 0.07	2.6 ± 0.1	2.53 ± 0.05	2.3 ± 0.06	2.1 ± 0.09	1.7 ± 0.04
N(jets)$\geq$ 2	2.2 ± 0.08	2.2 ± 0.07	2.2 ± 0.1	2.08 ± 0.04	1.8 ± 0.05	1.7 ± 0.08	1.3 ± 0.04
$\eta_{j1} \times \eta_{j2} < 0$	1.8 ± 0.07	1.8 ± 0.06	1.9 ± 0.1	1.68 ± 0.04	1.5 ± 0.05	1.4 ± 0.07	1.1 ± 0.03
$\Delta\eta_{j\tau} > 0$	1.6 ± 0.07	1.7 ± 0.06	1.7 ± 0.1	1.55 ± 0.04	1.4 ± 0.05	1.2 ± 0.07	1.0 ± 0.03
Central jet veto	1.4 ± 0.06	1.4 ± 0.05	1.4 ± 0.09	1.26 ± 0.05	1.2 ± 0.04	1.0 ± 0.06	0.87 ± 0.03
ANN > 0.3	0.81 ± 0.05	0.81 ± 0.04	0.80 ± 0.07	0.71 ± 0.04	0.68 ± 0.03	0.64 ± 0.05	0.50 ± 0.02
$m_{\tau\tau}$ =105-135 GeV	0.68 ± 0.04	0.71 ± 0.04	0.69 ± 0.06	0.60 ± 0.04	0.56 ± 0.03	0.47 ± 0.04	0.39 ± 0.02

List of Figures

List of Tables

Bibliography

[1] P. W. Higgs, *Broken Symmetries and the Masses of Gauge Bosons*, Phys. Rev. Lett. **13** (1964) 508–509.

[2] Y. Nambu and G. Jona-Lasinio, *Dynamical model of elementary particles based on an analogy with superconductivity. I*, Phys. Rev. **122** (1961) 345–358.

[3] J. Goldstone, A. Salam, and S. Weinberg, *Broken Symmetries*, Phys. Rev. **127** (1962) 965–970.

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

[5] Tevatron New-Phenomena and Higgs working group, G. Bernardi et al., *Combined CDF and D0 Upper Limits on Standard Model Higgs Boson Production at High Mass (155-200 GeV/c^2) with 3 fb^{-1} of data*, FERMILAB-PUB-08-270-E, August, 2008. (arXiv:0808.0534 [hep-ex]).

[6] ALEPH collaboration et al., *Precision Electroweak Measurements and Constraints on the Standard Model*, CERN-PH-EP-2008-020, CERN, Geneva, December, 2008. http://cdsweb.cern.ch/record/1142930/.

[7] A. Pich, *The Standard Model of Electroweak Interactions*, (arXiv:0705.4264 [hep-ph]).

[8] H. Fritzsch, M. Gell-Mann, and H. Leutwyler, *Advantages of the Color Octet Gluon Picture*, Phys. Lett. **B47** (1973) 365–368.

[9] S. Weinberg, *Nonabelian Gauge Theories of the Strong Interactions*, Phys. Rev. Lett. **31** (1973) 494–497.

[10] S. Weinberg, *Current algebra and gauge theories. 2. NonAbelian gluons*, Phys. Rev. **D8** (1973) 4482–4498.

[11] D. J. Gross and F. Wilczek, *Ultraviolet Behavior of Non-Abelian Gauge Theories*, Phys. Rev. Lett. **30** (1973) 1343–1346.

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

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

[14] A. Salam, *Weak and Electromagnetic Interactions*, originally printed in "Svartholm: Elementary Particle Theory, Proceedings of the Nobel Symposium held 1968 at Lerum, Sweden", Stockholm 1968, 367-377.

[15] F. Englert and R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons*, Phys. Rev. Lett. **13** (1964) 321–322.

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

[17] UA1 Collaboration, G. Arnison et al., *Experimental observation of isolated large transverse energy electrons with associated missing energy at* $\sqrt{s} = 540$ *GeV*, Phys. Lett. **B122** (1983) 103–116.

[18] UA1 Collaboration, G. Arnison et al., *Experimental observation of lepton pairs of invariant mass around 95 GeV/*c^2 *at the CERN SPS collider*, Phys. Lett. **B126** (1983) 398–410.

[19] UA2 Collaboration, M. Banner et al., *Observation of single isolated electrons of high transverse momentum in events with missing transverse energy at the CERN* $\overline{p}p$ *collider*, Phys. Lett. **B122** (1983) 476–485.

[20] UA2 Collaboration, P. Bagnaia et al., *Evidence for* $Z^0 \to e^+e^-$ *at the CERN* $\overline{p}p$ *collider*, Phys. Lett. **B129** (1983) 130–140.

[21] K. Riesselmann, *Limitations of a standard model Higgs boson*, DESY-97-222 (1997), (arXiv:hep-ph/9711456).

[22] Particle Data Group Collaboration, C. Amsler et al., *Review of Particle Physics*, Phys. Lett. **B667** (2008) 1.

[23] For a compilation of the most accurate theoretical results for SM and MSSM Higgs cross sections at the Tevatron and the LHC see the Webpage: http://maltoni.home.cern.ch/maltoni/TeV4LHC/.

[24] O. S. Brüning et al., *LHC Design Report - v.1: the LHC Main Ring*, CERN-2004-003-V-1, CERN, Geneva, 2004. http://cdsweb.cern.ch/record/782076/.

[25] M. S. Carena and H. E. Haber, *Higgs boson theory and phenomenology*, Prog. Part. Nucl. Phys. **50** (2003) 63–152, (arXiv:hep-ph/0208209).

[26] L. Evans et al., *LHC Machine*, JINST **3** (2008) S08001.

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

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

[29] CMS Collaboration, R. Adolphi et al., *The CMS experiment at the CERN LHC*, JINST **3** (2008) S08004.

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

[31] J. C. Barriere et al., *The alignment system of the ATLAS barrel muon spectrometer*, ATLAS note, ATL-MUON-PUB-2008-007, CERN, Geneva, January, 2008. http://cdsweb.cern.ch/record/1081769/.

[32] S. Aefsky et al., *The optical alignment system of the ATLAS muon spectrometer endcaps*, JINST **3** (2008) P11005.

[33] J. Knobloch et al., *LHC computing Grid: Technical Design Report*, CERN-LHCC-2005-024, CERN, Geneva, June, 2005. http://cdsweb.cern.ch/record/840543/.

[34] J. Dubbert et al., *Integration, commissioning and installation of monitored drift tube chambers for the ATLAS barrel muon spectrometer*, Nucl. Instrum. Meth. **A572** (2007) 53–56.

[35] J. Dubbert et al., *Test, integration, commissioning and installation of large drift tube chambers of the ATLAS barrel muon spectrometer*, NSS Conference Record, IEEE 2005 **2** (2005) 1019–1023.

[36] M. Groh, *Test von Driftrohrkammern für das Myonspektrometer des ATLAS-Detektors am Large Hadron Collider*, diploma thesis, Technische Universität München and Max-Planck-Institut für Physik, January, 2006.

[37] J. v. Loeben, *Test und Kalibrierung der Präzisionsdriftrohrkammern des ATLAS Myonspektrometers*, diploma thesis, Technische Universität München and Max-Planck-Institut für Physik, December, 2006.

[38] J. Schmaler, *Test and Alignment of the ATLAS Precision Muon Chambers*, diploma thesis, Technische Universität München and Max-Planck-Institut für Physik, March, 2007.

[39] B. Bittner, *Alignment of the ATLAS Muon Spectrometer Using Muon Tracks*, diploma thesis, Technische Universität München and Max-Planck-Institut für Physik, November, 2008.

[40] I. Potrap et al., *Alignment of the ATLAS Muon Spectrometer with straight tracks using MILLEPEDE method*, ATLAS internal note, ATL-COM-MUON-2007-017, CERN, Geneva, December, 2007. http://cdsweb.cern.ch/record/1073373/.

[41] B. Bittner et al., *Alignment of the ATLAS muon spectrometer with Muon Tracks*, NSS Conference Record, IEEE 2008 (2008) 2830–2833.

[42] J. v. Loeben, *First cosmic ray results of the ATLAS barrel muon spectrometer with magnetic field*, NSS Conference Record, IEEE 2007 (2007) 1891–1894.

[43] D. Rainwater, D. Zeppenfeld, and K. Hagiwara, *Searching for $H \to \tau\tau$ in weak boson fusion at the LHC*, Phys. Rev. **D59** (1999) 014037, (arXiv:hep-ph/9808468).

[44] T. Plehn, D. L. Rainwater, and D. Zeppenfeld, *A method for identifying $H \to \tau\tau \to e^{\pm}\mu^{\mp} p_{\mathrm{T}}$ at the CERN LHC*, Phys. Rev. **D61** (2000) 093005, (arXiv:hep-ph/9911385).

[45] M. Klute, *A study of the weak boson fusion, with $H \to \tau\tau$ and $\tau \to e(\mu)\nu\nu$*, ATLAS note, ATL-PHYS-2002-018, CERN, Geneva, March, 2002. http://cdsweb.cern.ch/record/685340/.

[46] K. Cranmer et al., *An update of the VBF $H \to \tau\tau$ and $\tau \to e(\mu)\nu\nu$ Cut Analysis*, ATLAS internal note, ATL-COM-PHYS-2003-002, CERN, Geneva, January, 2003. http://cdsweb.cern.ch/record/681356/.

[47] S. Asai et al., *Prospects for the search for a standard model Higgs boson in ATLAS using vector boson fusion*, Eur. Phys. J. **C32S2** (2004) 19–54, (arXiv:hep-ph/0402254).

[48] C. Foudas, A. Nikitenko, and M. Takahashi, *Observation of the Standard Model Higgs boson via $H \to \tau\tau \to$ lepton+jet Channel*, CMS note, CMS-NOTE-2006-088, CERN, Geneva, April, 2006. http://cdsweb.cern.ch/record/962047/.

[49] The ATLAS Collaboration, G. Aad et al., *Expected Performance of the ATLAS Experiment - Detector, Trigger and Physics*, CERN-OPEN-2008-020, CERN, Geneva, December, 2008. http://cdsweb.cern.ch/record/1125884/.

[50] ATLAS Computing Group, *ATLAS Computing: Technical Design Report*, CERN-LHCC-2005-022, CERN, Geneva, 2005. http://cdsweb.cern.ch/record/837738.

[51] GEANT4 Collaboration, S. Agostinelli et al., *GEANT4: A simulation toolkit*, Nucl. Instrum. Meth. **A506** (2003) 250–303.

[52] *Atlfast homepage*, http://www.hep.ucl.ac.uk/atlas/atlfast/.

[53] T. Sjostrand, S. Mrenna, and P. Skands, *PYTHIA 6.4 physics and manual*, JHEP **05** (2006) 026, (arXiv:hep-ph/0603175).

[54] G. Corcella et al., *HERWIG 6: An event generator for hadron emission reactions with interfering gluons (including supersymmetric processes)*, JHEP **01** (2001) 010, (arXiv:hep-ph/0011363).

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

[56] M. L. Mangano et al., *ALPGEN, a generator for hard multiparton processes in hadronic collisions*, JHEP **07** (2003) 001, (arXiv:hep-ph/0206293).

[57] T. Gleisberg et al., *SHERPA 1.alpha, a proof-of-concept version*, JHEP **02** (2004) 056, (arXiv:hep-ph/0311263).

[58] F. Krauss, R. Kuhn, and G. Soff, *AMEGIC++ 1.0: A matrix element generator in C++*, JHEP **02** (2002) 044, (arXiv:hep-ph/0109036).

[59] S. Frixione and B. R. Webber, *Matching NLO QCD computations and parton shower simulations*, JHEP **06** (2002) 029, (arXiv:hep-ph/0204244).

[60] S. Jadach et al., *The tau decay library TAUOLA: Version 2.4*, Comput. Phys. Commun. **76** (1993) 361–380.

[61] S. Hassani et al., *A muon identification and combined reconstruction procedure for the ATLAS detector at the LHC using the (MUONBOY, STACO, MuTag) reconstruction packages*, Nucl. Instrum. Meth. **A572** (2007) 77–79.

[62] G. C. Blazey et al., *Run II jet physics*, (arXiv:hep-ex/0005012).

[63] M. Heldmann and D. Cavalli, *An improved tau-Identification for the ATLAS experiment*, ATLAS note, ATL-PHYS-PUB-2006-008, CERN, Geneva, December, 2005. http://cdsweb.cern.ch/record/923980/.

[64] S. Corréard et al., *b-tagging with DC1 data*, ATLAS note, ATL-PHYS-2004-006, November, 2003. http://cdsweb.cern.ch/record/686346/.

[65] V. L. Korotkikh and A. M. Snigirev, *Double parton correlations versus factorized distributions*, Phys. Lett. **B594** (2004) 171–176, (arXiv:hep-ph/0404155).

[66] N. E. Adam et al., *Report of the Working Group on Higgs Bosons for the Workshop "Physics at TeV Colliders", Les Houches, France, 11–29 June, 2007*, (arXiv:0803.1154 [hep-ph]).

[67] T. Plehn, D. L. Rainwater, and D. Zeppenfeld, *Determining the structure of Higgs couplings at the LHC*, Phys. Rev. Lett. **88** (2002) 051801, (arXiv:hep-ph/0105325).

[68] C. Ruwiedel, N. Wermes, and M. Schumacher, *Prospects for the measurement of the structure of the coupling of a Higgs boson to weak gauge bosons in weak boson fusion with the ATLAS detector*, Eur. Phys. J. **C51** (2007) 385–414.

[69] S. Kaiser et al., *Study of Track Jets in VBF $H \to WW$ Production with Pile-up*, ATLAS internal note, ATL-COM-PHYS-2008-019, CERN, Geneva, February, 2008. http://cdsweb.cern.ch/record/1092110/.

[70] A. Ahmad et al., *Search for the Standard Model Higgs boson via Vector Boson Fusion production process in the di-tau channels with ATLAS*, ATLAS internal note, ATL-COM-PHYS-2008-220, CERN, Geneva, November, 2008. http://cdsweb.cern.ch/record/1140156/.

[71] A. Höcker et al., *TMVA - Toolkit for Multivariate Data Analysis*, CERN-OPEN-2007-007, CERN, Geneva, March, 2007. http://cdsweb.cern.ch/record/1023102/.

[72] R. A. Fisher, *The use of multiple measurements in taxonomic problems*, Annals Eugen. **7** (1936) 179–188.

[73] Y. Freund and R. E. Schapire, *A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting*, Journal of Computer and System Sciences **55** (1997) 119–139.

[74] L. Breiman et al., *Classification and Regression Trees*. Wadsworth and Brooks, Monterey, CA, 1984.

[75] F. James, *Statistical methods in experimental physics*. World Scientific, 2006.

[76] K. Melnikov and F. Petriello, *W Boson Production Cross Section at the Large Hadron Collider with $O(\alpha_s^2)$ Corrections*, Phys. Rev. Lett. **96** (2006) 231803, (arXiv:hep-ph/0603182).

[77] R. Bonciani, S. Catani, M. L. Mangano, and P. Nason, *NLL resummation of the heavy-quark hadroproduction cross-section*, Nucl. Phys. **B529** (1998) 424–450, (arXiv:hep-ph/9801375).

[78] M. Schmitz, *Studie zur Bestimmung des Untergrundes aus Daten und der Higgs-Boson-Masse in Vektorbosonfusion mit mit dem ATLAS-Detektor am LHC*, diploma thesis, Universität Bonn, June, 2006. BONN-IB-2006-07.